Water Detection Framework for Industrial Electric Arc Furnaces

by

Hamzah Alshawarghi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2016

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis develops a framework for water detection in an industrial electric arc furnace. The objective of the framework is to prevent water leak furnace explosions. This framework consists of a hybrid algorithm and a fault detection method. The hybrid algorithm consists of a mechanistic model and an empirical model. The hybrid algorithm and the fault detection method developed in this work are implemented on two industrial AC electric arc furnaces. The names of the plants and details of the operations were withheld for confidentiality reasons.

The first problem treated in this work was collecting the required data. The data required for this work included EAF operational data and off-gas composition. Both melt-shops did not have off-gas analysis systems and hence an off-gas analyzer with an HMI/SCADA data collection system was installed for each furnace. EAF operational data was sent to the data HMI/SCADA collection system installed at each melt-shop. The off-gas compositions measured in both melt-shops were CO, CO_2, O_2, H_2, N_2, and H_2O. Once all required data was collected then the framework to detect water was developed. In order to test the water detection framework developed in this work, industrial trials were completed where water was intentionally added into the furnace by increasing the electrode spray water flow rate.

The mechanistic model is completed by performing a mass balance on the furnace. The model provides a boundary with upper and lower limits in real-time of the expected EAF off-gas water vapor leaving the furnace. The mechanistic model of the hybrid algorithm has shown in both industrial EAFs that it provides a valuable on-line monitoring tool to the operator on what boundary to expect for the off-gas water vapor.

There are many input variables and historical heats in an EAF operation; hence before building the empirical predictive component of the hybrid algorithm, heats selection model and input variables selection model are constructed based on latent variable methods. The outcome of the heats selection model is heats with normal

operation. The outcome of the input variables selection model is variables that are highly correlated with the off-gas water vapor. Once the heats and the input variables are selected, then the empirical predictive models are developed.

Empirical predictive models investigated in this work are: statistical fingerprinting, artificial neural network, and multiway projection to latent structures. Robustness issues with each method are discussed and a performance comparison between the methods is presented. The last section of this thesis proposes a novel approach to detecting water leaks in the furnace.

Acknowledgments

I would like to thank my supervisors, Professor Ali Elkamel and Professor Behzad Moshiri, for their encouragement and support throughout my thesis work. I would like to thank my supervisory committee, Professor Othman Basir, Professor Yuning Li, and Professor Xianshe Feng for providing valuable information during the initial stages of this work.

I would like to thank Tenova Goodfellow Inc. for providing all the data, time, and support needed to complete this work, and I would like to express my appreciation to all my friends working at Tenova Goodfellow Inc. for their friendship and helpful discussions over the years.

Finally, special thanks to my family especially my parents, Mudar and Mouna, for their continuous prayers and endless love. I would also like to thank my grandfather, Taha, for his continuous encouragement. I would like to express my sincere appreciation to my wife Arwa for all her support throughout my thesis.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

a^l - Artificial neural network activation vector

AC - Alternating current

ANN - Artificial neural network

b^l - Artificial neural network bias vector in the l^{th} layer

C (w,b) - Artificial neural network quadratic cost function

CUSUM - Cumulative sum control chart

df - Degree of freedom

DRI - Direct reduced iron

E - Residuals

EAF - Electric arc furnace

HMI - Human machine interface

kWh - Kilo watts hour

MAD - Median absolute deviation

MPCA - Multiway principal components analysis

MPLS - Multiway projection to latent structures

MVA - Mega volt amperes

$\dot{n}_{carbon}$ - Total carbon molar flow rate (Mol/s)

$\dot{n}_{CC}$ - Charged carbon molar flow rate (Mol/s)

$\dot{n}_{CH4}$ – Total methane molar flow rate (Mol/s)

$\dot{n}_{dry}$ – Dry off-gas molar flow rate (Mol/s)

$\dot{n}_{H2}^{dry}$ – Dry off-gas hydrogen molar flow rate (Mol/s)

$\dot{n}_{H2}$ - Off-gas hydrogen molar flow rate (Mol/s)

$\dot{n}_{H2O}$ - Expected water vapor molar flow rate leaving the furnace (Mol/s)

$\dot{n}_{H2OElectrode}$ - Electrodes water cooling molar flow rate (Mol/s)

$\dot{n}_{H2OInjection}$ - Water Injection molar flow rate (Mol/s)

$\dot{n}_{IC}$ - Total injected carbon molar flow rate (Mol/s)

$\dot{n}_{N2}^{dry}$ – Dry off-gas nitrogen molar flow rate (Mol/s)

$\dot{n}_{O2}^{dry}$ - Dry off-gas oxygen molar flow rate (Mol/s)

$\dot{n}_{TotalH2}$ - Total Hydrogen (H_2 + H_2O) molar flow rate leaving the furnace (Mol/s)

$\dot{n}_{Wet}$ - Off-gas wet molar flow rate (Mol/s)

OPC - Open platform communications

PLC - Programmable logic controller

RMSEP - Roots mean squared error of prediction

s_a^2 - Variance of each latent component "a"

SCADA - Supervisory control and data acquisition

SPE - Squared prediction error

SSX_a - Sum of squares in the X matrix after "a" components

SVM - Support vector machine

T, U - Latent variables method score vectors

VIP - Variables importance to prediction

w^l - Artificial neural network weight matrix connecting to the l^{th} layer of neurons

W, C, P - Latent variable method loading vectors

y_{CO2}^{dry} - Dry off-gas carbon dioxide molar percent

y_{CO}^{dry} - Dry off-gas carbon monoxide molar percent

y_{H2}^{dry} - Dry off-gas hydrogen molar percent

y_{O2}^{dry} - Dry off-gas oxygen molar percent

y_{H2O} - Molar percent of the expected water vapor leaving the furnace

z^{l} - Artificial neural network weighted input to the neurons in layer l

List of Symbols

σ - Artificial neural network activation function

$\mathcal{E}$ - Carbon injection efficiency factor (%)

δ^L - Artificial neural network intermediate error variable associated with layer L

Chapter 1

Introduction

Electric Arc Furnaces (EAFs) are used in the steel industry to produce liquid steel. Approximately 26 percent of the global steel produced today is produced by the EAF process (worldsteel.org). EAFs are used to convert different iron materials such as recycled scrap or direct reduced iron (DRI) to liquid steel. The feed iron material is melted using electrical and chemical energy in the furnace, and the molten steel chemistry is adjusted to obtain the desired grade specifications. The electric arc furnace is a batch process producing batches of liquid steel known as heats. The electrical energy is added to the furnace through electrodes in the form of electric arc, and the chemical energy is added using a fuel source such as methane, oxygen, and carbon. Typical heats in electric arc furnaces vary greatly because of the different operating conditions, but modern operations aim for a heat cycle less than one hour with electric energy consumption in the range of 380-400 kWh/ton (Jones et al., 2005).

During the last decade there have been significant advances in the EAF technology that focused on increasing productivity leading to lower cost steel production. However, recently, due to the severe consequences of furnace explosions caused by water leaks inside the furnace there has been a growing demand for safety. EAF side walls and roof are typically water cooled due to the high temperature inside the furnace, and water leaks typically occur from those panels. Water leaks have historically posed serious safety concerns for every steel plant. Hence, there is an industrial need for an effective water leak detection methodology.

1.1 EAF Process Description

The electric arc furnace heat sequence consists of the following steps: grade selection, bucket preparation, furnace charging, melting, refining, de-slagging, tapping, and furnace turn-around. The heat steps are discussed in more details by Fruehan (1998) and Jones et al. (2005). The primary raw material used for EAF steelmaking is scrap. Scrap is a valuable commodity, and it comes from three different sources: obsolete scrap such as demolished buildings, industrial scrap, and scrap produced during the steelmaking process. Scrap varies in chemical composition and it can contain contaminants that are undesirable for steelmaking such as copper. Steelmaking facilities that produce higher quality products typically use cleaner iron raw material such as DRI, which contains low contaminants.

The first step in a heat is to select the steel grade to be produced; next the scrap quality in the bucket is prepared based on the chosen steel grade to ensure that the grade specifications are met at the end of the heat. The second step is to prepare the scrap bucket. The operator layers the scrap in the bucket according to the size and density of the scrap so that the molten steel is formed faster in the furnace. Moreover, lime and carbon can be added to the bucket with the scrap, or they can be injected into the furnace during the heat.

The third step is charging the buckets into the furnace, where the roof and the electrode are raised and moved to the sides to allow the crane to charge the scrap bucket into the furnace. Once the operator finishes charging the scrap, the roof and the electrodes swing back and are lowered to start the electrical arc. If the steelmaking facility uses DRI as iron raw material, then typically DRI is continuously fed through the roof of the furnace during the heat. Modern scrap furnaces aim to operate with two or three charge buckets of scrap, because charging is a dead-time where the furnace is not melting, and also there are radiation losses every time the roof opens.

The fourth step is melting which is the core in EAF operations; modern EAF designs maximize the melting efficiency of the furnace. Melting is accomplished by supplying electrical energy and chemical energy to the furnace. The electrodes are used in the furnace to supply the electrical energy, where in the beginning of the heat, an intermediate voltage tap is used to allow the electrodes to bore into the scrap. Once enough liquid is formed, then a high voltage tap (Long arc) is selected. A long arc allows more energy to be transferred to the scrap through the radiation of the arc than a short arc. Moreover, at the start of melting the arc is unstable. However, once a molten bath forms, the arc becomes stable and the energy input to the steel bath increases. Chemical energy during the melting period of the heat is supplied by different sources such as conventional burners and oxygen lances. A conventional burner burns fuel such as methane using oxygen to generate chemical energy to melt the scrap near the burners, where oxygen lances inject oxygen directly into the bath. This injected oxygen reacts with components in the steel bath such as aluminum, silicon, carbon, and iron. These oxidation reactions are exothermic, and hence they supply additional chemical energy to heat the steel bath. The metallic oxides formed are removed from the steel bath into the slag layer. The charging process is repeated once enough scrap has been melted to accommodate the subsequent bucket. Once the final scrap bucket is charged and melted, the formation of a foamy slag is critical to bury the arc and protect the furnace sidewalls. The foamy slag is formed by injecting carbon and oxygen which forms CO bubbles in the slag. Moreover, once all scrap is melted and flat bath conditions are reached, a shorter electrical arc is used to minimize exposing the furnace sidewalls to the arc radiations.

The refining phase of the heat starts when flat bath conditions are reached. The operator's first objective is to inject oxygen to lower bath carbon, aluminum, silicon, and manganese contents to the desired level for tapping, where oxygen reacts with these elements to form metallic oxides that float out of the steel bath and into the slag layer. The operator's second objective during refining is to increase bath temperature using electrical energy to the desired tapping temperature. The de-slagging phase is then carried out to remove the slag that accumulated in the furnace during refining. Once the desired steel grade composition and temperature are achieved in the furnace, tapping is carried

out which is to discharge the steel into a ladle to be transferred to the next operation. The last step of the heat is the furnace turn-around which is the period that follows tapping during which the operator inspects the furnace interior for any refractory damages or water leaks from the panels.

Table 1.1 shows a typical heat balance for a modern EAF with data taken from Jones (2014). The total theoretical energy required for the steel bath to reach tapping temperature (e.g. 1600 degree Celsius) is approximately 370 kWh/ton. However, modern EAF is only on average 60% efficient, hence 616 kWh/ton total energy must be supplied to the furnace. Table 1.1 shows that electrical energy supplies 65% and chemical energy provides the remaining 35% of the energy input to the heat. Energy losses to the furnace water cooled panels are typically 10%, approximately 20% to the off-gas, and 7% to the slag layer. Miscellaneous losses include energy losses that occur when the operator opens the slag door to de-slag.

Table 1.1: Modern EAF Energy Balance (Jones, 2014)

Input	Electrical Energy	65 %
	Burners (i.e. Fuel and Oxygen Combustion Reaction)	5 %
	Chemical Reactions (i.e. Oxidation Reactions)	30 %
	Total	100 %
Output	Steel Bath	60 %
	Slag Layer	7 %
	Panels and Roof Cooling Water Losses	10 %
	Off-gas Losses	20 %
	Miscellaneous Losses	3 %
	Total	100 %

There are two components to the energy losses through the off-gas: sensible energy losses; and the chemical energy losses. The sensible energy loss to the off-gas is a function of the off-gas temperature and the heat capacity of the constituents of the off-gas. The off-gas consists primarily of oxygen (O_2), carbon dioxide (CO_2), carbon monoxide (CO), hydrogen (H_2), nitrogen (N_2), and water vapor (H_2O). Chemical energy losses are calculated as the potential energy that would have been recovered if carbon monoxide (CO) and hydrogen (H_2) are combusted in the furnace. Furthermore, because

in many EAF melt-shops this chemical energy is combusted after the off-gases leave the furnace in the off-gas extraction system, this operating practice is inefficient; because this energy is not used within the furnace to melt the scrap and heat the steel to the desired temperature. Alshawarghi et al. (2015) shows an example of the economic benefits achieved when the chemical energy is recovered in the furnace at an EAF melt-shop at Kanto Steel in Japan.

The electric arc furnace structure consists primarily of a hearth, roof, and the shell. The hearth consists of the refractory that lines the lower bowl, which is not water cooled because the refractory material can withstand high melting temperature (above 1600 degree Celsius). The roof supports the furnace delta in the center, where one or more of the graphite electrodes enter the furnace and the roof which is typically water cooled, and the shell which consists of water cooled side walls and lower steel bowl. Typically, water leaks into the furnace can occur from the panels or the roof. In addition, the off-gas generated in the furnace during the heat is handled by the direct-furnace shell -evacuation system that provides an off-gas extraction. This off-gas extraction system provides adequate pollution control and minimizes dust build-up in the melt-shop. Figure 1.1 shows a basic schematic of an electric arc furnace:

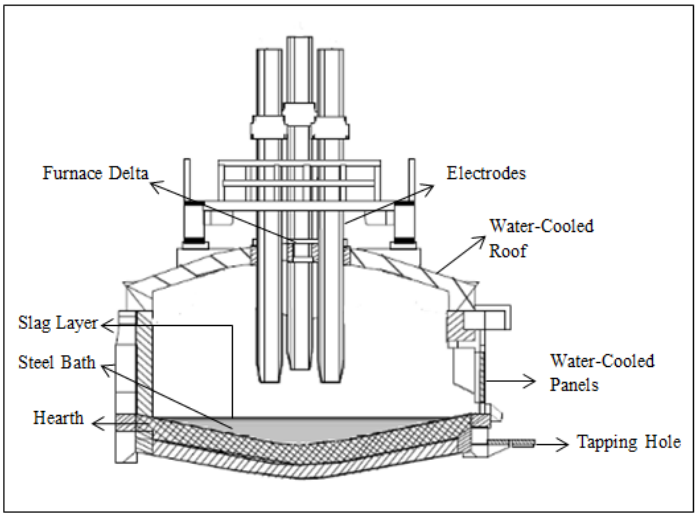

Figure 1.1: EAF Schematic (Fruehan, 1998).

1.2 Motivation and Goals

In modern EAFs, the cooling water system is an essential part of the furnace used to cool the roof and the sidewall panels. The water flows at a continuous rate of approximately 165-185 liters/min/m^2 of cooled area. The total cooling water flowrate requirement for a typical EAF ranges between 16,650 and 23,850 liters/min (Quiroga, 2013). Due to the high flowrate of water in the furnace panels, a leak in any one of the water cooled panels can quickly result in significant amounts of water in the furnace. Furthermore, if this water leaks into the EAF and comes in contact with molten steel, there is the potential for a severe explosion. There are two methods of explosions: one is a steam explosion from the mixing of water with molten steel where water trapped underneath molten material evaporates and violently expels hot material from the furnace; the other method is the dissociation of water into hydrogen gas resulting in the formation of an explosive mixture of gases that could ignite in the presence of oxygen (Zuliani et al., 2014). Regardless of the mechanism, water leaks in the EAF presents a serious and dangerous situation. Personnel safety, damaged equipment, and production losses are possible effects of water leaks in the furnace.

An example of an EAF water leak accident was the explosion that happened in ArcellorMittal Coatesville, Pennsylvania on May 26, 2007, where three operators were hospitalized and one of them died the following day. The furnace was down for days to complete the furnace repair and the accident investigation (OSHA.com). Another example of a recent EAF water leak accident was the explosion at Carbide Industries in Louisville, Kentucky on March 21, 2011, that killed two workers and injured two others, and the furnace was down for days to repair the furnace and complete the investigation (OSHA.com). These water leak explosions typically result in fatalities. The frequency of water leak accidents may vary from once every few months to once every few years. Reasons for the frequency difference include safety standards implemented in the melt-shop and technologies such as flow meters installed on the water cooled panels. The benefits for accurately detecting water leaks can minimize the risk of such furnace explosions. Most furnaces today rely on water flow meter system in the panels to

indicate to the operator if there is a water leak from the panels into the furnace; however, the measurement noise limits its effectiveness to larger water leaks in the range of 90-180 liters/min (Zuliani et al., 2014).

The main focus of this work involves the application of mathematical modeling techniques to develop a framework to detect water leaks in an industrial EAF. The water detection framework can be divided into two main sections, with the second section building on the development of the first section. The objective of the first section is to develop a model that can calculate expected off-gas water vapor leaving the furnace; the objective of the second section is to develop a fault detection algorithm to indicate if there is a water leak into the furnace. The nature of the model developed in the first section is hybrid in nature, hence it is a combination of mechanistic and empirical models. The first section can be divided into three main sub-sections: construct an EAF expected off-gas water vapor boundary limits; develop an EAF input variables and heats selection models; and evaluate different empirical methods to predict EAF off-gas water vapor. The outcome of the mechanistic component of the hybrid model is the boundary limit. The mechanistic component is a simplified mass balance that takes into account the process inputs contributing to the formation of the off-gas water vapor inside the furnace and provides the operator with boundary limits of the expected water vapor leaving the furnace. The empirical component of the hybrid model consists of a variables and heats selection models that are based on latent variable methods and a water vapor predictive model used to predict water vapor leaving the furnace. Therefore, the outcome of the first section is a boundary and a prediction for the expected off-gas water vapor.

A critical input required by the water leak detection method developed in this work is the EAF off-gas analysis. Hence, part of this work is to install an off-gas analyzer in each of the two melt-shops to measure the off-gas composition. The calculated EAF water vapor by the hybrid model along with the off-gas water vapor measured by the analyzer are used in the fault detection algorithm developed in this work to provide an alarm to the operator if there is a potential water leak situation in the furnace. The water detection method developed in this work must be sufficiently detailed so it can detect

small water leaks. A minimum significant water leak in the furnace is normally 30 liters per minute, and the objective of this work is to develop a method capable of detecting such leaks and hence prevent water leak explosions. The water detection framework developed in this work is implemented on two industrial EAFs.

1.3 Main Contributions

The complexity of the electric arc furnace process has hindered the development of practical models that can be used to improve the overall furnace operation. There were few attempts in the past (e.g. Logar et al., 2012; and Macrosty and Swartz, 2005) to develop mechanistic models that can be used for EAF control and optimization. However, the modeling approach followed in this work is different because it is hybrid in nature and because the objective here is to calculate the off-gas water vapor which is not addressed in most of the previous developed EAF models. Figure 1.2 summarizes the flow chart of the water detection framework developed in this work:

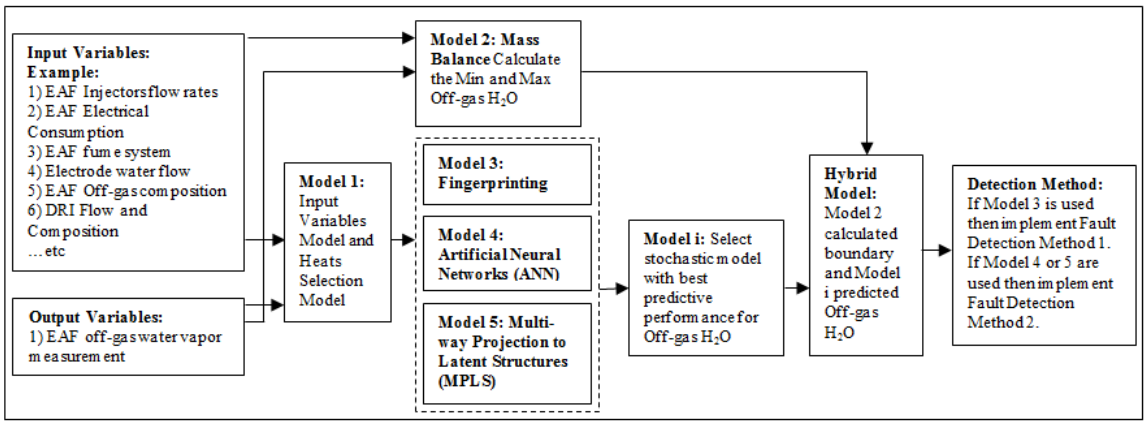

Figure 1.2: Water Detection Framework Developed in this Work.

The work presented in this thesis attempts to address the issue of water leaks in the furnace by providing a framework for a hybrid model to predict off-gas water vapor leaving the furnace and then to develop a fault detection method that can be used with the hybrid model. Different empirical methods are going to be compared in this work to

determine the one with the best performance for the two industrial EAFs. Some of the key features of this work are the following:

- Mechanistic model to accurately calculate the boundaries of the expected water measurement. This step is required as a check for the calculated off-gas water leaving the furnace.
- Input variable selection model that is capable of selecting the variables that are highly correlated with the furnace off-gas water vapor measurement. In a typical EAF operation there are normally more than 50 input variables, and hence it is essential to build a variables selection model. Moreover, a heat selection model is also developed in this work that is capable of selecting normal operating heats and excluding outlier heats.
- Three empirical models (Statistical Fingerprinting, Artificial Neural Network, and Multiway Projection to Latent Structures) are developed to predict the expected off-gas water vapor. These prediction methods capabilities are compared in order to choose the method with optimum performance for this application.
- A novel detection method is developed based on a comparison between measured and predicted off-gas water vapor.
- This approach is implemented on two different industrial electric arc furnaces. The first one is a 100% scrap furnace, and the second furnace is mostly a DRI furnace. The reason that these two EAF melt-shops are selected for this work is because they represent the two common operation modes.

1.4 Thesis Overview

Chapter 2 - Review of EAF Modeling and Detection Methods

This chapter covers previous and current work related to the modeling of the electric arc furnace, and different fault detection approaches are reviewed. This chapter discusses the strengths and weaknesses of each approach and the advantages of the hybrid approach in the electric arc furnace modeling.

Chapter 3 – Industrial Experiments and Data Collection

This chapter discusses the two industrial EAFs that are used in this work in detail, as well as the equipment that is used to measure the off-gas composition, then a description of the experiments that are performed in each of the two melt-shops.

Chapter 4 – Boundary Formulation Model, Heats Selection Models, and Input Variables Selection Models

The boundary formulation model to estimate the boundaries of the expected water vapor leaving the furnace is developed in the first part of the chapter, and then the heat and input variables selection models are developed in the second part of the chapter. The heats and inputs selection models are developed based on latent variable methods (MPCA and MPLS).

Chapter 5 - Predictive Models and Fault Detection Methodology

Three different empirical methods are explored: statistical fingerprinting, multiway projection to latent structures (linear method), and artificial neural network (non-linear method). The robustness of the methods is compared in this chapter. Finally, this chapter discusses the novel fault detection approach developed to detect minimum significant water leaks in the furnace. The approach is tested on the two industrial furnaces.

Chapter 6 - Conclusions and Recommendations

A summary of this work is presented in this chapter with highlights on the main results. Recommendations for future work are also discussed in this chapter.

Chapter 2

Review of EAF Modeling and Detection Methods

Developing an accurate and comprehensive mechanistic EAF model that characterize all the mass and energy balances is difficult due to the lack of online sensors for measuring some of the primary variables (e.g. scrap composition or DRI composition in real-time) required for the models and the presence of nonlinearities due to the batch nature of the process. Moreover, to the author's knowledge there has not been a published EAF method capable of accurately detecting water leaks in the furnace because primarily most of these models do not address EAF off-gas water vapor.

Currently, EAF operators use the flow of water inside the EAF water cooled panels to provide an indication if there is water leak from the panels inside the furnace. However, this method is prone to high signal noise and poor response time (Zuliani et al. 2014). In addition, there are off-gas systems suppliers (e.g. Grieshaber, K. and F. Martinez, 2015) that have developed systems to measure the off-gas water vapor, and they use the changes due to unexpected additional water sources to detect for leaks. However, due to the high variability in the EAF process, it may be insufficient to use only the measured off-gas water vapor and compare it against a static set-point determined from historical heats to detect water leaks. Variability in the EAF process include charging wet scrap or scrap that contains high hydrocarbons (e.g. turnings), or variation in humidity from air ingress entering from a slag door into the furnace because operators in different shifts tend to open the slag door at different times in a heat. The solution proposed in this work includes a reliable measurement of off-gas composition

and a dynamic calculation of water vapor based on process variables in real-time to accurately detect potential water leaks and minimize false alarm rates.

There have been efforts to develop mechanistic models for the electric arc furnaces (e.g. Logar et al. (2012), MacRosty and Swartz (2005), and Bekker et al. (1999)) to be used for different applications such as simulations, control, and optimization.

2.1 EAF Modeling Approaches

Logar et al. (2012) presented an approach to the mathematical modeling of an AC electric arc furnace. The objective of the EAF model was to be used for control, optimization of the energy consumption, and to develop a simulator to train operators. The authors considered the furnace as a combination of electrical, hydraulic, chemical, thermal, and mass sub-processes. Each sub-process was modeled and all the models together made-up the EAF model. The electrical model was described as a 3-phase, non-linear electric circuit. The electric arcs were the non-linearity in the electrical model. The EAF electric-circuit model was represented as a three-phase, star AC circuit. The authors used a form of randomness to the mathematical equations in the electrical model to minimize the difference between the measured and the simulated data.

The second sub-process model developed was the electrode control (hydraulic sub-process). The purpose of the electrode control system was to control the resistance and the power of each arc, where the power of the arc was related to the arc length. The input to the electrode control model developed by the authors was the controller outputs in %, and the outputs of the model were the arc resistance, the scrap height, and the arc length. The third sub-process modeled was the heat and mass processes in the EAF. The author divided the EAF into different zones, where the components of each zone possessed equal thermal, chemical, and physical properties. This assumption was clearly inaccurate because the properties of the components vary within the zone, but the authors had to make this assumption to simplify the calculation. The EAF zones considered by the authors were solid scrap zone, liquid scrap zone, solid slag zone, liquid slag zone, gas

zone, roof zone, and wall zone. The authors considered all modes of heat transfer (i.e. conduction, convention, and radiation) in the thermal model. A consequence of the heat-balance equations was the temperature calculation in each EAF zone. Also, the mass balance equations for each of the EAF zones were constructed to track materials in each zone. Finally, the fourth sub-process modelled by Logar et al. (2012) was the chemical processes in the EAF. This model included common chemical reactions that occur during a heat, such as the oxidation and reduction of iron, carbon, silicon, manganese, chromium, and phosphorous. The model included electrodes oxidation, the oxidation of combustible materials, the oxygen burners, and the slag foaming processes. The chemical model assumed that all the chemical reactions, except CO post-combustion and CH_4 oxidation, occurred in the liquid metal and slag zones. The off-gas zone compositions considered in the EAF model were N_2, O_2, CO, CO_2, and CH_4. H_2O was ignored. Therefore, this model was incapable of detecting water leaks in the furnace. The EAF model developed was based on the 80 MVA AC furnace installed in Slovenia.

MacRosty and Swartz (2005) developed a detailed model of the EAF based on first principles. The objective of the model was to use it within an optimization framework. The EAF was modelled as a system of four equilibrium zones: solid scrap zone, molten metal zone, slag metal interaction zone, and gas zone. The solid scrap zone was modelled as a mass of scrap melting according to the heat transferred from the molten steel, off-gas, and the electrical arc. The molten metal zone was modelled as scrap in the solid scrap zone melted, they entered this zone and then they would leave to enter the slag metal zone. The slag metal zone included the slag constituents (i.e. metallic oxides except CO). The species included in the gas zone were: CO, CO_2, O_2, CH_4, H_2, H_2O, N_2, and C_9H_{20}. The EAF model developed also included a heat model that considered the radiation and convective heat transfer between different EAF zones, the furnace components, and the electrical arc. Since radiation was an important mechanism of heat transfer in the EAF, it was important to predict the contribution of radiation in the furnace. The objective of this component in the heat model was to find the radiation between the different surfaces in the furnace based on their surface temperature, emissivity, and surface area. In the heat model, the authors assumed that an initial cone-

frustum shaped void was melted into the scrap by the electrodes. A consequence of the heat model was the temperature calculation in each EAF zone. Also, the mass balance equations for each of the EAF zones were constructed to track materials in each zone. Although this work included H_2 and H_2O in the gas model, MacRosty and Swartz (2005) made a few assumptions in order to develop the model that might have affected the overall accuracy of the model such as: each zone consisted of a unique composition and conditions; injected carbon went into the solution; and no reactions in the molten metal zone because of the absence of O_2 in the molten-metal zone. Furthermore, not addressing issues on the radiative and conductive heat transfer between the steel, the slag and the gas zones, and the CO post combustion might affect the overall accuracy of the model.

Modigell et al. (2001) developed an EAF model. The objective of the model was to be used as a simulation tool. The model included four reaction zones that were assumed to be in a state of chemical equilibrium. The flow of material between reaction zones was directed by concentration gradients and mass transfer coefficients. However, the authors did not disclose enough details about the model.

Bekker et al. (1999) developed an EAF model from first principles of thermo-chemistry for the purpose of control system design. Due to the complex nature of the EAF process, the authors used empirical relationships and assumptions to simplify EAF mechanisms that were not measured. The authors assumed that the radiative energy from the arc and the energy from exothermic reactions were only added to the liquid phase of the heat, and then that the liquid phase transferred that energy to the solid phase by conduction. The authors also assumed equivalent temperature between the liquid metal, molten slag, and gas phases of the heat. Liquid steel temperature increased by the chemical and electrical energy. The energy available to melt the steel was assumed to be equivalent to the ratio of scrap temperature to the molten temperature, and the remaining energy was used to heat the solid scrap. Reactions considered by Bekker et al. (1999) in his model were oxidations of Fe, C, and Si and reduction of FeO. Moreover, the authors assumed that the oxygen injected into the furnace reacts with Fe, C, and Si. The only gas-phase elements were CO, CO_2 and N_2. All O_2 react on entry, and

H_2O was ignored. Therefore, the EAF model developed by Bekker et al. (1999) was incapable of detecting water leaks in the furnace.

Matson et al. (1999) developed a model that approximated the furnace as three separate phases: bath, slag, and gas reactors. The authors assumed chemical equilibrium in the individual phases and transport limitations between each phase to compute the rate of reaction. A dynamic elemental balance was used in each phase to track the flow of the components. Gibbs free energy minimization was implemented in the equilibrium algorithm. The chemical equilibrium problem was solved via a subroutine. Mass transfer between the phases was modeled as diffusion across a concentration gradient. The authors modeled the scrap as a group of spheres. The temperature profile of the spheres was determined from the sensible heating of the spheres as a function of its radius. The surface temperature of the scrap was monitored at each time step in the algorithm. In order to reach an acceptable accuracy with this method, small discretization steps were required.

Cameron et al. (1998) developed an EAF model to be used as a dynamic simulation tool that could be used to optimize EAF operating practices. The model included four phases: metal phase, slag phase, organic solid phase, and gas phase. The model also included six interfaces between the metal, slag, gas, and carbon material. Moreover, the authors assumed chemical equilibrium at each interface. Mass transfer between the phases and interfaces was driven by a concentration gradient, with the chemical equilibrium at the interface computed by minimizing the Gibbs free energy. EAF off-gas chemistry data was used to test the model.

Due to the complex nature of the EAF process and the lack of online sensors for measuring some of the primary variables (e.g. scrap composition or DRI composition in real-time) required to develop an accurate EAF model based on first principles, most of the authors had to make assumptions to be able to develop mechanistic models, and most of the models developed did not calculate the water vapor leaving the furnace, and hence these models were incapable of detecting water leaks in the furnace. This work proposes

a hybrid modelling approach to overcome the limitations of the mechanistic approach to build a water detection algorithm, where a simplified mass balance model is used to provide upper and lower limits of the off-gas water vapor and an empirical model to calculate the water vapor leaving the furnace. The mechanistic model calculates a boundary from first principles around the water vapor prediction determined by the empirical model. This model is powerful because it provides the operator with a tool to monitor the quality of the prediction where, if there is a shift in the EAF process operation, then the empirical model should be retrained with a new heats dataset.

2.2 Fault Detection Approaches

Fault detection methods can be classified as model-based methods or data-driven methods. Model-based methods rely on fundamental understanding of the process where data-driven models rely on historical data. Model-based fault detection methods have been around for many years but their contribution to the industrial practice is limited to the cost and time required to develop accurate models for complex industrial processes. The data-driven approach requires less time and lower cost to develop. Empirical methods commonly used for data-driven fault detection approaches include artificial neural network (Chetouani, 2007), multiway principal component analysis (Nomikos and MacGregor (1994)), and Bayesian approach (Yu, 2012). Furthermore, Freeman et al. (2013) compared both approaches to a small unmanned aerial vehicles (UAV) platform.

Nomikos and MacGregor (1994) developed a multivariate statistical method for monitoring batch processes where the only information required were good historical batches. The empirical method used was a multiway principal component analysis (MPCA). MPCA was used to extract the information from the multivariate dimensions and projected them onto lower-dimensional space defined by principal components. The method used by the authors to calculate the principal component was the NIPALS algorithm. Moreover, due to the three dimensional array (batches, measurements, and time) nature of the batch data, the authors unfolded the three-dimensional array to a two-dimensional array, and then they built the MPCA model. The authors determined that

three principal components were necessary to describe adequately the normal operation of a batch. Monitoring plots generated by the MPCA method were the score plots and loading plots for the latent variables of the MPCA model. These plots included control limits corresponding to 95% and 99% confidence regions, calculated based on the reference heats. The MPCA monitoring plots were used online to monitor the progress of a new batch in real-time. The MPCA method was based on the concept that future behavior of a process was monitored by comparing it against the past when the process was performing well. Two fundamental assumptions were necessary for the MPCA method to work: the first assumption was that the reference database was representative of the process operation, and hence if something changes in the process, then a new MPCA model must be built on the new batches. The second assumption was that the fault event must be observable from the measurements collected in order for the MPCA model to detect it.

Yahya Chetouani (2007) developed an artificial neural network (ANN) approach for real-time detection of faults. This approach combined ANN and CUSUM statistical test for fault detection. The ANN model developed was a one layer perceptron network, and the process used in this work was a reactor-exchanger setup. The training algorithm used to develop the ANN model was the back-propagation training function for feed-forward networks using momentum and adaptive learning technique. The author used the CUSUM statistical test for fault detection, where this test was performed as a cumulative sum test, and where jumps in the mean occur at unknown time instants. The reactor-exchanger used to test this method was a glass-jacketed reactor with a tangential input for heat transfer fluid.

Sheibat-Othman et al. (2014) proposed a hybrid data/model-based approach for fault detection for chemical reactions. Two stirred tank jacketed chemical reactors were used. The reactor was equipped with temperature probes and the feed mixture was put on a balance to calculate the feed flowrate. Temperature sensor faults and actuator faults (capacity for heating and cooling) were used to investigate the proposed hybrid methodology. The process model developed for the system was a heat balance of a semi-

continuous stirred tank reactor and its jacket. The authors used two class support vectors machine (SVM) for the data driven model. Furthermore, it was found that it would require a great number of data to train the SVM model because the reactions in the reactor were highly nonlinear reactions. Therefore, a simplified process model was used as a starting point to develop an observer for fault isolation, and information from the SVM model was used to correct the simplified process model when no faults were detected. It was also found that the SVM model alone was sufficient to detect faults if the process dynamics were linear.

Freeman et al. (2013) designed and applied a model-based residual generation and data driven fault detection approaches to a small unmanned aerial vehicles (UAV) platform. The electric powered airplane had a 1.3 meters wingspan and a weight of 1.3 kg. The model based fault detection strategy used linear filtering methods to reject faults. Raw flight data was used to develop the data driven algorithm without knowledge of system dynamics. An $H\infty$ filter was constructed to detect aileron faults. A data driven detector was developed by processing the control error signals logged from the flight data and consequently to create an error score related to the probability of a fault. Both approaches successfully detected different aileron faults during maneuvers and in the presence of environmental disturbances. However, the performance of the data driven detector suffered in the linear simulations with high model uncertainty and did not always detect faults. The system knowledge built in the model-based design allowed for better performance.

The next chapter gives an overview of the two industrial furnaces selected to develop and test the water detection framework developed in this work. In addition, the chapter also discusses the off-gas analyzer equipment installed at both melt-shops to measure EAF off-gas composition which was required to build the water detection framework.

Chapter 3

Industrial Experiments and Data Collection

An important measurement required for this work was the furnace off-gas composition (i.e. CO, CO_2, H_2, O_2, N_2, and H_2O); therefore, proprietary off-gas analyzer was installed for each furnace as a part of this work.

3.1 Industrial EAFs Description

The water detection method developed in this work was implemented on two different AC industrial electric arc furnaces. The first EAF was a scrap charging furnace, and the second one was mostly a DRI furnace. Table (3.1) shows the overview of the two industrial furnaces:

Table 3.1: Overview of the Two Industrial Electric Arc Furnaces (2014)

Parameter	EAF 1	EAF 2
Annual EAF Production	800,000 Tons	600,000 Tons
Total Metallic Charge Weight	190 Tons/heat	90 Tons/heat
Total Tapped Liquid Weight	170 Tons/heat	82 Tons/heat
Feed Material	Scrap	Mostly DRI
Number of Buckets per Heat	3 or 4	0 and sometimes 1 scrap bucket with DRI
EAF Injectors	4 coherent burners (Natural gas and Oxygen), 2 Carbon Injectors, and 1 Lime Injector. Use of charged carbon (Coal).	3 coherent burners (Natural gas and Oxygen), 3 Carbon Injectors, and 1 Lime Injector.
Heat tap to tap time	95 min	45 min
Power on time	70 min	37 min

EAF 1 produced approximately 800,000 tons of raw steel in 2014. The plant delivered a wide range of products for construction, aircraft and aerospace, energy, and heavy equipment. The scrap charge weight was 190 tons per heat, and the furnace tapped 170 tons per heat. This furnace had four coherent burners, two carbon powder injectors, and one lime injector. Carbon powder was also charged with the scrap. The heat total cycle was 95 minutes, and the power on time was 70 minutes. EAF 2 produced, in 2014, approximately 600,000 tons of tubes and delivered a wide range of products for the energy market and the automotive sector. The DRI total weight was 90 tons per heat, and the furnace tapped 82 tons per heat. This furnace had three coherent burners, three carbon powder injectors, and one lime injector. The heat total cycle was 45 minutes, and the power on time was 37 minutes. The names of the plants and details of the operations were withheld for confidentiality reasons.

The coherent burners in both furnaces were wall-mounted injectors with a nozzle that delivered oxygen at supersonic velocity into the steel bath. The coherent burners had ports for main oxygen, secondary oxygen, and natural gas. The main oxygen port was used for the oxygen jet. The injector nozzle kept the jet of oxygen coherent. The jet of oxygen was kept coherent by forming a shroud flame (envelope) around it through the combustion of the secondary oxygen with natural gas. The injector delivered a specific amount of oxygen to the steel bath with minimal splash. When the coherent jet of oxygen produced by the nozzle impinged on the steel bath, the concentrated momentum of the oxygen jet dissipated in the steel as fine bubbles, providing deep penetration and slag-metal mixing. The nozzle also operated as a conventional sidewall burner during the first few minutes after scrap was charged to melt scrap faster which improved furnace productivity and decreased electrical power consumption. The carbon and lime injectors were wall-mounted pipes. The flow rate set-points for the coherent burners, carbon injector, and lime injector were predetermined based on an optimized chemical profile.

3.2 EAF Off-gas Analyzer System

The proprietary analyzer measures and analyzes real-time off-gas chemistry from the forth-hole of the Electric Arc Furnace (EAF). Figure 3.1 below gives a basic overview of the off-gas system:

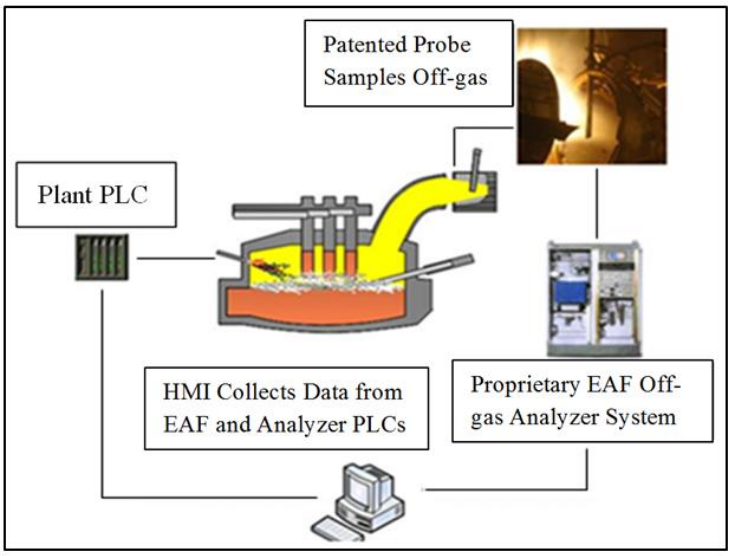

Figure 3.1: Schematics of the Proprietary Continuous EAF Off-gas System (Alshawarghi et al. (2015)).

The off-gas system extracts the off-gas sample, analyzes the sample, and sends the off-gas values to the HMI/SCADA computer in the control room. The HMI/SCADA also collects data from the plant PLC. The water detection model set-points are then sent from the HMI/SCADA computer to the Plant PLC. Figure 3.1 and 3.2 show the major components of the proprietary off-gas system installed at both melt-shops. The major components of the system are a patented water cooled probe, a heated sample line, a multi-gas analyzer with a conditioning cabinet, the HMI/SCADA computer, analyzer control system, and the analyzer room. The EAF off-gas sample is collected from the water cooled off-gas sample probe mounted in the water cooled D1-duct on the furnace elbow. The sample gasses are drawn under vacuum through the water cooled probe and into the filter inside the probe where the off-gas is filtered from dust and then through the heated sample line to the conditioning sub-system inside the analyzer cabinet. The

reason that the gases are transported through a heated line is to maintain sample integrity until it reaches the analyzer cabinet. The conditioning and analysis sub-systems are installed inside the analyzer cabinet which is located in an analyzer room built specifically for this project. The analyzer room is built approximately 20 meters away from the furnace in a clean and safe area. The reason the room is kept in close proximity to the furnace is to minimize delays in sample extraction and analysis. The sample delay at both installations is kept below 20 seconds from the time the sample is drawn from the furnace to the time the analysis is displayed on the HMI in the control room.

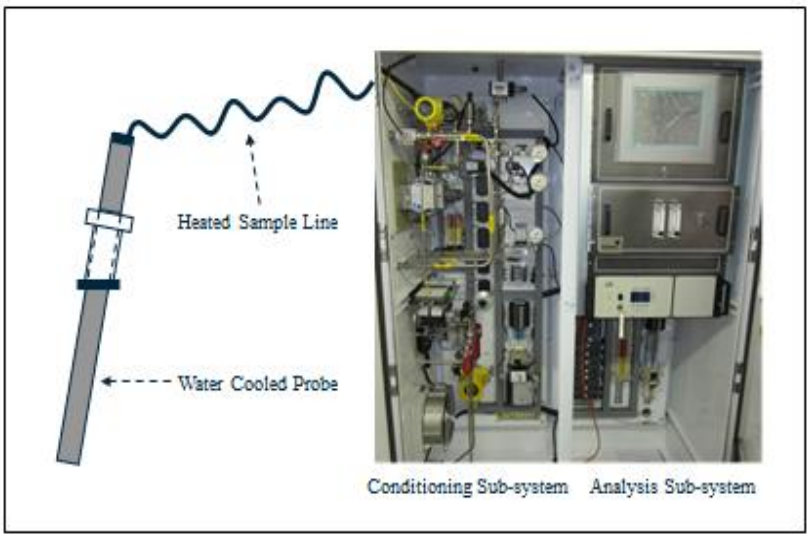

Figure 3.2: EAF Off-Gas Analyzer System.

3.3 EAF Off-gas Sample Conditioning

The sample conditioning sub-system draws the off-gas sample from the process through the patented water-cooled probe via a vacuum pump located in the conditioning sub-system. The pump draws the off-gas sample continuously under a vacuum pressure. The sample then goes under positive pressure once it passes through the head of the vacuum pump. The conditioning sub-system then removes the dust before passing a small portion of the off-gas sample to the water vapor measurement detection sub-system. The moisture in the off-gas sample is then removed, and a small portion of the off-gas sample

is sent to the analysis sub-system. The sample conditioning sub-system generates by-products from its operation that are discharged from the analyzer drain and off-gas vent ports. Harmful gases are vented and drained from the analyzer, and these gases are to be transported to a safe location outside of the analyzer room. During the normal course of operation, the probe plugs as dust accumulates on the inside walls of the probe. In order to ensure continuous analysis, the probe, the heated sample line, and the filters inside the conditioning cabinet are purged during the periods of time when off-gas analysis is not required, such as when the furnace is being charged with scrap or during tapping.

3.4 EAF Off-gas Sample Analysis and Data Collection

The off-gas analysis sub-system analyzes the dirt and moisture free off-gas sample supplied by the conditioning sub-system. The measurement accuracy of the analyzer is maintained through regular calibration. Hydrogen is detected by thermal-conductivity cell. The oxygen detector is an electrochemical cell. The two gas species are measured in a dual microprocessor controlled non-dispersive infra-red cell. The off-gas water vapor is measured in a laser spectrometer system. The multi-gas analyzer measures levels of O_2, CO, CO_2, H_2, and H_2O in molar %, and then the analyzer PLC sends the data to the off-gas HMI/SCADA computer. The off-gas HMI/SCADA computer is interfaced with the plant's network to receive and log off-gas analysis from the analyzer PLC and receive plant process data from the plant PLC. Figure 3.3 shows the configuration of the off-gas analyzer HMI/SCADA and the analyzer PLC computer on the plant's network.

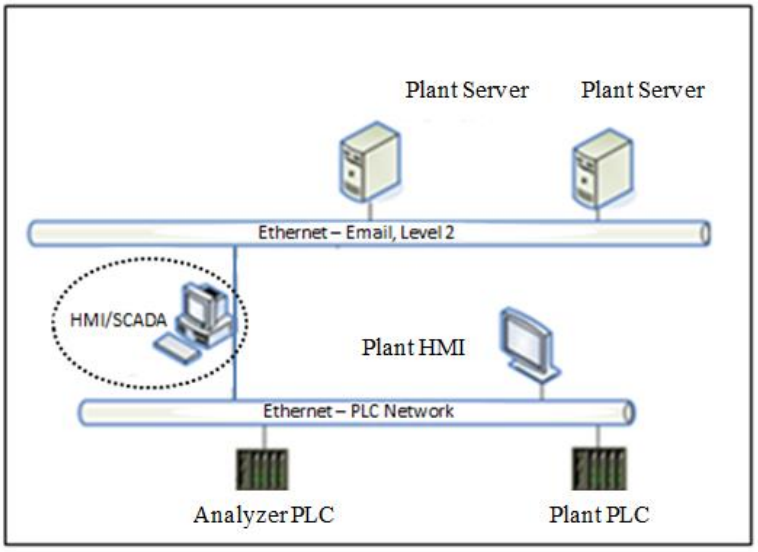

Figure 3.3: Configuration of the Off-gas HMI/SCADA Computer and Off-gas Analyzer PLC on the Plant's Network (Nikkanen et al. (2012)).

The HMI/SCADA platform installed in both facilities is GE iFix. HMI and SCADA refer to industrial control systems. HMI is a component of SCADA. SCADA is the supervisory control and data acquisition, and it allows for a direct control and communication with PLCs, data storage systems, and process control systems. HMI stands for Human Machine Interface, and it allows for interactions between the operators and industrial equipment such as the furnace and the analyzer. The PLC is the control system, and it stands for programmable logic controllers. The analyzer PLC controls all the functions of the analyzer. All operational functions of the analyzer are controlled and monitored by the analyzer PLC. The analyzer PLC installed at both melt-shops is a Siemens PLC. The HMI/SCADA computer interacts with the plant PLC and with the analyzer PLC via an OPC server.

Off-gas systems potential damages include: probe plugging, air leakage in the sample line, general mechanicals damage, PLC-HMI communication loss, power blackout. The above mentioned possible damages have been monitored using the historical data stored in the HMI/SCADA database since the beginning of the project, and it was observed that none of these events occurred over a period of one year. In addition, some general rules have been developed for a fast check of sample reliability such as:

oxygen should be less than 25%; the sample line negative pressure should be within standard limits; passing the chemistry algorithm in the HMI/SCADA system that detects for air in-leakage, low chemistry concentration, chemistry errors, and calibration errors. All of these conditions have to be simultaneously satisfied in order to consider the sample as good chemistry. The off-gas analyzers have worked reliably in both installations for a period over one year giving good off-gas chemistry measurements.

3.4.1 EAF 1 Off-gas Chemistry

Figure 3.4 shows an example of a typical heat off-gas chemistry measured by the proprietary off-gas system at EAF 1:

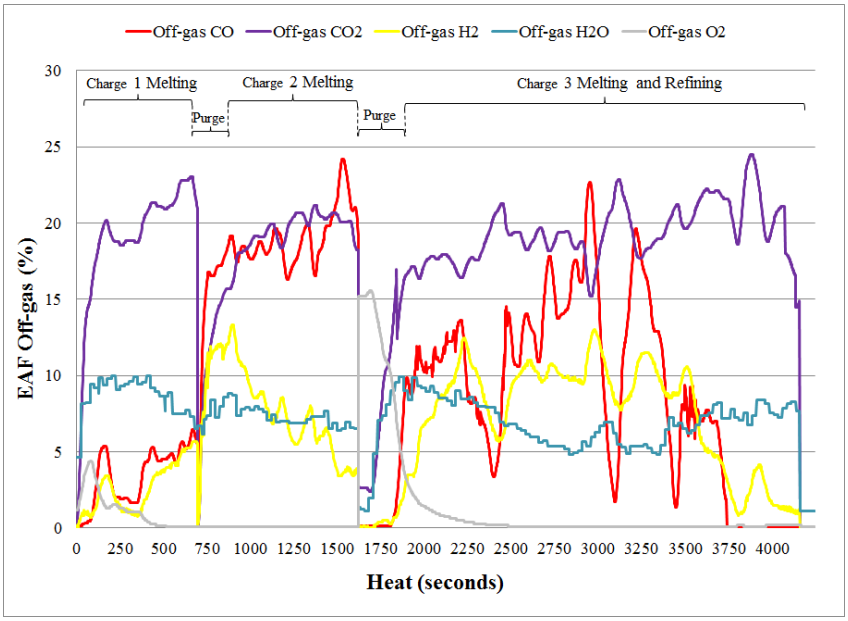

Figure 3.4: Typical Heat Off-gas Analysis for EAF 1.

The three stages of the EAF operation (i.e. Charge 1 Melting, Charge 2 Melting, Charge 3 Melting and Refining) are indicated in Figure 3.4. The probe purge is initiated during scrap charging and tapping at the end of the heat when the off-gas analysis is not required. The off-gas chemistry measured by the off-gas analyzer system indicates that the furnace freeboard operates under slightly reducing condition. This slightly reducing

condition is indicated by the medium concentration of carbon monoxide (CO – red) and hydrogen (H_2 – yellow) in this heat. The medium concentration of CO and H_2 at any point in the heat is typically 5-25%. During the first few minutes of scrap melting in all three charges, the off-gas analysis shows an increase in the amount of CO and H_2. The sources for off-gas CO in the melting phase are the combustion of natural gas with oxygen, the burning of the hydrocarbons that come with the scrap, and the burning of carbon powder charged with the scrap bucket. Sources for H_2 and H_2O are combustion of natural gas and oxygen, burning of the hydrocarbons that come with the scrap, electrode spray water, and humidity in the air drafted into the furnace from the slag door. The refining phase of the process shows an increase in CO. The evolution of CO during refining represents the de-carburization process of the steel bath where lance oxygen from the coherent burners is used to remove carbon from the steel bath. The off-gas water vapor (light blue) generally shows an increase in the beginning of each charge melting, and then it slightly decreases. This increase is due to the higher usage of natural gas in the beginning of the charge when there is a significant amount of scrap in the furnace. Natural gas produces a significant amount of water vapor in the furnace. Furthermore, towards the end of refining, water vapor increased because two of the three coherent burners were switched from the oxygen lancing mode to the burner mode because carbon in the bath was decreasing rapidly. Off-gas O_2 (gray) is typically zero during melting and refining because off-gas oxygen cannot co-exist with CO and H_2 at the high off-gas temperature. The source of CO_2 (purple) is the combustion of CO with O_2.

3.4.2 EAF 2 Off-gas Chemistry

Figure 3.5 shows an example of a typical heat off-gas chemistry measured by the proprietary off-gas system at EAF 2:

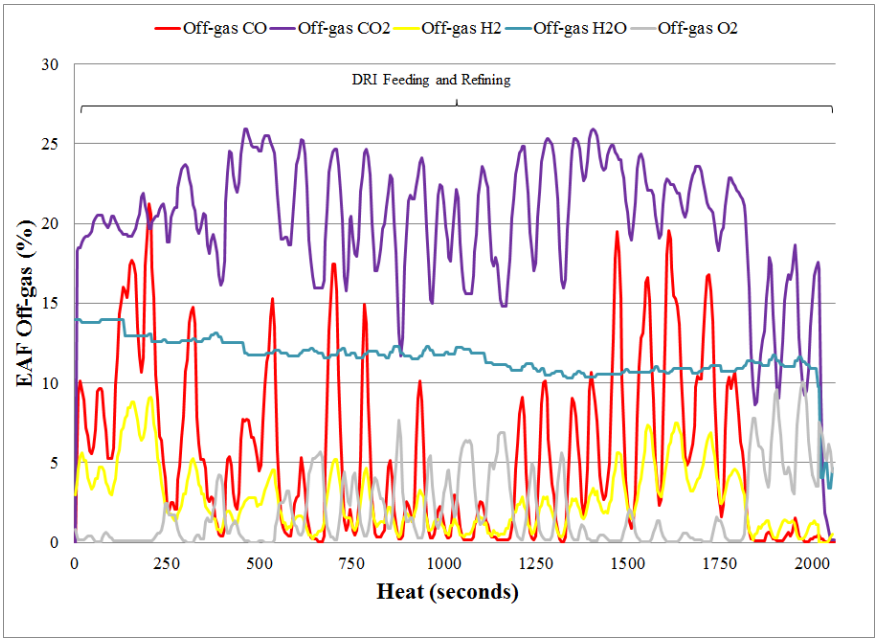

Figure 3.5: Typical Heat Off-gas Analysis for EAF 2.

DRI is fed continuously to EAF 2 during the heat. The probe purge is initiated during tapping at the end of the heat cycle. The off-gas chemistry measured by the off-gas analyzer indicates that the furnace freeboard operates under slightly reducing condition. This slightly reducing condition is indicated by the low-medium concentration of carbon monoxide (CO – red) and hydrogen (H_2 – yellow). Low-medium concentration at any point in the heat is typically 0-20%. During continuous DRI feed, the off-gas shows a medium concentration of CO due to the carbon content that comes with the DRI. The refining phase of the process shows a higher presence of CO and a dynamic behavior of CO compared to the CO during the refining period in a scrap furnace (Figure 3.4). The reason that CO has a dynamic behavior is due to the dynamics of DRI melting.

The off-gas H_2O (light blue) follows a less nonlinear behavior which is different from the off-gas H_2O behavior in EAF 1. The main reason for the different behavior is that in a continuous DRI EAF process, the coherent burners are only used in the oxygen lance mode because the steel bath is mostly in flat bath conditions, where in a scrap furnace the coherent burners are used as a conventional burners during scrap melting; and when enough scrap has melted, the coherent burner operates as an oxygen lance, hence

the amount of natural gas used in both modes varies, also wet scrap or scrap with hydrocarbons contribute to the off-gas water vapor where in a DRI furnace, the water vapor behavior is less nonlinear because the main source of water vapor is the natural gas, electrode spray water, and humidity from the air being drafted into the furnace through the slag door and hence most of these variables are consistent throughout the heat. The off-gas water vapor is higher in EAF 2 than EAF 1, and this is due to the higher natural gas consumption because the coherent burners in EAF 2 have higher firing rates; also the electrode spray water in EAF 2 has higher water flow rates. The off-gas water vapor slightly decreases towards the end of refining in Figure 3.5 because the operator stopped one of the oxygen lances because the amount of carbon in the bath was decreasing rapidly. Off-gas O_2 (gray) increased towards the end of the heat because the operator stopped all oxygen lances and opened the slag door to take the steel bath carbon and temperature sample. The source of CO_2 (purple) is the combustion of CO with O_2.

Controlled trials of injecting additional water into the two furnaces were conducted after installing the off-gas analyzer system. The purpose of these trials was to test whether the water detection framework developed in this work can detect the injected water. The method used to inject additional water into the EAF was by increasing the electrode spray water. The next step of this work is to develop the expected upper and lower boundaries of the expected water leaving the furnace and the heats and input variables selection models.

Chapter 4

Boundary Formulation Model, Heats Selection Models, and Input Variables Selection Models

The first part of this chapter discusses the simplified mass balance approach followed in this work to develop the upper and lower boundaries of the expected off-gas water vapor leaving the furnace. This simplified approach is implemented on the two industrial furnaces and results are presented. The second part of this chapter discusses the heats and variables selection models which are critical in building empirical models.

4.1 Introduction

There are four main zones in an electric arc furnace: off-gas freeboard zone, slag zone, steel bath zone, and solid scrap zone. For the purpose of this work, the freeboard zone is the only zone considered. The chemical reactions that are considered in this work are the reactions that influence the amount of water vapor in the off-gas. Equations 4.1-4.4 summarize the main reactions related to water vapor in the furnace freeboard zone:

$$\text{Fuel Combustion } CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \quad (4.1)$$

$$\text{Water Vaporization } H_2O_{(l)} \rightarrow H_2O_{(v)} \quad (4.2)$$

$$\text{Water Dissociation } 2H_2O \rightarrow 2H_2 + O_2 \quad (4.3)$$

$$\text{Water Shift Reaction } H_2O + CO \rightarrow H_2 + CO_2 \quad (4.4)$$

Small quantities of water in the EAF off-gas are typical and can result from a variety of sources such as methane combustion with oxygen from the oxy-fuel burners (Eq. 4.1), water vaporization from the water used to cool the electrodes and the electrodes' delta (Eq. 4.2). Dissociation of water vapor into hydrogen and oxygen (Eq. 4.3) is also possible at high temperatures in the electric arc furnace. In addition, water reacting with carbon monoxide at low temperatures in the electric arc furnace producing hydrogen and oxygen (Eq. 4.4) is also possible. Other reactions that can contribute to smaller amounts of off-gas water vapor include combustion of hydrocarbons on oily scrap, water vapor from humid air drawn into the furnace, and lime which can hydrate and be a source of water vapor. All of these reactions are possible to occur in the EAF; however, the extent of the reaction is determined by the EAF conditions (i.e. reducing vs. oxidizing off-gas and temperatures). Figures 3.4 and 3.5 in Chapter 3 have shown that off-gas chemistry in a typical normal heat contains off-gas water vapor and hydrogen.

A serious source of water entering the furnace results from leaks that develop in the sidewalls and roof of the EAF. When water leaks into the EAF during high temperature conditions, a portion of the water can exist as vapor (H_2O), and a portion can chemically dissociate or react to form H_2 and CO_2. Figure 4.1 shows the input and the output from the freeboard zone:

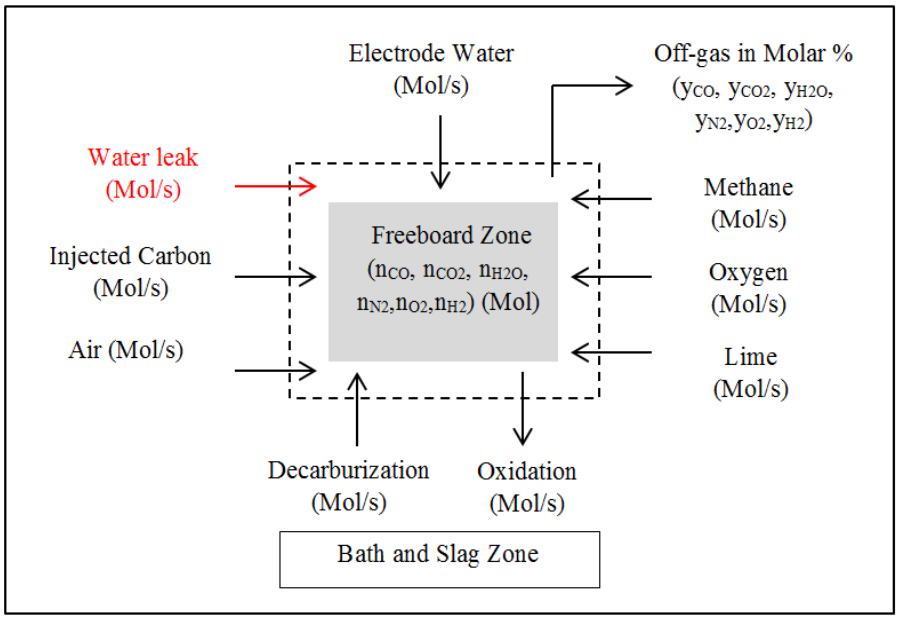

Figure 4.1: EAF Off-gas Freeboard Zone.

The next section in this chapter provides a simplified approach for calculating the boundary for the expected water vapor leaving the furnace based on normal operation. The purpose of the boundary is to provide the operator with upper and lower limits of the expected water vapor leaving the furnace.

4.2 Water Vapor Boundary Formulation

4.2.1 Method

The approach followed in this section to develop a boundary for the expected water vapor leaving the furnace is a steady state mass balance approach. Hence, this model is mechanistic in nature. The objective of this model is twofold: first to provide the operator with the upper and lower limits of the expected water vapor leaving the furnace, and then to serve as a check for the water vapor prediction which is empirical in nature. The approach developed to calculate the boundary is simplified because a rigorous approach is not required for the objective of this step. An assumption that has been made to carry out this work is that there exists a homogenous freeboard zone, both

in temperature and composition. The total carbon molar flow rate ($\dot{n}_{\text{carbon}}$) is obtained with Eq. (4.5):

$$\dot{n}_{\text{carbon}} = \dot{n}_{CH4} + \dot{n}_{CC} + (\mathcal{E} \times \dot{n}_{IC}) \tag{4.5}$$

$$\dot{n}_{x} = \frac{\dot{m}_{x}}{M_{x}} \tag{4.6}$$

$$\dot{n}_{cc} = \frac{Total\ Charged\ Carbon\ (mol)}{Power\ on\ Time\ (s)} \tag{4.7}$$

The plant measures mass flow rates in real-time for injected carbon and methane. The molar flow rate is calculated using the mass flow rate ($\dot{m}_x$) and the molecular mass (M_x) (Eq. 4.6). $\dot{n}_{\text{CH4}}$ is the total methane molar flow rate (moles/s). $\dot{n}_{\text{CC}}$ is the average charged carbon molar flow rate. EAF 1 adds charged carbon in the middle of every bucket, and hence the carbon stays longer in the melting period of the heat. The model assumes that all the charged carbon is oxidized during the heat, where the flow rate is calculated as the ratio between the total moles of charged carbon and the time in seconds when electrical power is on (Eq. 4.7). EAF 2 feeds DRI into the furnace after charging one bucket and does not use charged carbon. DRI at EAF 2 typically contains 2.5 % carbon content and hence the model replaces the charged carbon with DRI carbon. $\dot{n}_{\text{IC}}$ is the total injected carbon molar flow rate, and $\mathcal{E}$ is the carbon injection efficiency factor. The efficiency factor is the fraction of injected carbon that enters the furnace freeboard and does not leave as dust. Factors that affect the efficiency of carbon injection include carbon injectors design and time of injection in the heat. This efficiency factor can vary between 0 and 1, and for this work based on experience with the carbon injectors used at both melt-shops is assumed to be 0.5. The model assumes that there are no other sources of carbon. This assumption, while clearly not accurate, because there is in fact carbon due to electrode consumption and from the scrap material, is still valid for the purpose of calculating a boundary for the expected water vapor leaving the furnace, as it is shown later, the upper boundary is conservatively higher than the real off-gas flow and the lower boundary is conservatively lower than the real off-gas flow.

The EAF off-gas analyzer installed at both facilities removes the water vapor after measuring the water vapor content in the off-gas. Hence, the off-gas compositions (i.e. CO, CO_2, H_2, N2, and O_2) are measured on a dry basis. The off-gas dry molar flow rate ($\dot{n}_{dry}$) leaving the furnace is obtained with Eq. (4.8):

$$\dot{n}_{dry} = \frac{\dot{n}_{Carbon}}{y_{CO2}^{dry} + y_{CO}^{dry}} \tag{4.8}$$

y_{CO2}^{dry} is the dry off-gas carbon dioxide molar percent measured by the off-gas analyzer, and y_{CO}^{dry} is the dry off-gas carbon monoxide molar percent measured by the off-gas analyzer. The dry off-gas hydrogen, oxygen, and nitrogen molar flow rates are obtained with Eq. (4.9-4.11):

$$\dot{n}_{H2}^{dry} = \dot{n}_{dry} \times y_{H2}^{dry} \tag{4.9}$$

$$\dot{n}_{O2}^{dry} = \dot{n}_{dry} \times y_{O2}^{dry} \tag{4.10}$$

$$\dot{n}_{N2}^{dry} = \dot{n}_{dry} \times (100 - y_{H2}^{dry} - y_{O2}^{dry} - y_{CO2}^{dry} - y_{CO}^{dry}) \tag{4.11}$$

Where $\dot{n}_{H2}^{dry}$, $\dot{n}_{O2}^{dry}$, $\dot{n}_{N2}^{dry}$ are the molar flows (moles/sec) of hydrogen, oxygen, and nitrogen leaving the EAF, respectively. y_{H2}^{dry} is the dry off-gas hydrogen molar percent, y_{O2}^{dry} is the dry off-gas oxygen molar percent. The molar flow of the total hydrogen (*H_2 and H_2O*) expected to leave the EAF ($\dot{n}_{TotalH2}$) is obtained with Eq. 4.12

$$\dot{n}_{TotalH2} = 2 \times \dot{n}_{CH4} + \dot{n}_{H2OElectrode} + \dot{n}_{H2OInjection} \tag{4.12}$$

Eq. 4.12 includes the hydrogen from the methane, the hydrogen from the electrode water cooling, and the hydrogen due to a water injection into the furnace. $\dot{n}_{H2OElectrode}$ is the molar flow of water for electrode cooling, $\dot{n}_{H2OInjection}$ is the molar flow of water from a water injection. The model does not include humidity from air drafted into the furnace. However, the assumption is valid for the purpose of calculating a boundary for the expected water vapor leaving the furnace. For the

calculation presented in this work, the injection is varied from 0 liters per minute (i.e. minimum boundary) to 60 liters per minute (i.e. upper boundary) to calculate the lower and upper boundaries for the expected water vapor leaving the furnace. The molar flow rate of the expected water vapor is obtained with Eq. (4.13):

$$\dot{n}_{H2O} = \dot{n}_{TotalH2} - \dot{n}_{H2} \tag{4.13}$$

Moreover, the wet off-gas molar flow ($\dot{n}_{Wet}$) and molar percent of the expected water vapor (y_{H2O}) are obtained with Eq. 4.14 and Eq. 4.15, respectively:

$$\dot{n}_{Wet} = \dot{n}_{H2O} + \dot{n}_{Dry} \tag{4.14}$$

$$y_{H2O} = \frac{\dot{n}_{H2O}}{\dot{n}_{Wet}} \times 100\ \% \tag{4.15}$$

4.2.2 EAF 1 Boundary Model Results

Figure 4.2 shows the comparison between the calculated water vapor upper and lower boundaries and the measured water leaving EAF 1 during a heat with 3 charges and a refining period.

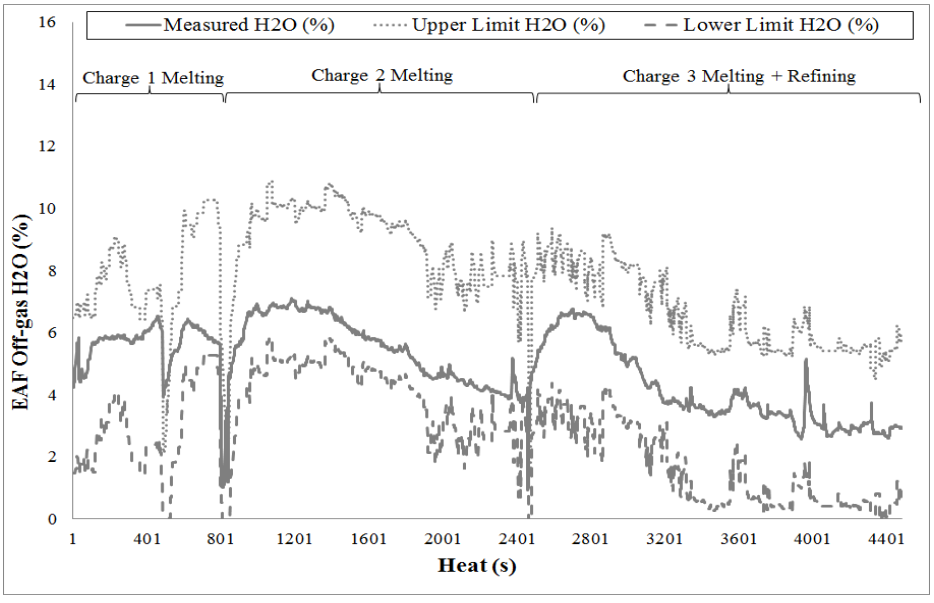

Figure 4.2: Comparison between Measured Off-gas Water Vapor and the Calculated Boundary for a Normal Heat for EAF 1.

This heat is a normal heat where there is no leak in the furnace. The boundaries are calculated based on equations 4.5-4.15. As shown in Figure 4.2, the calculated boundaries follow a similar trend as the measured water vapor. Typically in the beginning of the charge melting, the water vapor increases due to oxy-fuel burners (i.e. natural gas combustion), and then it decreases gradually because less natural gas is used. The sharp spikes seen in the upper and lower limits are due to spikes in some of the algebraic model input variables such as off-gas CO. However, the difference between measured and calculated ranges between 1% and 6% during the heat. This difference is due to the assumptions in the simplified algebraic mass balance approach. Figure 4.3 shows the comparison between the calculated water vapor upper and lower boundaries and the measured water vapor leaving EAF 1 in the off-gas during a trial heat with 4 charges and a refining period.

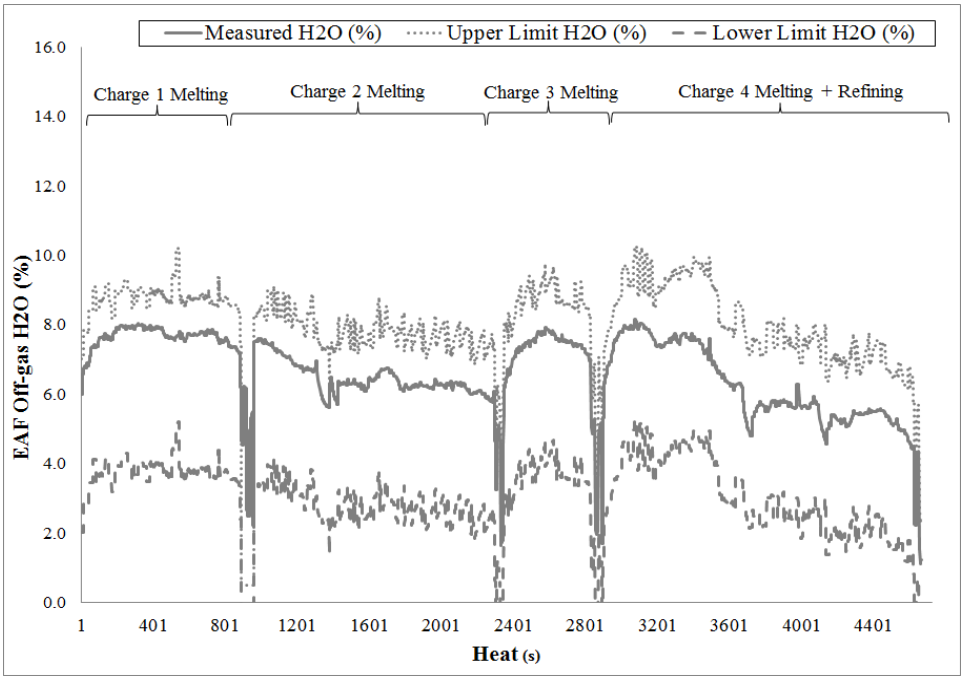

Figure 4.3: Comparison between Measured Off-gas Water Vapor and the Calculated Boundary for a Trial Heat for EAF 1.

In this trial heat an additional 60 liters per minute was injected into the furnace from the electrode spray water from the beginning until the end of the heat to test if less difference can be observed between measured water vapor and the calculated upper

boundary. The measured water showed an increase of roughly 2% going from a range of 4-6% in the previous normal heat to 6-8 % due to the additional of water added to the furnace. Figure 4.3 shows that the calculated boundaries follow a similar trend as the measured water vapor. However, since the algebraic model does not account for any additional water being injected into the furnace, the measured water vapor is shifted towards the upper limit due to the additional increase in water. The calculated boundaries are conservative due to the assumptions made in the algebraic mass balance model; hence, the measured water vapor does not go above the upper limit, albeit water was injected into the furnace.

4.2.3 EAF 2 Boundary Model Results

Figure 4.4 shows the comparison between the calculated water vapor boundaries and the measured water leaving EAF 2 in the off-gas during a heat with DRI feeding and refining.

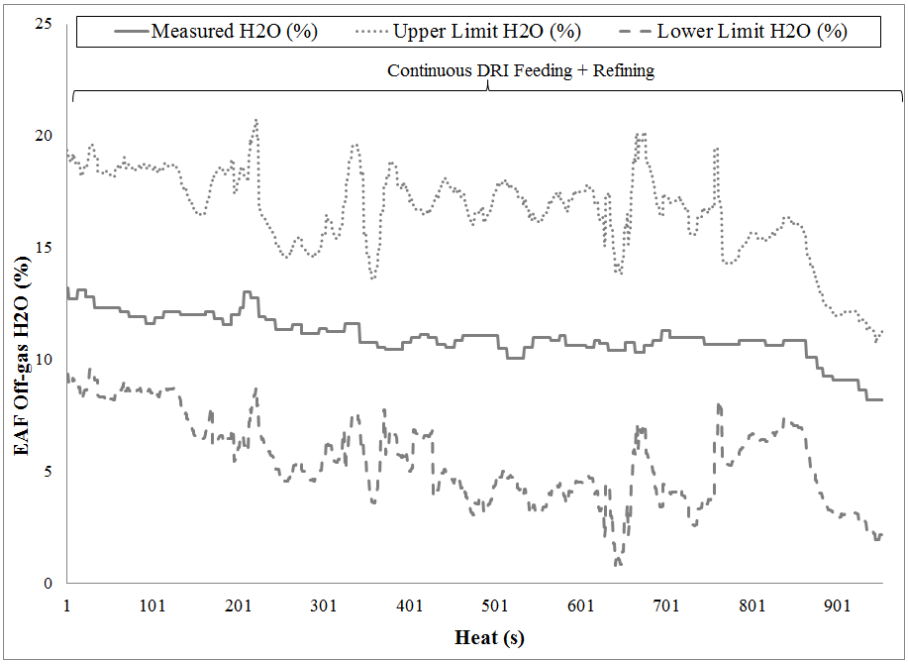

Figure 4.4: Comparison between Measured Off-gas Water Vapor and the Calculated Boundary for a Normal Heat for EAF 2.

This is a normal heat where there is no additional water injected into the furnace. As shown in Figure 4.4, the calculated boundaries follow a similar trend as the measured water vapor. Typically, oxy-fuel burners are used less often in DRI operations compared to scrap operations because the steel bath is near flat bath condition. The reason that water vapor in this furnace is higher than EAF 1 is due to higher coherent burners firing rates, hence more natural gas is used and the electrode spray water flow rate is higher. The sharp spikes seen in the upper and lower limits are due to spikes in some of the algebraic model input variables such as off-gas CO. However, the difference between measured and calculated ranges between 4-7%. This difference is because of the assumptions in the simplified algebraic mass balance approach.

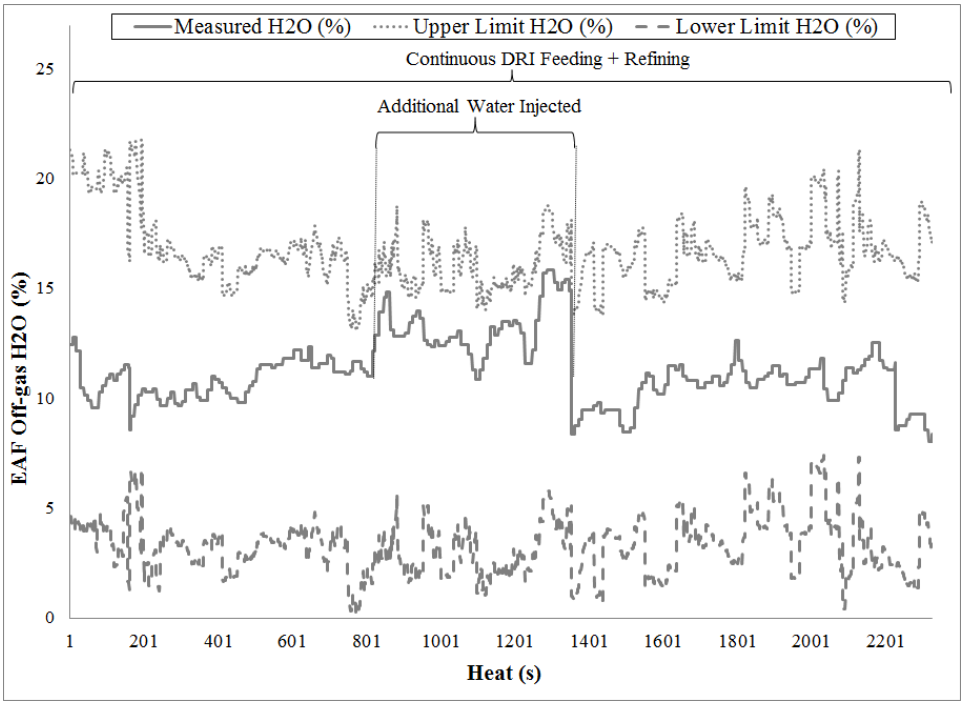

Figure 4.5: Comparison between Measured Off-gas Water Vapor and the Calculated Boundary for a Trial Heat for EAF 2.

Figure 4.5 shows the comparison between the calculated water boundaries and the measured off-gas water vapor leaving EAF 2 during a trial heat. In this heat an additional 60 liters per minute was injected from the electrode spray water in the middle of the heat for approximately 8 minutes to test if less difference can be observed between measured water vapor and the calculated upper boundary. The measured water showed

an increase of roughly 3%. Figure 4.5 shows the calculated boundaries follow a similar trend as the measured water vapor. However, the measured water vapor is shifted towards the upper limit due to the additional increase in water. The calculated boundaries are conservative due to the assumptions made in the algebraic mass balance model where the measured water vapor does not go above the upper limit even though the injected water into the furnace was large.

In both industrial cases, the boundary formulation model illustrates that an upper and lower limits for expected water vapor are developed from furnace process information and off-gas analysis to provide a useful tool to the operator. The next step in developing the hybrid model proposed in this work is to develop the empirical model. The first step in developing the empirical model is to determine the set of normal heats and important variables that are used to build the predictive model.

4.3 Heats and Variables Selection Models

4.3.1 Introduction

In a typical EAF heat in both melt-shops, there are more than 30 variables that are being measured in real-time throughout the heat. This condition presents a highly dimensional problem with noisy and collinear variables because some of these variables are correlated with each other and some of them have low signal to noise ratio. The correlation between the measured EAF variables and the off-gas water vapor varies, where some of these variables are more correlated with the off-gas water vapor than other variables. Moreover, there is a cause and effect relationship between some of the variables and off-gas water vapor, such as electrode spray water and EAF fuel flowrate, where if electrodes spray water increases, the off-gas water vapor would also increase.

The data in this work is collected in matrices X and Y. Figure 4.6 shows the three dimensional nature of the heat data blocks (X and Y). The three dimensional X array consists of N historical heat rows with K variables columns measured in real time at J

time samples. Each heat has a different heat duration (i.e. power on time in minutes) because of the different scrap and DRI qualities. Smaller scrap pieces require less energy to melt and hence require less heat time than larger pieces, and scrap with more impurities requires more heat time in order to reach final liquid steel composition; therefore, the J time samples are different from heat to heat. Moreover, the output variable (EAF off-gas water vapor) is measured by the off-gas analyzer from the beginning until the end of the heat.

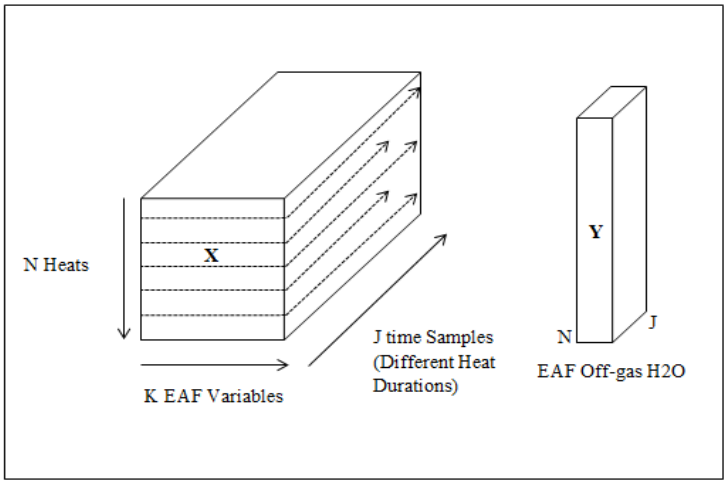

Figure 4.6: EAF Heat Data Structure.

The number of historical heats included in this analysis is 24 heats for EAF 1 and 51 heats for EAF 2. The historical heats for both electric arc furnaces are selected to span different operating practices that are considered typical operations. Table 4.1 shows the historical heats input variables and the output variable for EAF 1 and 2. Input variables that do not have variability in them are ignored when building a data driven model. Therefore, total shroud oxygen flow rate and electrode flow rate were ignored in the input variables for EAF 1 because their values were constant throughout the heat.

Table 4.1: Heats Data Set for EAF 1 and EAF 2

EAF	Number of Heats	Input Variables	Output Variable
EAF 1	24	Total Main Oxygen Flow rate, Total Fuel Flow rate, Total Carbon Flow rate, Off-gas CO_2, Off-gas CO, Off-gas O_2, EAF Pressure, EAF Fume System Damper Position, EAF Current Phase 1, EAF Current Phase 2, EAF Current Phase 3, EAF Active Power, Last Charge Indicator, EAF Transformer Position	EAF off-gas H_2O
EAF 2	51	Total DRI Flow rate, Total Shroud Oxygen Flow rate, Total Main Oxygen Flow rate, Total Fuel Flow rate, Total Carbon Flow rate, Electrode Flow rate, Off-gas CO_2, Off-gas CO, Off-gas Analyzer Purge H_2O, EAF Current Phase 1, EAF Current Phase 2, EAF Current Phase 3, EAF Transformer Position	EAF off-gas H_2O

High dimensional problems have received significant attention where it has been studied extensively in literature (Lin et al. (2011)). Latent variable methods such as MPCA (Multiway Principal Component Analysis) and MPLS (Multiway Projection to Latent Structures) are well-known techniques in dealing with high dimensional batch data with many, noisy, and collinear variables (Wold et al., 1987). MPCA first unfolds the three dimensions batch data into two dimensions and then it is used on the unfolded heat to capture the correlation structure on the measured variables (i.e. X Array) and projects this into a lower dimensional latent structure (Wold et al., 1987; Nomikos and MacGregor, 1994). MPLS extends the MPCA method to incorporate the output Y dataset. Similarly, MPLS first unfolds the three dimensions batch data into two dimensions, and then it is used on the unfolded batch dataset to model the relationship between the two matrices, unfolded X and Y, and project them into a lower dimensional latent structure (Wold et al., 1987; Nomikos and MacGregor, 1995). MPCA is used in this work to analyze the historical heats and then to select the normal heats dataset for EAF 1 and EAF 2. Consequently, MPLS is used in this chapter to select the important input variables in correlation with the output off-gas water vapor. The normal heats and the input variables selected in this chapter are used in the following chapter by the MPLS

model and the ANN model to predict off-gas water vapor for both furnaces. Particular problems discussed in this chapter are:

1. Data preprocessing, including alignment, unfolding method, centering, and scaling;
2. Latent Variable Methods (MPCA and MPLS) descriptions;
3. Classification of a heat as normal on the basis of its process data using MPCA;
4. Identify important input variables related to the normal Heats using MPLS;

4.3.2 Heats Data Preprocessing

The objective of aligning the heat trajectories is to ensure that the variables at any time during one heat correspond to those at the same time in other heats (Wold et al. (2009)). The approach followed in this work to align the heats is linear warping against heat time (Rodrigues et al. (2006)). After aligning the heats for both furnaces, the dataset (X and Y) in the heats are unfolded from a three dimensional array to a two dimensional matrix, and then the heat data is mean-centered and scaled. Mean-centering removes arbitrary bias from the measurements by moving the data into the center of the coordinate system, and scaling removes the different units from the measurements by making them unitless.

4.3.2.1 EAF 1 Heat Data Preprocessing

The indicator used in the linear warping alignment for EAF 1 heats is the centered and scaled specific electrical consumption (kWh/ton). The first step followed to align the heats is to choose a reference heat that represents a normal heat with a typical melting practice. The reference heat is divided into three phases because there are three buckets. Figure 4.7 shows the three phases in the reference heat. The centered and scaled specific electrical consumption for the first bucket is 0 to 0.25, for the second bucket is 0 to 0.55, and for the third bucket and refining is 0 to 1. In this plant, three buckets heats are the common practice. However, any four or two buckets heats are excluded from this analysis.

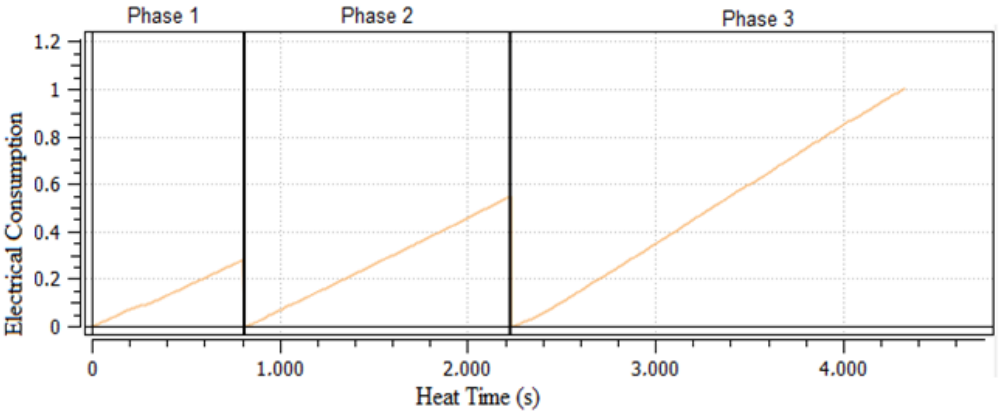

Figure 4.7: Different Phases for the Specific Electrical Consumption (kWh/ton) in the Reference Heat for EAF 1.

Figure 4.8 shows the before alignment plot and Figure 4.9 shows the after alignment plot for all the EAF 1 historical heats for a centered and scaled EAF fuel flowrate. The bold black curve in both plots is the reference heat. Figure 4.8 shows how historical heats before alignment have different heat times because the EAF fuel flow rate ends at different points in the three phases, and after alignment all historical heats end at the time in each phase as the reference heat. Furthermore, all input variables in the X array are aligned in the same approach. The heat time is normalized from seconds to %.

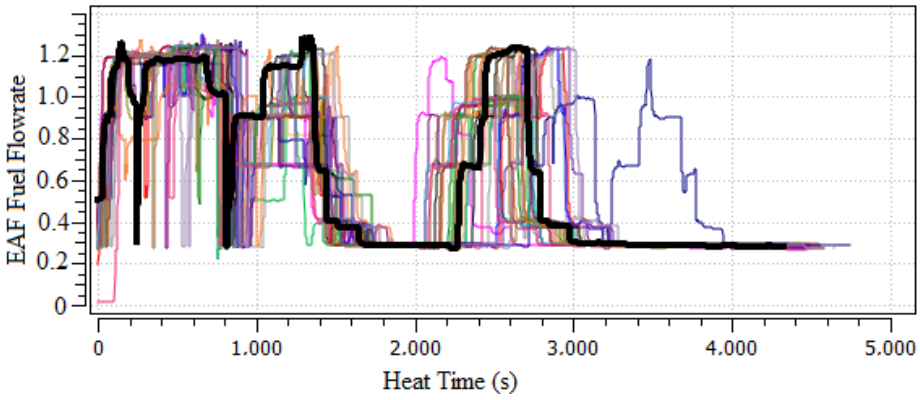

Figure 4.8: Pre Alignment Plot for all Heats for EAF 1 Fuel Flowrate.

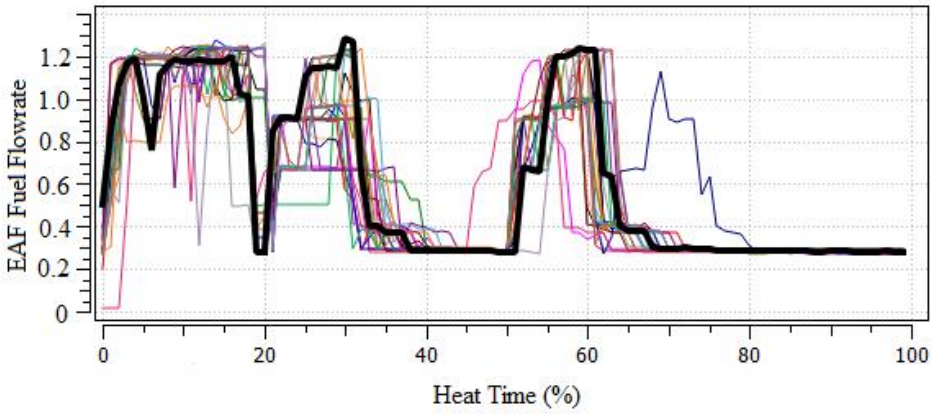

Figure 4.9: Post Alignment Plot for all Heats for EAF 1 Fuel Flowrate.

Figure 4.10 shows the before alignment plot and Figure 4.11 shows the after alignment plot for all the EAF 1 historical heats for the EAF off gas water vapor. The bold black curve in both plots is the reference heat. Figure 4.10 shows how historical heats before alignment have different heat times because the EAF off-gas water vapor ends at different points in the three phases, and after alignment all historical heats end at the same time as the reference heat. The heat time is normalized from seconds to %.

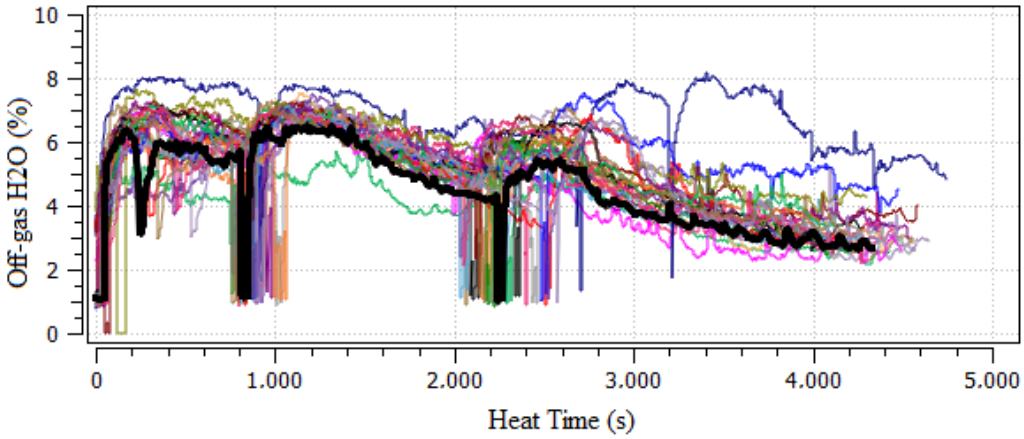

Figure 4.10: Pre Alignment Plot for all Heats for EAF 1 Off-gas H_2O.

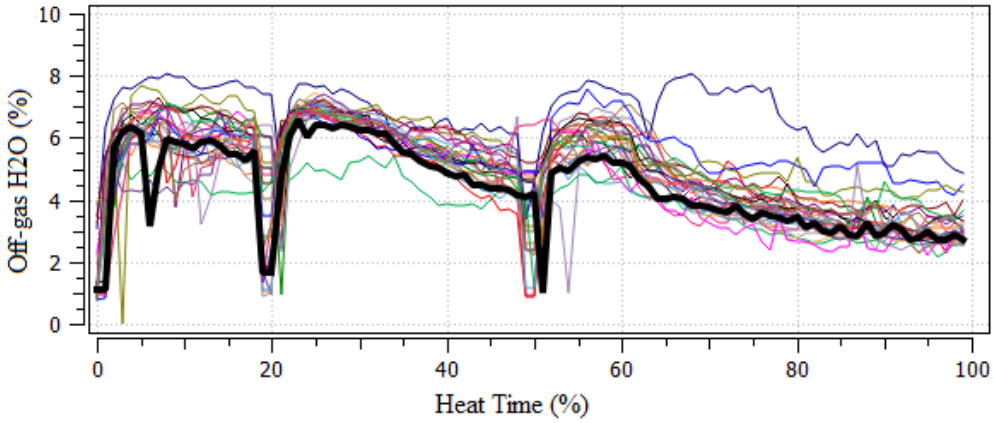

Figure 4.11: Post Alignment Plot for all Heats for EAF 1 Off-gas H_2O.

4.3.2.2 EAF 2 Heat Data Preprocessing

The indicator used in the linear warping alignment for EAF 2 heats is the % DRI fed to the furnace which normally ranges from 0 to 100%. However, the end point can range from 90 to 110% based on the DRI composition fed into the furnace. The first step followed to align the heats is to choose a reference heat that represents a normal heat with % DRI fed ranges from 0 to 100% and with a typical heat practice. Figure 4.12 shows the five phases in the reference heat. Each phase represents a DRI feed rate mode. In this plant, this % DRI mode is the common practice; hence other heats with different practices are excluded from this analysis.

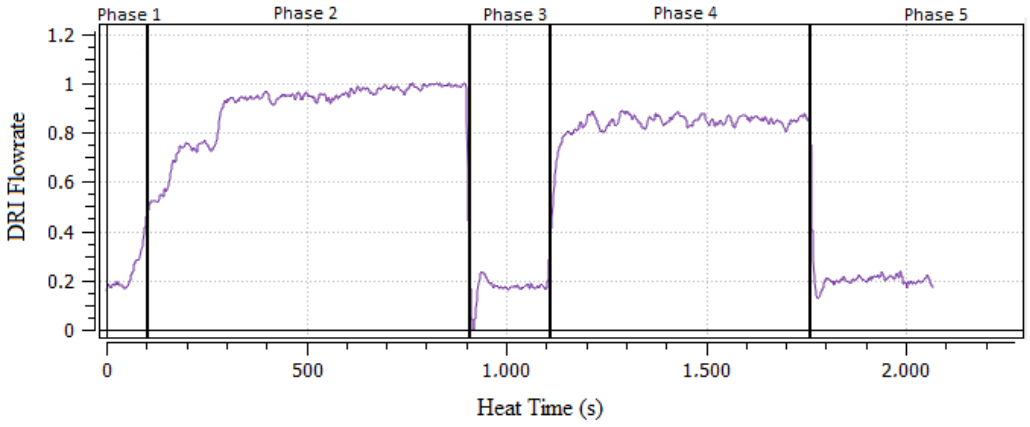

Figure 4.12: Different Phases for the DRI Feeding Flowrate in the Reference Heat for EAF 2.

Figure 4.13 shows the before alignment plot and Figure 4.14 shows the after alignment plot for all the EAF 2 historical heats for a centered and scaled EAF fuel flow rate. The bold black curve in both plots is the reference heat. Figure 4.13 shows how historical heats before alignment have different heat times because the EAF fuel flow rate ends at different points in the five phases, and after alignment all historical heats end at the time in each phase as the reference heat. The heat time is normalized from seconds to %. Furthermore, all EAF 2 input variables in the X array are aligned in the same approach.

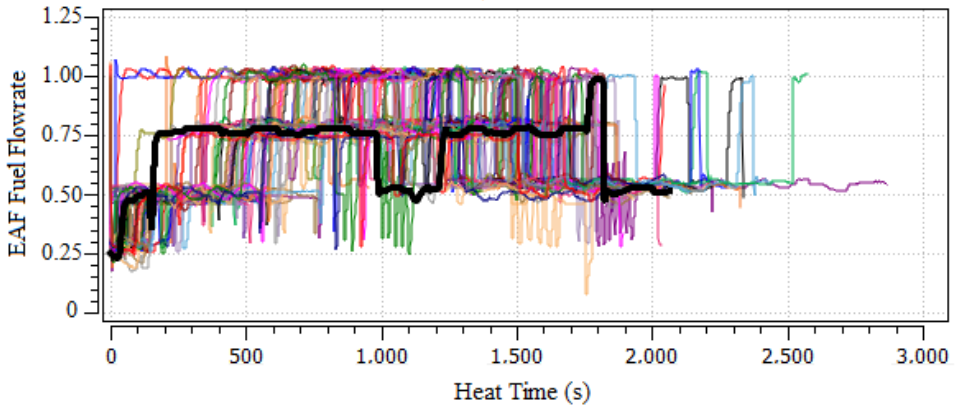

Figure 4.13: Pre Alignment Plot for all Heats for EAF 2 Fuel Flowrate.

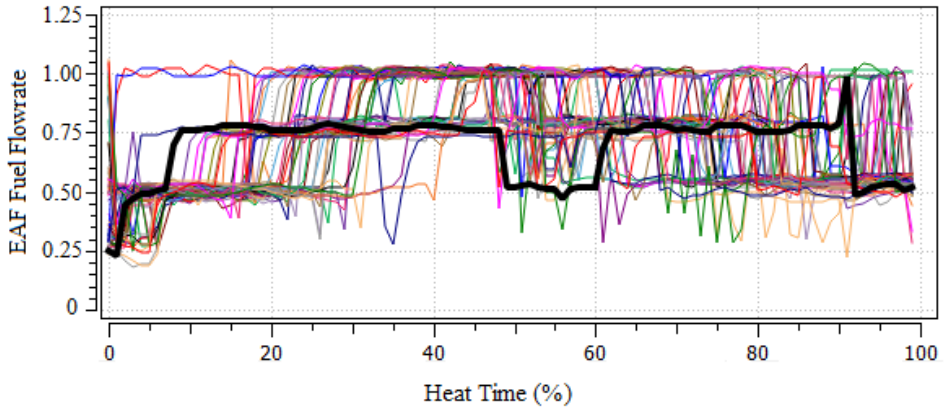

Figure 4.14: Post Alignment Plot for all Heats for EAF 2 Fuel Flowrate.

Figure 4.15 shows the before alignment plot and Figure 4.16 shows the after alignment plot for all the EAF 2 historical heats for off -gas water vapor. The bold black

curve in both plots is the reference heat. Figure 4.15 shows how historical heats before alignment have different heat times because the EAF off-gas water vapor ends at different points in the five phases, and after alignment all historical heats end at the time as the reference heat. The heat time is normalized from seconds to %.

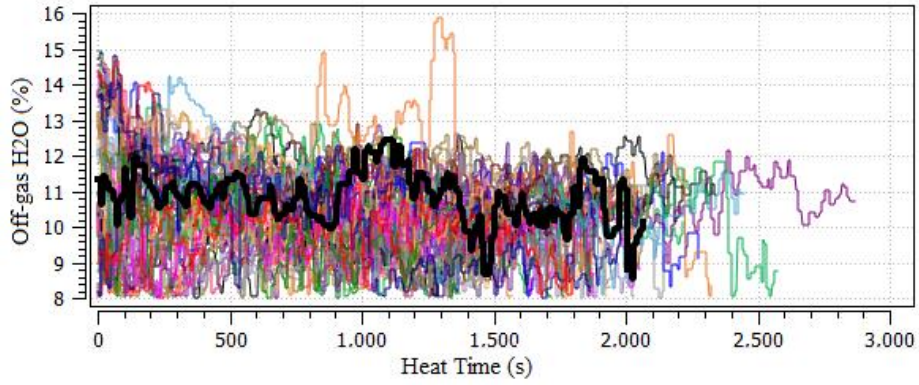

Figure 4.15: **Pre Alignment Plot for all Heats for EAF 2 Off-gas Water Vapor.**

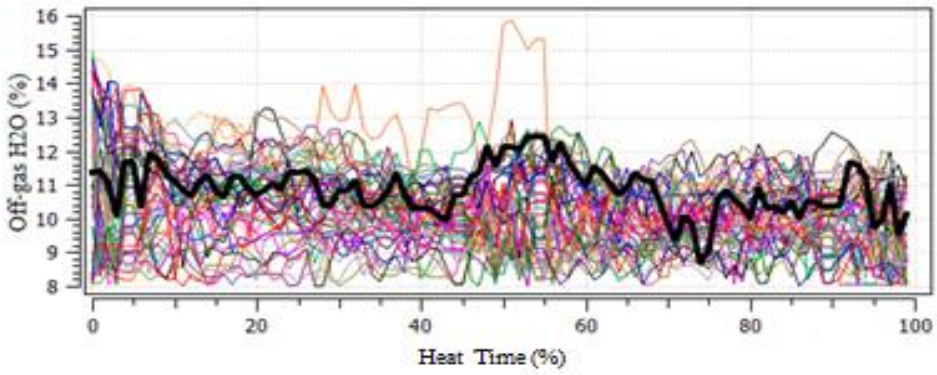

Figure 4.16: **Post Alignment Plot for all Heats for EAF 2 Off-gas Water Vapor.**

The subsequent step after alignment of the variables is to unfold the three dimensional X and Y arrays into two dimensions. The approach followed in this work to unfold the arrays is heat-wise unfolding, which is, to unfold the array such that all the information for each heat is contained in one row as shown in Figure 4.17. This unfolding method allows variability analysis between heats and captures non-linearity which is present in batch processes (Nomikos and MacGregor (1994)). The Y array is unfolded following the same approach described in Figure 4.17. The Y variable in this

process is measured continuously from the beginning of the heat until the end. Hence, in every heat N at every time sample J there is Y variable.

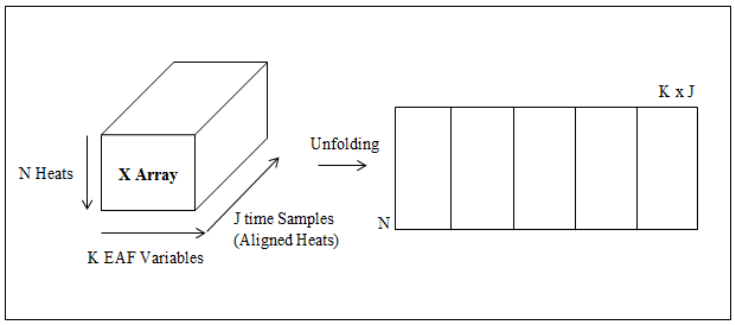

Figure 4.17: Unfolding Approach of the X Array (Nomikos and MacGregor (1994)).

Each column is then mean centered and scaled. The mean trajectory of each variable in the unfolded matrix is removed via centering, and the variation remaining is the variation of all the variables about their mean trajectories. It is important to center and scale the data properly before building the MPCA and the MPLS models because centering brings the data to the origin and scaling removes the variable variance contribution due to units of measurements. In this work, centering and scaling are calculated by the following equations:

Centering: $$x_{k,center} = x_{k,raw} - mean\,(x_{k,raw}) \quad (4.16)$$

Scaling: $$x_k = \frac{x_{k,center}}{standard\ deviation\,(x_{k,center})} \quad (4.17)$$

Thereafter, MPCA and MPLS models are built on the unfolded, centered, and scaled arrays summarizing the major sources of variation among the different heats, and such analysis allows efficient heat to heat comparison (Wold et al., 1987).

4.3.3 Multiway Principal Component Analysis (MPCA)

Multiway principal component analysis (MPCA) is a multivariate statistical method (Mardia et al., 1989; Jackson, 1991), which has the objective of best explaining

the variance in the input variables (X). MPCA reduces the dimensions of the X matrix ($K \times N$), where there are K input variables and N observations in the X matrix, into a lower dimension latent vector space. The latent vector space represents a new coordination system determined by projecting the original noisy and collinear data into a reduced space which contains most of the relevant information about the process (Wold et al. 2001). MPCA provides a simpler description of the data variability than the original data. Figure 4.18 shows the breakdown of the MPCA structure:

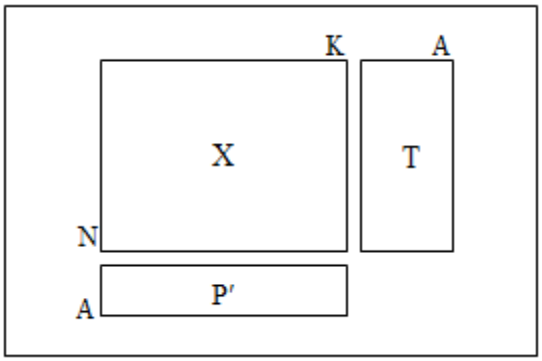

Figure 4.18: MPCA Structure Diagram (Dunn, 2014).

The MPCA latent vectors in this work are calculated using a nonlinear iterative partial least squares (NIPALS) algorithm (Geladi and Kowalski, 1986; Wold et. al., 1987a). This algorithm is used because it handles missing data and because the principal components are calculated sequentially. The number of latent vectors required to explain the variability in the data is determined by cross validation (Wold, 1978; Eastment and Krzanowski, 1982). Each latent vector is described by loading vectors (P) and score vectors (T). The loading vectors are orthogonal, and they are in the direction of maximum variability; the scores are the coordinates for the observations in the reduced space. The latent vectors measure the latent structure in the original data.

Geometrically, the latent vectors are oriented in the direction of greatest variability in X, and they are orthogonal to each other. Hence the objective is to find the direction that minimizes the residual distance from each observation to the model plane. The direction of the latent vector is defined by the *loadings P,* and they are constrained to

a unit length. Each observation is projected onto the loading vectors, and the distance from the origin to the projected point on the latent vector in the X space is the *t score*. The perpendicular distance from each observation to the latent vector is the residual for that observation (Dunn, 2014). Figure 4.19 shows a simple illustration of a two component MPCA model plane in a 3 dimensions data space (X_1, X_2, *and* X_3). The first component vector passes through the origin and is oriented in the direction that best explains the observation points in the three dimension space. The second component passes through the origin, orthogonal to the first component, and oriented in the direction that best explains the observation points in the 3 dimension space. Both components are constrained to a unit length. Now, the two loading vectors define a PCA model plan with two components.

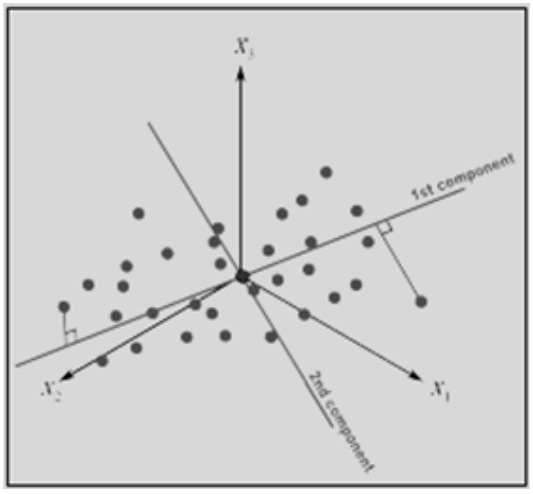

Figure 4.19: Geometric Representation of the Steps in the MPCA Model (Dunn, 2014).

The MPCA model breaks the raw heats data into a latent variable model (loadings *p*, and scores *t*) and the residual error. Mathematically, the scores (*T*), loadings (*P*), and the residuals (*E*) for the X matrix are written as follows:

$$T = XP \tag{4.18}$$

$$(N \times A) = (N \times K)\,(K \times A)$$

$$X = TP' + E \tag{4.19}$$

The best prediction of the original unfolded X array (X_{pred}) and the residual error vector of the ith observation (e_i) are calculated using the following equations:

$$X_{pred} = TP' \tag{4.20}$$

$$(N \times K) = (N \times A)\,(A \times K)$$

$$e_i = x_i - x_{ipred} \tag{4.21}$$

Important parameters used to analyze MPCA models are the residuals (R^2), squared prediction error (SPE), and Hotelling's T^2 (Nomikos and MacGregor, (1994)). SPE is the distance from the MPCA model's plane and is the square root of the sum of squares for each residual. Hence, when an observation has a large SPE value that indicates the observation has a large residual and that is inconsistent with the correlation structure of the MPCA model. Hotelling's T^2 is the directed distance from the origin to where the point is projected on the MPCA model's plane. Hence, when an observation has a large Hotelling's T^2 value that indicates the observation is consistent with the correlation structure of the MPCA model; however, the score values are larger than the normal observations score values. Hotelling's T^2 summarizes all the score values for a given observation (*i*). These three parameters are calculated using the following equations:

$$SPE_i = \sqrt{e_i' e_i} \tag{4.22}$$

$$T_i^2 = \sum_{a=1}^{a=A} \left(\frac{t_{i,a}}{s_a}\right)^2 \tag{4.23}$$

s_a^2 = variance of each component a

$$R^2 = 1 - \frac{Var(X - X_p)}{Var(X)} \tag{4.24}$$

Hotelling's T^2 has an F-distribution (Nomikos and MacGregor, (1994)), and in this work, the 95% and the 99% confidence limits are calculated and shown on the Hotelling's T^2 chart for all the MPCA models developed for the heats selection work for EAF 1 and EAF 2. Eq. 4.25 shows the general equation used to calculate the confidence limits:

$$T^2_{A,\alpha} = \frac{(N-1)(N+1)A}{N(N-1)} \times F_\alpha(A, N-A) \tag{4.25}$$

Where A is components, N is observations, and 100(1-α) % is the confidence limit. Scores are assumed to be normally distributed since the historical heats dataset is large, and in this work, the 95% and the 99% confidence limits are calculated and shown on the scores plot for all the MPCA models developed for the heats selection work for EAF 1 and EAF 2. Eq. 4.26 shows the general equation used to calculate the confidence limits:

$$\begin{gathered} t_a \sim N\,(0, s_a) \\ 100\ \times (1-\ \alpha)\%\ limit = \ \pm\left(\ t_{\frac{\alpha}{2},df}\right) s_a \\ df = N - 1 \end{gathered} \tag{4.26}$$

Where s_a is the standard deviation of score column a, *df* is the degree of freedom. SPE has Chi-squared distribution (Nomikos and MacGregor, (1994)), and in this work, the 95% and the 99% confidence limits are calculated and shown on the SPE chart for all the MPCA models developed for the heats selection work for EAF 1 and EAF 2. Eq. 4.27 shows the general equation used to calculate the confidence limits:

$$\begin{gathered} SPE_i \sim g\chi^2(h) \\ g = \frac{v}{2m} \\ h = \frac{2m^2}{v} \\ m = \ mean(SPE) \\ v = \ var\ (SPE) \end{gathered} \tag{4.27}$$

A useful diagnosing tool used in this work to compare an outlier heat to the normal set of heats is the *Contribution Plot* (Nomikos and MacGregor, (1994)). In other words, this plot shows which input variables contributed for the outlier heat to be different than the rest of the heats set. Eq. 4.28 shows the equation used to calculate the contributions:

$$contrib\ (x_k) = \left(x_{i,k}^{(to)} - x_{i,k}^{(from)}\right) \times \sqrt{\sum_a (p_{k,a} \times \frac{t_{i,a}^{(to)} - t_{i,a}^{(from)}}{s_a})^2} \qquad (4.28)$$

4.3.4 Multiway Projection to Latent Structures (MPLS)

Multiway projection to latent structure (MPLS) is an extension to the MPCA by incorporating the output (Y) array (EAF off-gas H_2O). MPLS is an established multivariate statistical method (S. Wold et al., 1984; Geladi and Kowalski, 1986) which has the objective of best explaining the variance in the input variables (X), the variance in the output variables (Y), and the covariance between X and Y. MPLS reduces the dimensions of the X matrix ($K \times N$) and the Y matrix ($M \times N$), where there are K input variables and N observations in the X matrix and M output variables and N observations in the Y matrix, into a lower dimension latent vector space. The latent vector space represents a new coordination system determined by projecting the original noisy and collinear data into a reduced space which contains most of the relevant information about the process (Wold et al. 2001). Figure 4.20 shows the breakdown of the MPLS structure:

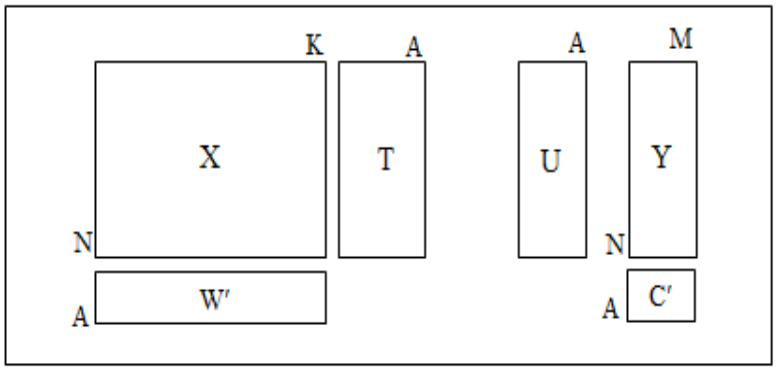

Figure 4.20: MPLS Structure Diagram (Dunn, 2014).

As in the MPCA model, the MPLS latent vectors space in this work is also calculated using nonlinear iterative partial least squares (NIPALS) algorithm, and the number of latent vectors required to explain the variability in the data is also determined by cross validation. Each latent vector is described by loading vectors (*W and C*) and score vectors (*T and U*). The loading vectors are orthogonal, and they are in the direction of maximum variability; the scores are the coordinates for the observations in the reduced space.

Geometrically, the latent vectors are oriented in the direction of greatest variability in X and Y and best correlation between X and Y. The direction of the latent space is defined by the loadings (*W in X and C in Y*), and they are constrained to a unit length. Each observation is projected onto the loading vectors, and the distance from the origin to the projected point on the latent vector in the X space is the *t score*, and the distance from the origin to the projected point on the latent vector in the Y space is the *u score*. The scores (*T and U*) summarizes both spaces (*X and Y*). The perpendicular distance from each observation to the latent vector is the residual for that observation (Dunn, 2014). Figure 4.21 shows a simple geometrical illustration of a two component MPLS model plane in a 3 dimensions data space (X_1, X_2, X_3, and Y_1, Y_2, Y_3). The first component vector passes through the origin and is oriented in the direction that best explain the covariance between X and Y. The second component passes through the origin, orthogonal to the first component, and oriented in the direction that best explain the covariance between X and Y. The latent vectors are constrained to a unit length.

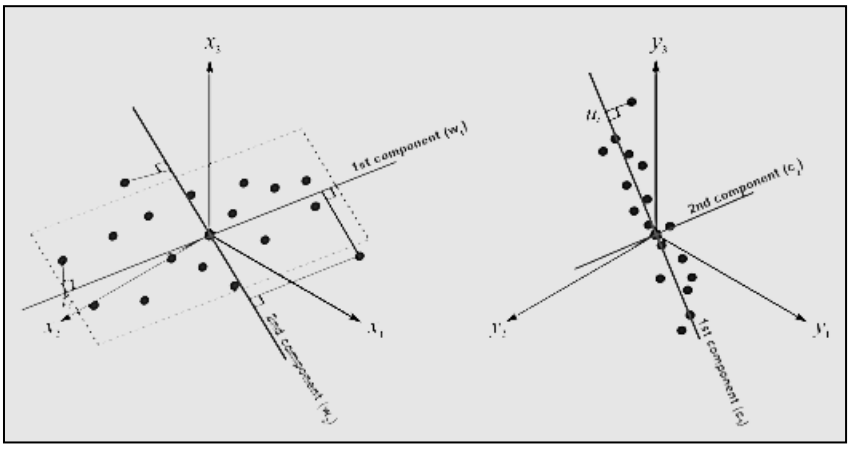

Figure 4.21: Geometric Representation of the Steps in the MPLS Model (Dunn, 2014).

Mathematically, the objective for MPLS is to determine the scores (T in X and U in Y) that have maximum covariance. The covariance is calculated as:

$$Cov(t_a,u_a) = Correlation\ (t_a,u_a) \times \sqrt{t'_a t_a} \times \sqrt{u'_a u_a} \tag{4.29}$$

Hence maximizing the covariance between t_a and u_a is maximizing the variance in the X space ($t'_a t_a$), the variance in the Y space ($u'_a u_a$), and the relationship between X and Y spaces (Correlation (t_a, u_a)), simultaneously. The MPLS model breaks the raw heats data into a latent variable model (loadings *W and C*, and scores *T and U*) and the residual error. Mathematically, the scores (*T*) and loadings (*W*) for the X matrix are calculated as follows (Wold et al., 1987):

$$T = XW^* \tag{4.30}$$

$$(N \times A) = (N \times K)\,(K \times A)$$

$$W^* = W(P'W)^{-1} \tag{4.31}$$

Mathematically, the scores (*U*) and loadings (*C*) for the Y matrix are calculated as follows (Wold et al., 1987):

$$U = YC \tag{4.32}$$

$$(N \times A) = (N \times M)\,(M \times A)$$

However, since they have maximal covariance, interpreting one of them is sufficient. The *T* scores are always available in the model building stage, and when the model is used on a new dataset, whereas the *U* scores are only available in the model building step, hence the *T* scores are used in this work. Important plots that are used in this work to analyze relationships between the unfolded X and Y arrays are the score plot, which shows the relationship between the scores *T* and *U*, and the weight and loading plot (*W*C*) which shows the relationship between the loadings *W* and *C*. As with the MPCA model, important parameters used to analyze MPLS models are the residuals (R^2), squared prediction error (SPE), and Hotelling's T^2. A tool that is used to analyze which input variable is important in correlation with the off-gas water vapor is the variables importance to prediction (VIP) plot. Importance of variable *k* using *A* components in MPLS is calculated using Eq. 4.33:

$$VIP_{A,k}^2 = \frac{K}{SSX_0 - SSX_A} \times \sum_{a=1}^{A} (SSX_{a-1} - SSX_A) W_{a,k}^2 \tag{4.33}$$

Where SSX_a is the sum of squares in the X matrix after *a* components.

In Chapter 5, the MPLS model built in this chapter will be used to predict off-gas H_2O on new testing heats to validate the accuracy of the off-line model. The following steps summarize the methodology followed in Chapter 5 to predict off-gas water vapor on testing heats:

1) Build the MPLS model off-line using historical good heats (Chapter 4)
2) Center and Scale $x_{new,raw}$ using the same mean and standard deviation of the training heats set to get x_{new}
3) Calculate t scores: $t_{new} = x_{new}W^*$

4) Calculate projected x : $x_{new,p} = t_{new}P'$

5) Calculate Residual $e_{new} = x_{new} - x_{new,p}$

6) Calculate SPE from e_{new} and check with the limits from the offline model. If the SPE is lower than 99% confidence limit then continue to step 7.

7) Calculate Hotelling's T^2 from t_{new} and check with the limits from the offline model. If the Hotelling's T^2 is lower than 99% confidence limit then continue to step 8.

8) Calculate the prediction: $y_{new,p} = t_{new} C'$

9) Uncenter and Unscale $y_{new,p}$ back to the off-gas water vapor units (%)

4.3.5 Heats Selection Using MPCA

MPCA is used in this section to perform an analysis on the historical heats to discriminate between normal heats and abnormal heats. This analysis is critical because abnormal heats are removed from the training set and because major source of heat to heat variations are studied.

4.3.5.1 EAF 1 Heat Selection Model

MPCA model is built on the processed EAF 1 input variables data. The numbers of principal components obtained are 7 components using cross validation. The calculated model R^2 is 82%. Figure 4.22 shows the score plot for the first two components that explains approximately 61% of the variability in the 24 heats data. The remaining 5 components in the MPCA model explain approximately 21% of the variability in the historical heats data and hence showing the first two components is sufficient. The score plot shows that all heats are distributed evenly in the four quadrants, and hence there is no clustering in a certain region on the plot. Moreover, heat 14 was above the 99% limit. Figure 4.22 shows the 95% and 99% confidence interval of the model for the first two principal components (T1 and T2):

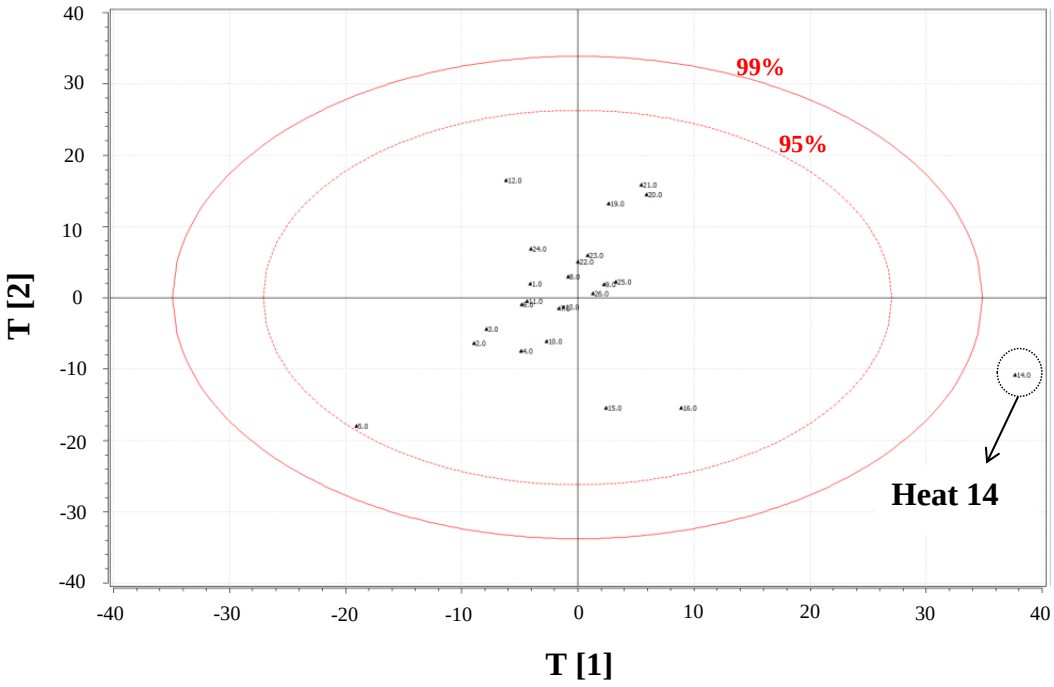

Figure 4.22: MPCA Model Score Plot for the First Two Components (T1 and T2) for all EAF 1 Historical Heats.

The contribution plot shown in Figure 4.23 shows the difference of the trajectory between heat 14 and the average of the rest of the historical heats from the beginning of the heat until the end using bar plot for each variable. The contribution plot shows a significant number of bars in the EAF fuel flow and main oxygen flow in the last segment of the heat, which is towards refining that caused more CO to be generated, which caused the fume system to react to this event. Figure 4.24 shows the EAF main oxygen trajectory between heat 14 and the average of the rest of the heats. The figure clearly shows a difference during refining between heat 14 and the rest of the heats. The heat time is normalized from seconds to %.

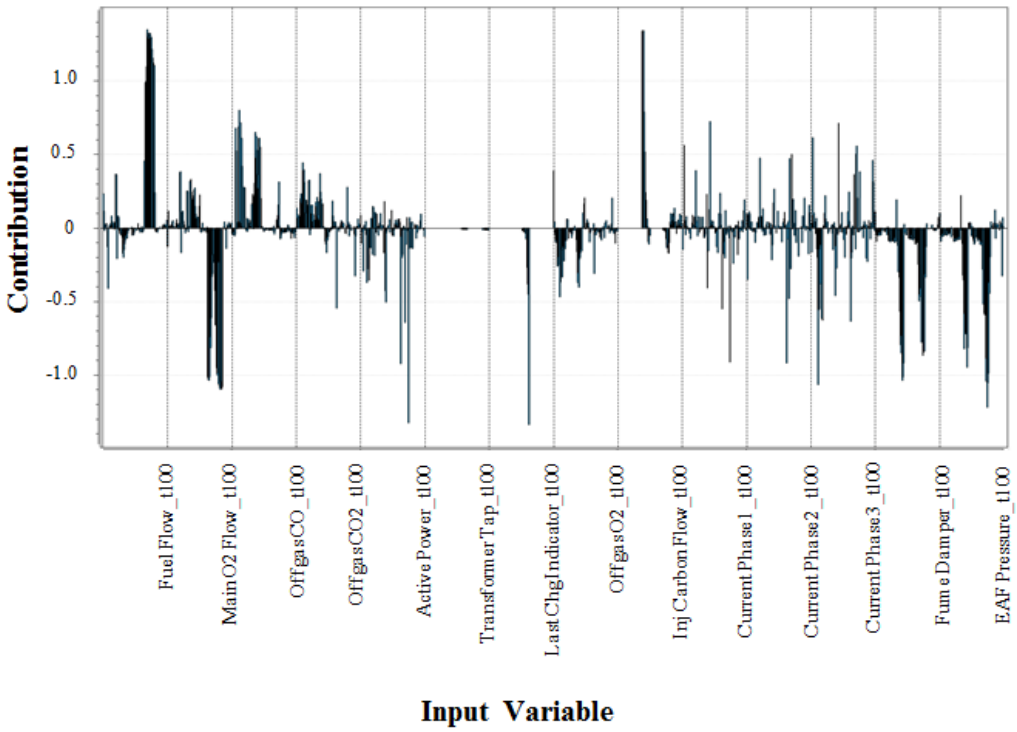

Figure 4.23: MPCA Model 1 EAF 1 Heat 14 Contributions Plot.

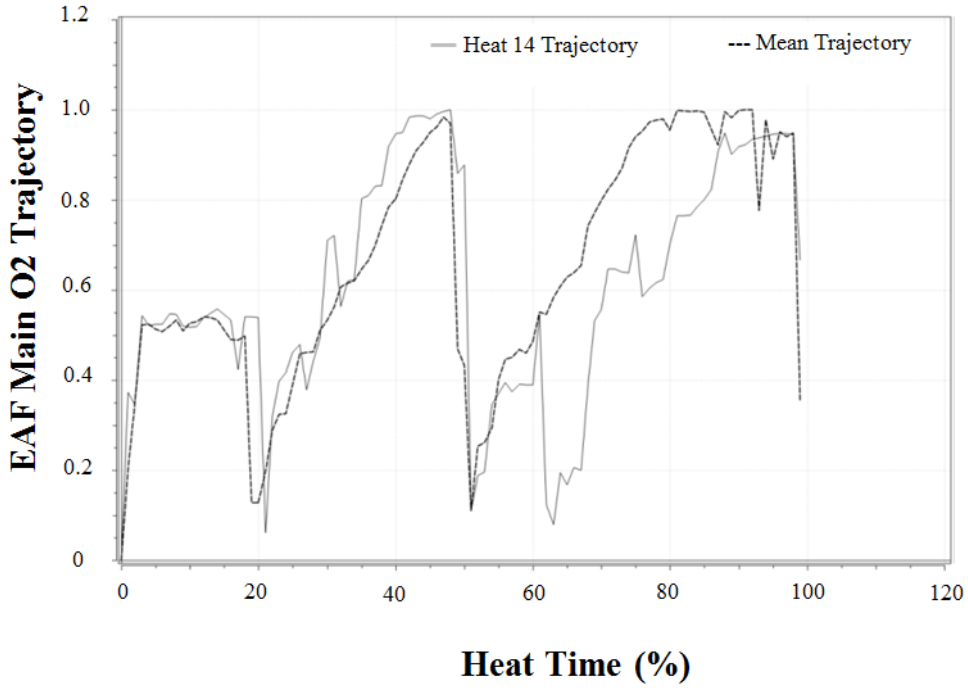

Figure 4.24: MPCA Model 1 Heat 14 Main Oxygen Flow rate Trajectory vs. Average Heats Trajectory for EAF 1.

Figure 4.25 shows the Hotelling's T^2 plot for the MPCA model. The plot shows that all the heats are below the 95% confidence interval. Figure 4.26 shows the SPE plot for the MPCA model. The plot also shows that all the heats are below the 95% confidence interval. Since Heat 14 is below the 99% in both charts, then it is an indication that the heat is not an outlier, albeit there is a difference during refining in the EAF fuel and main oxygen practice. Therefore, heat 14 is kept in the data set.

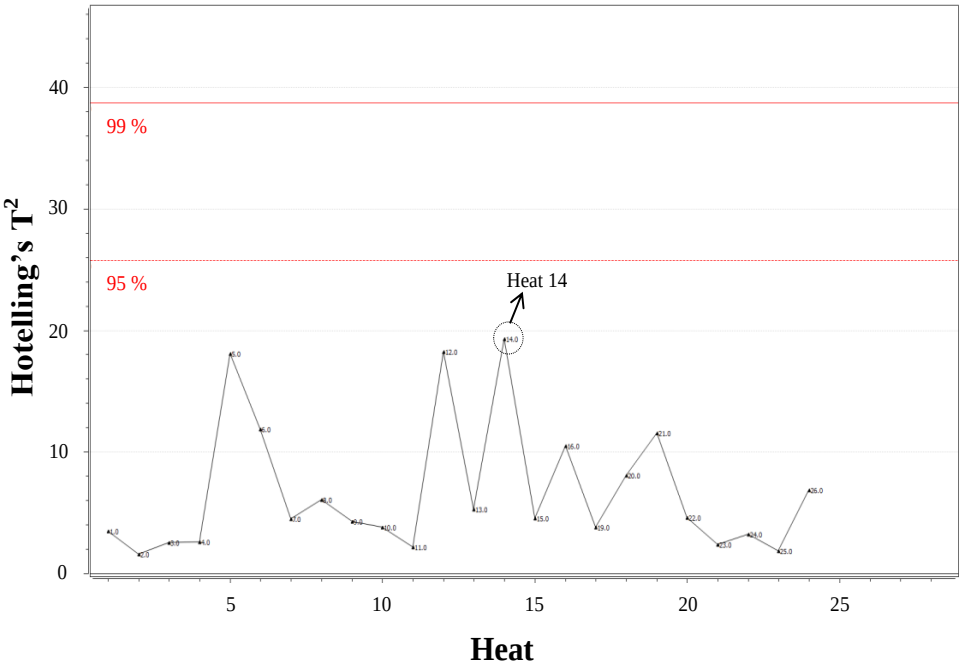

Figure 4.25: MPCA Model Hotelling's T^2 Plot for EAF 1 Historical Heats.

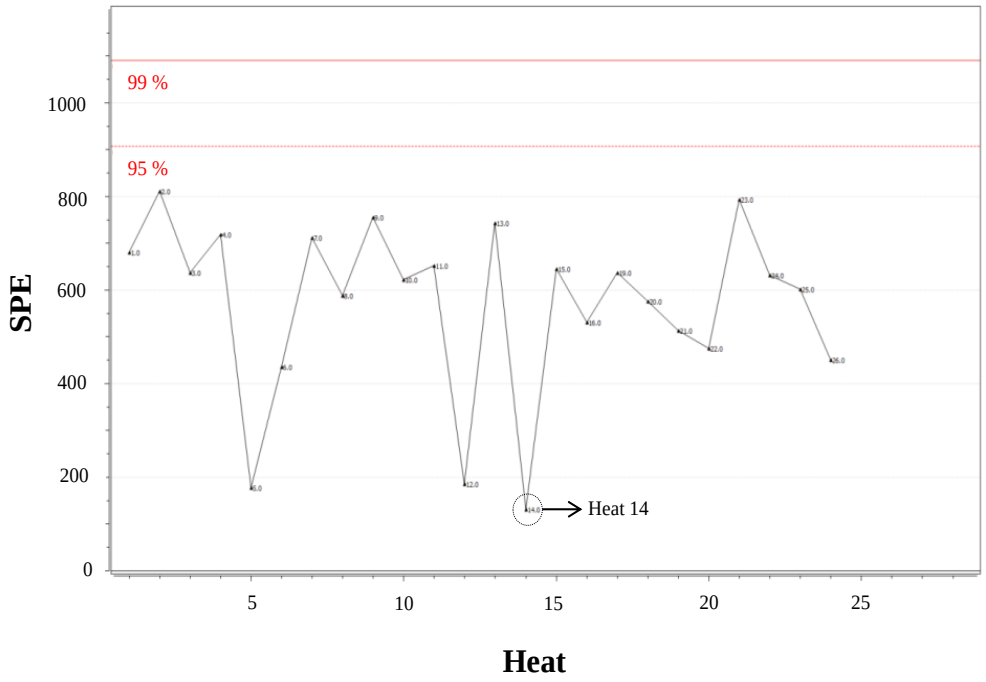

Figure 4.26: MPCA Model SPE Plot for EAF 1 Historical Heats.

The MPCA model built in this section for EAF 1 heats attempts to explain the predictable variation in dataset. The scores plot, Hotelling's T^2 plot, and SPE plot show that the MPCA model describes sufficiently the normal EAF 1 heats dataset, where these historical heats exhibit normal statistical properties. The next step for EAF 1 dataset is to select the most important variables in correlation with the off-gas water vapor that are going to be used in the next chapter to build the MPLS model and the ANN model to predict the off-gas water vapor.

4.3.5.2 EAF 2 Heat Selection Model

MPCA model is built on the processed EAF 2 heat data, and the number of components obtained is 8 using cross validation. The R^2 obtained is 71%. Figure 4.27 shows the score plot for the first two components that explain the approximately 50% of the variability in the 51 historical heats data. The remaining 6 components in the MPLS model explain approximately 21 % of the variability in the historical heats data, and hence showing the first two components is sufficient.

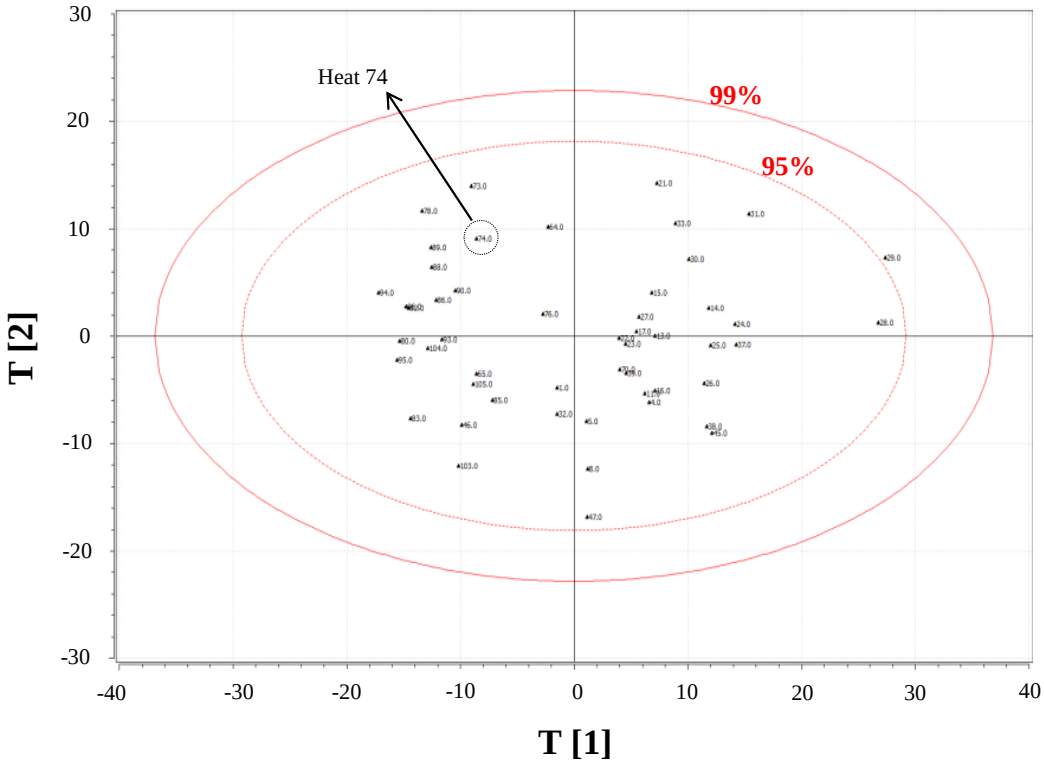

Figure 4.27: MPCA Model 1 Score Plot for the First Two Components (T1 and T2) for all EAF 2 Historical Heats.

The dotted elliptic shows the 95% confidence interval and the solid elliptic shows the 99% confidence interval. The score plot shows that all heats are distributed evenly in the four quadrants, and hence there is no clustering in a certain region on the plot. Moreover, none of the heats are above the 99% limit in the first two components score plot. The scatter feature in the scores plot indicates that these heats belong to the same normal population.

Figure 4.28 shows the Hotelling's T^2 plot for the MPCA model for EAF 2. The dotted line is the 95% confidence interval, and the solid line is the 99% confidence interval line. This plot also shows that all the heats are below the 99% confidence interval.

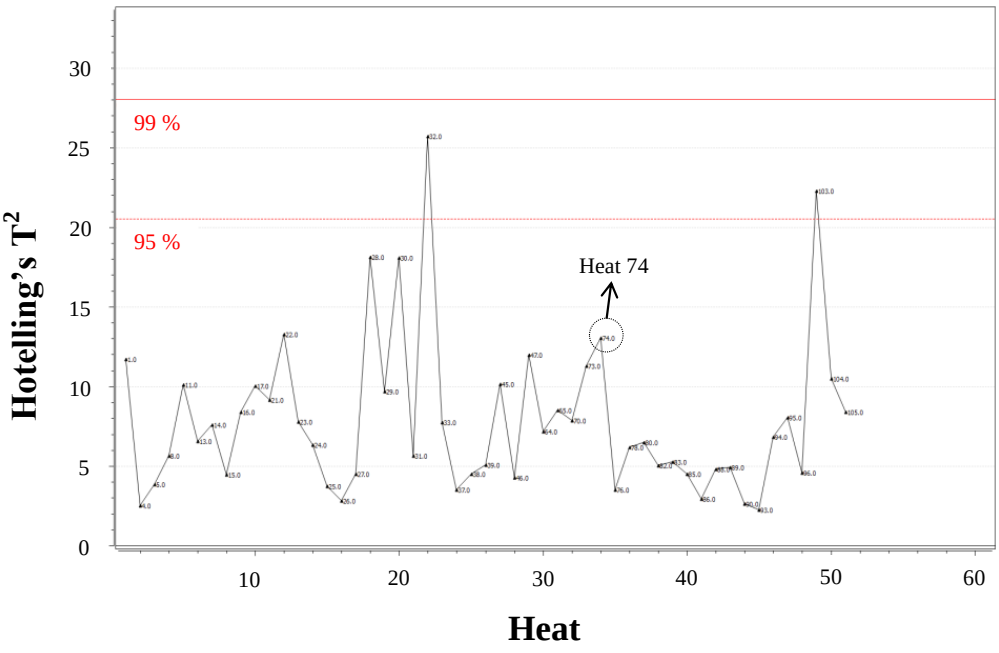

Figure 4.28: MPCA Model 1 Hotelling's T^2 Plot for EAF 2 Historical Heats.

Figure 4.29 shows the SPE plot for the MPCA model 1 for EAF 2. The dotted line is the 95% confidence interval and the solid line is the 99% confidence interval line. This plot shows that Heat 74 is above the 99% confidence limit.

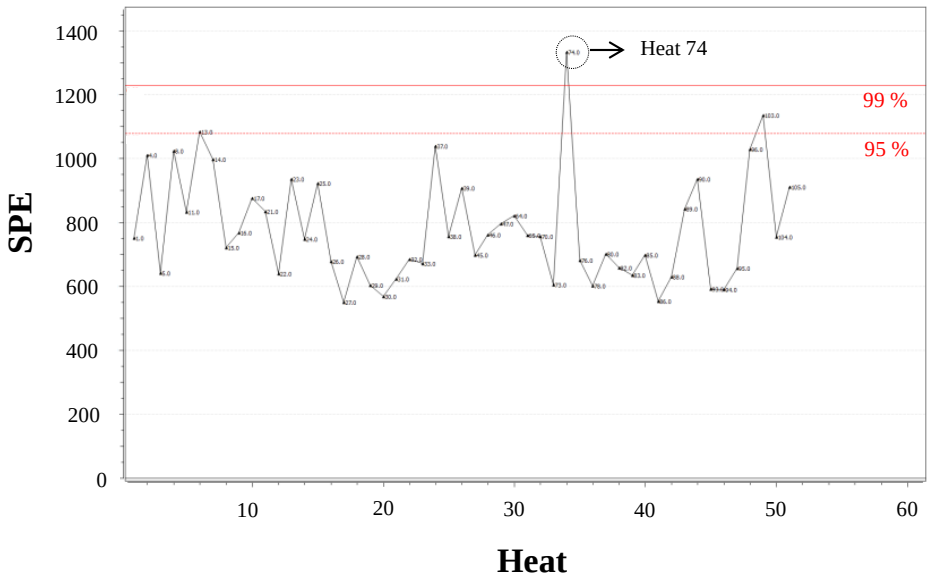

Figure 4.29: MPCA Model 1 SPE Plot for EAF 2 Historical Heats.

Figure 4.30 explains the contributions of the all input variables that caused heat 74 to have a SPE value above 99% confidence limit. The contribution plot shows the difference of the trajectory between heat 74 and the average of the rest of the historical heats from the beginning of the heat until the end using bar plot for each variable. The contribution plot shows a significant number of bars in the EAF shroud oxygen. Figure 4.31 shows both trajecotories on a heat time basis (Heat 74 and the average of the MPCA model) indicating that there is a difference in the EAF shroud oxygen practice between heat 74 and the rest of the heats. The heat time is normalized from seconds to %.

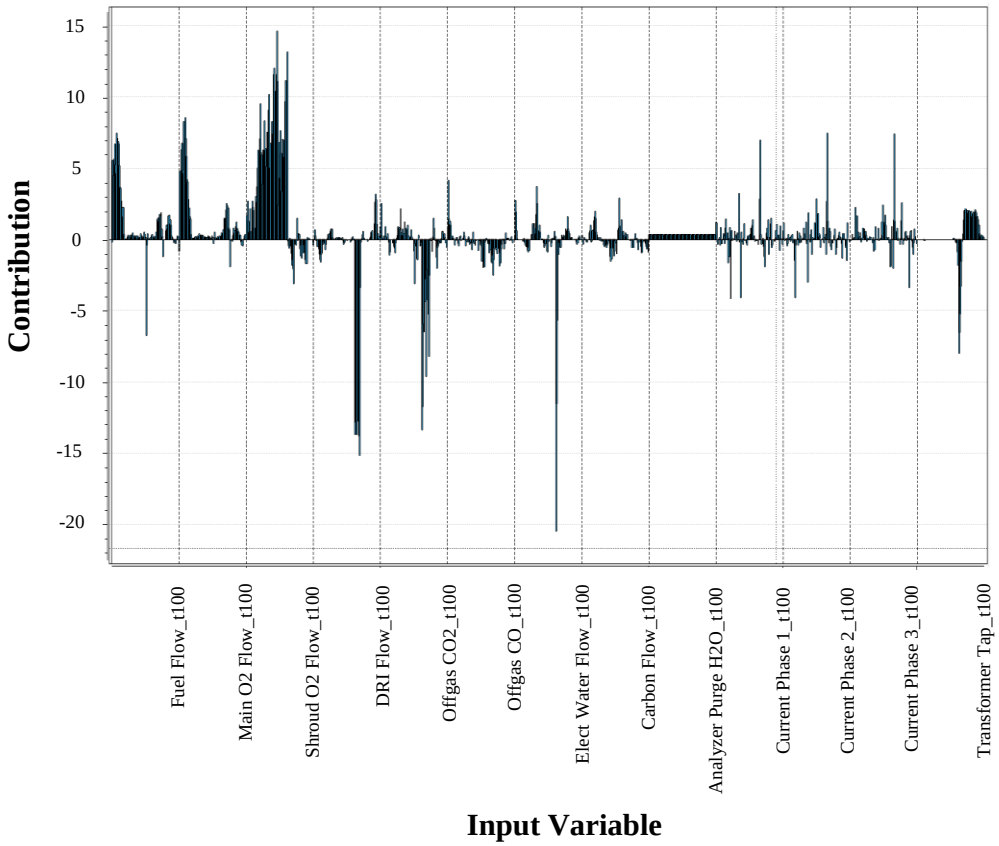

Figure 4.30: MPCA Model 1 EAF 2 Heat 74 Contributions Plot.

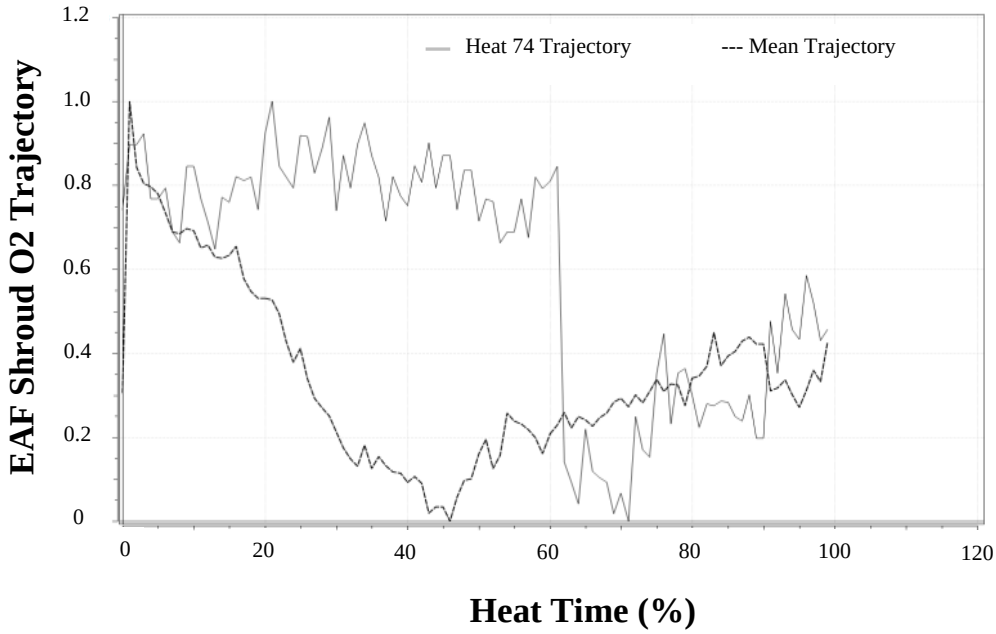

Figure 4.31: MPCA Model 1 Heat 74 EAF Shroud Oxygen Flow rate Trajectory vs. Average Heats Trajectory for EAF 2.

The SPE chart indicates how far the heat is above the model plane, and since heat 74 is above the 99% confidence limit, heats 74 is removed from the data set. The MPCA model is rebuilt on the remaining 50 heats, and the number of components obtained is 8 components explaining 72% of the variability. The first principal component explains approximately 35% of the variability, and the second component explains an additional 10%. Figure 4.32 shows the score plot for the first two components that explain approximately 47% of the variability in the historical 50 heats data set. The score plot shows that all 50 heats are distributed evenly in the four quadrants, and hence there is no clustering in a certain region on the plot. Figure 4.33 shows the Hotelling's T^2 plot for the MPCA model 2 for EAF 2. This plot also shows that all the heats are below the 99% confidence interval. Figure 4.34 shows the SPE plot for the MPCA model 2 for EAF 2. This plot also shows that all the heats are below the 99% confidence interval.

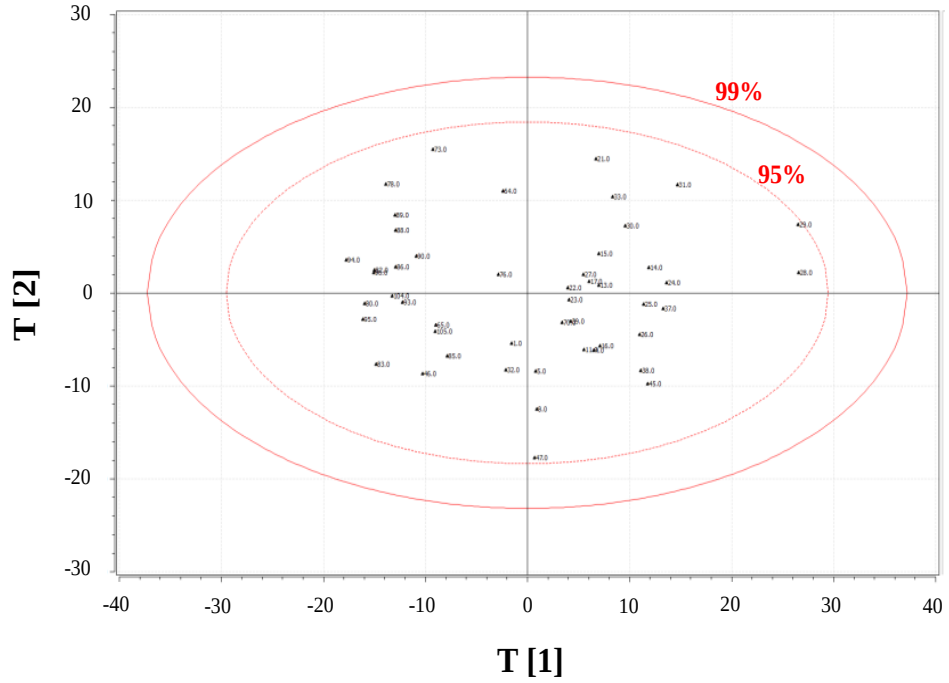

Figure 4.32: MPCA Model 2 Score Plot for the First Two Components (T1 and T2) for all EAF 2 Historical Heats.

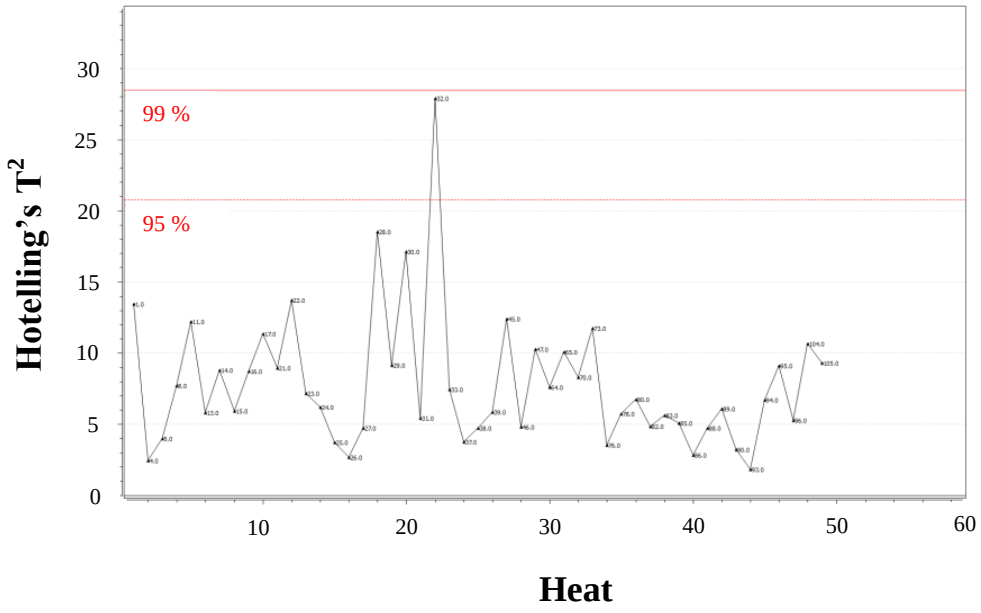

Figure 4.33: MPCA Model 2 Hotelling's T^2 Plot for EAF 2 Historical Heats.

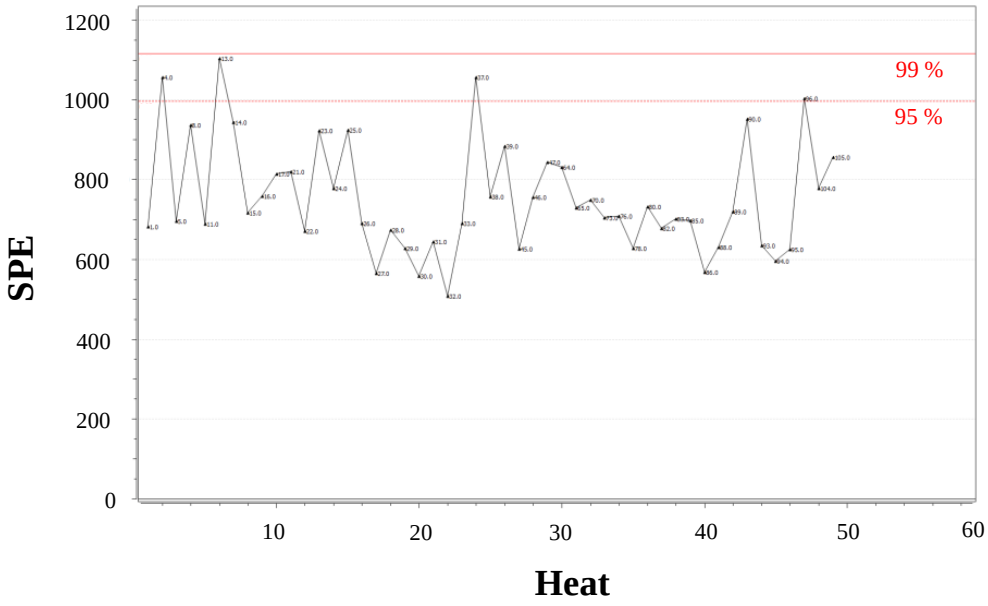

Figure 4.34: MPCA Model 2 SPE Plot for EAF 2 Historical Heats.

The scores plot, Hotelling's T^2 plot, and SPE plot show that the MPCA model 2 describes sufficiently the normal EAF 2 heats dataset, where these historical heats exhibit normal statistical properties. The next step for EAF 2 dataset is to select the most important variables in correlation with the off-gas water vapor that are going to be used in the next chapter to build the MPLS model and the ANN model to predict the off-gas water vapor.

4.3.6 Variable Selection Using MPLS

MPLS is used in this section to perform an analysis on the correlation between the different input variables and EAF off-gas water vapour. This analysis is critical in building an empirical predictive model because in a typical EAF operation there are many variables with varying correlations with the off-gas water vapour. The objective in this section is to select the most important variables in correlation with the off-gas water vapour for EAF 1 and EAF 2.

4.3.6.1 EAF 1 Variable Selection Model

MPLS model is built on the processed EAF 1 input variables data and the 24 selected heats from the heat selection model developed in the previous section. The number of principal components obtained is 7 components using cross validation. The calculated model R^2 is 82%. The first two components explain approximately 61% of the variability in the 24 heats. The remaining 5 components in the MPLS model explain approximately 21% of the variability in the dataset.

Figure 4.35 shows the correlation structure throughout the heat between each of the input variables and the off-gas water vapor on the first component. The chart is divided into sections. Each variable section starts from "t0" which is the beginning of the heat and ends at "t100" which is the end of the heat. Each section shows the correlations of the variable compared with the other variables. The bars in each section move up and down based on the correlation with the other variables. If the bars in the two variables sections are in the same direction, then they are positively correlated; if they are in opposite directions, then they are negatively correlated; and if there are no bars, then there is no correlation. Figure 4.35 shows that the correlation between the different input variables and the off-gas water variable varies through the heat. Figure 4.35 also shows that a variable such as a transformer tap position has a weaker correlation with the off-gas H_2O because the values are almost zero in the transformer tap position section from t0 to t100.

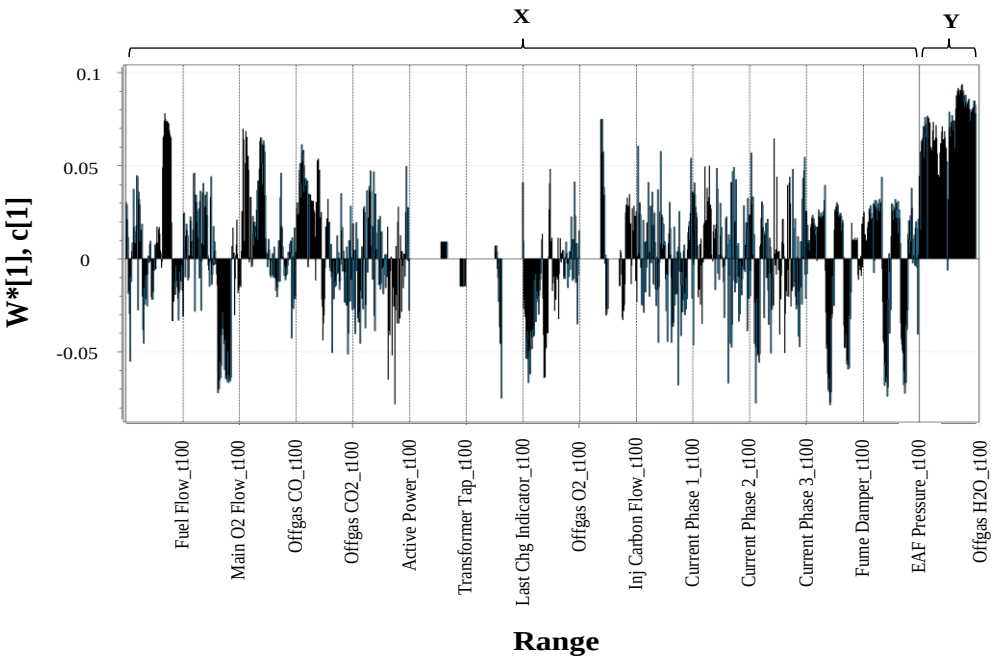

Figure 4.35: MPLS Model First Principal Component Loading Plot for EAF 1.

Figure 4.36 shows the MPLS model variable importance plot in correction with the off-gas water vapor for EAF 1. The plot shows that variables such as EAF pressure, main O2 Flow, and EAF fuel flow have stronger correlation with the off-gas water vapor than last charge indicator or transformer tap position. EAF pressure is measured at the furnace elbow, and it is a measurement of the pressure in the furnace freeboard. Hence, there is a physical correlation between the measured pressure in the freeboard and the amount of water vapor in the freeboard. Main oxygen and natural gas flow injected into the furnace directly influence the amount of off-gas water vapor generated in the furnace as shown in the mechanistic mass balance model developed in chapter 4; hence there is a cause and effect relationship. Furthermore, variables such as last charge indicator and transformer tap position have a weaker correlation with the off-gas water vapor.

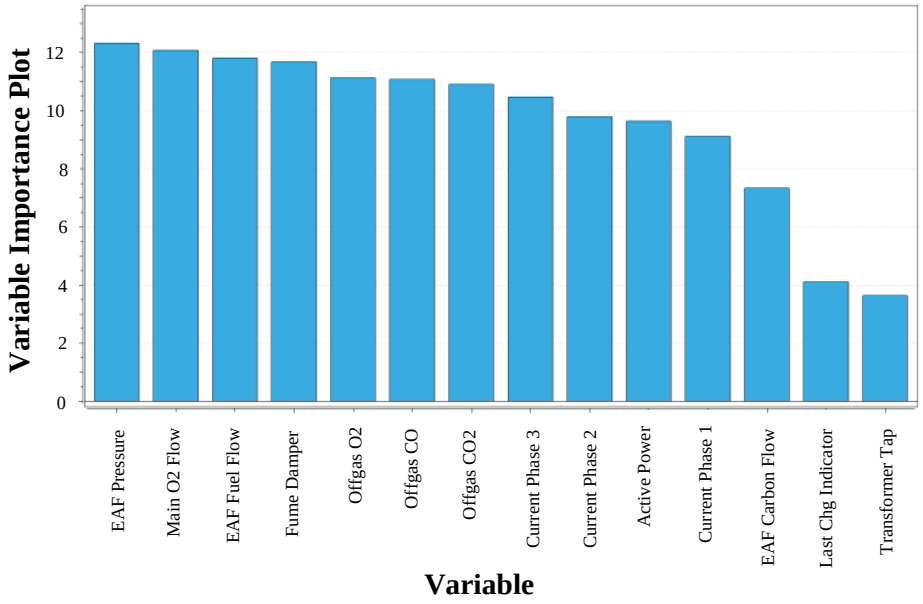

Figure 4.36: MPLS Model VIP Plot for EAF 1.

Based on the above MPLS model VIP plot, the following variables are removed when building an empirical predictive model because there is a weaker correlation with the off-gas water as shown in the above VIP plot: Current phase 1, 2, 3, last charge indictor, EAF Active Power, and EAF transformer position. EAF carbon is kept because from a process point of view, carbon powder may contain some humidity, so it is kept as an input to the predictive model.

4.3.6.2 EAF 2 Variable Selection Model

MPLS model is built on the processed EAF 2 input variables data and on the 50 selected heats from the heat selection model developed in the previous section. The number of principal components obtained is 7 components using cross validation. The calculated model R^2 is 72%. The first two components explain approximately 45% of the variability in the 50 heats. The remaining 5 components in the MPLS model explained approximately 27 % of the variability in the dataset.

Figure 4.37 shows the correlation structure throughout the heat between each of the input variables and the off-gas water vapor on the first component. Similar to the MPLS model for EAF 1, the chart is divided into variable sections. Figure 4.37 shows that the correlation between the different input variables and the off-gas water variable varies through the heat. Figure 4.37 also shows that a variable such as s transformer tap position have a zero correlation from t0 to t50 and then a negative weak correlation from t50 to t100 with the off-gas water vapor. Moreover, there is a strong positive correlation between the analyzer purge water vapor and the off-gas water vapor.

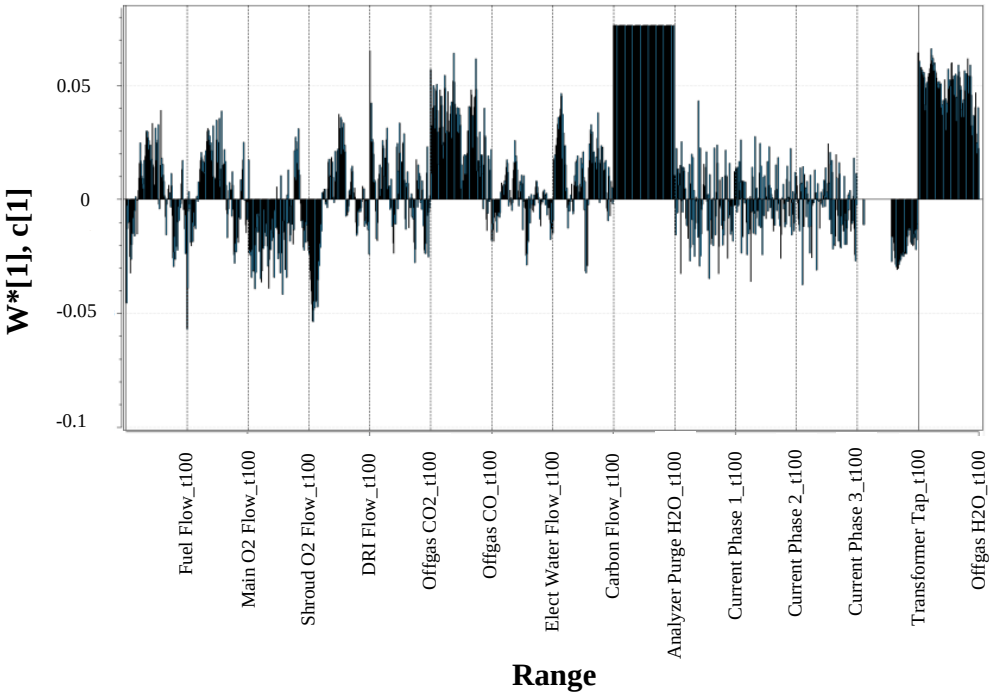

Figure 4.37: MPLS Model First Principal Component Loading Plot for EAF 2.

Figure 4.38 shows the MPLS model variable importance plot in correction with the off-gas water vapor for EAF 2. The plot shows that variables such as electrode water flow and EAF fuel flow have stronger correlation with the off-gas water vapor then transformer tap position. Electrode water flow contributes to the water vapor in the off-gas. The electrode water flow varies in this furnace, and hence it is included in the

MPLS model, where in EAF 1, this variable was constant throughout the heat; hence the MPLS model ignored it. Furthermore, variables such as off-gas CO and CO_2 are correlated with the off-gas water vapor, as shown in the mechanistic model in Chapter 4. However, variables such as transformer tap position have a weaker correlation with the off-gas water vapor, as shown in Figure 4.38.

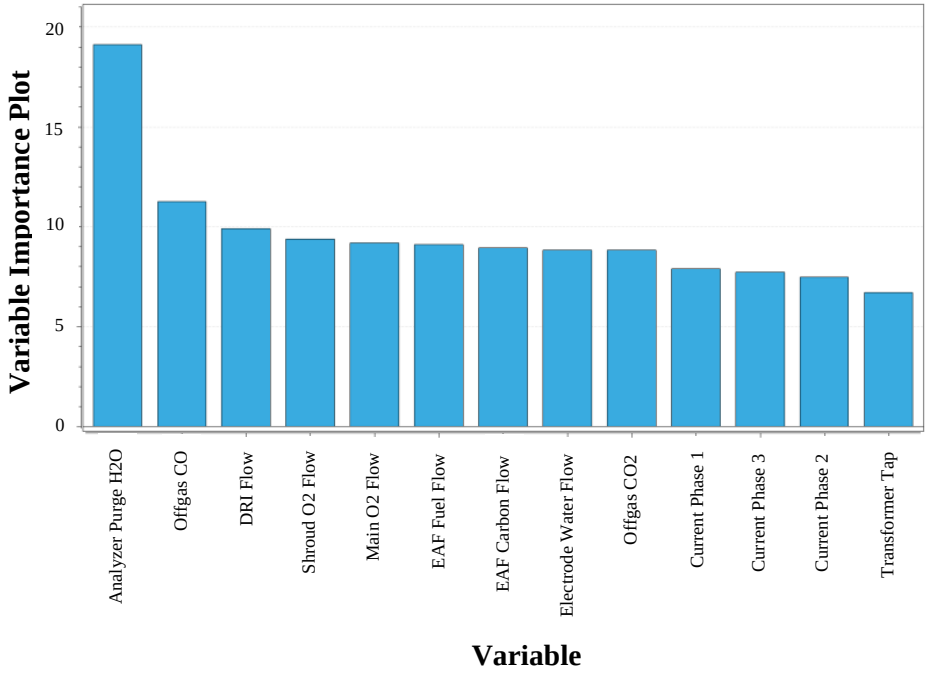

Figure 4.38: MPLS Model VIP Plot for EAF 2.

Based on the above PLS model VIP plot, the following variables are removed when building an empirical predictive model because there is weaker correlation with the off-gas water vapor as shown in the above VIP plot: Current phase 1, 2, 3, and EAF transformer position.

Chapter 5

Predictive Models and Fault Detection Methodology

5.1 Introduction

Chapter 5 presents three different empirical methods to predict the off-gas water vapor leaving the electric arc furnace. These methods are statistical fingerprinting, artificial neural network (ANN), and multiway projection to latent structures (MPLS). Statistical fingerprinting uses the off-gas water vapor measurement in its algorithm, while ANN and MPLS use the input variables and heats selected in Chapter 4 to develop the predictive models. The heats are divided into training heats and testing heats. The training heats are used to build the models, and the testing heats are used to validate the models. The testing set includes heats where water is intentionally injected in the furnace as artificial water leaks to validate if the leak can be detected by the model. These three empirical methods are implemented on both industrial EAFs. Afterwards, the prediction capability of ANN and MPLS are compared. This chapter also discusses fault detection methods developed to alarm the operator if there is a potential water leak event in the furnace.

5.2 Statistical Fingerprinting

5.2.1 Method

There are few EAF operations where the process is well controlled, and hence there is a minimal variability in the process. In these operations, scrap quality does not vary significantly, the operator uses the same chemical program, and there is limited weather variation. These conditions allow for the off-gas water vapor to behave similarly from heat to heat. Consequently, in these EAF operations, the off-gas measurement of water vapor can provide adequate indication and metrics to distinguish between normal and abnormal levels of water vapor in the EAF. Although the off-gas water vapor concentration varies throughout the melting and refining phases of the heat, the statistical fingerprinting method has been developed to characterize the off-gas water vapor over a number of heats with similar operating conditions such as number of charges, scrap recipe, and chemical program (Zuliani et al., 2014). For example, Figure 5.1 shows the typical off-gas water vapor trend for the first charge melting on a kWh basis for several heats with similar operating conditions. Figure 5.1 also shows a baseline average curve for the off-gas water vapor. The heat time shown in Figure 5.1 is kWh which is a typical energy clock used by melt-shops to pace the heat.

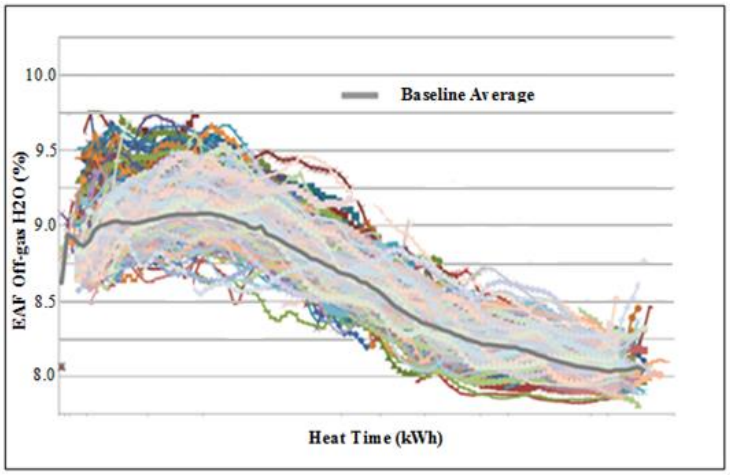

Figure 5.1: Typical Off-gas Water Vapor Trend of a First Charge on a kWh Basis for Several Heats with Similar Operating Conditions.

The first step in this method is to start with a set of heats with similar operating conditions. This set is used as a training set to calculate the fingerprints thresholds, which is then used on future heats to detect for potential water leaks. In this work, the training heats selected in chapter 4 from the MPCA model are used to calculate the fingerprints thresholds. The second step in this method is to manually divide the charge time into bins of similar dynamics. This step enables the off-gas water vapor to be characterized across multiple heats in a single bin. The third step is to compute a fingerprint threshold value in each bin that represents an upper limit. The statistic used in the fingerprinting method is the median and median absolute deviation (MAD). The median is a measure of sample location, and is computed by sorting the data and taking the middle value. The median is a robust estimator of the sample location and MAD is a robust measure of variation (Dunn, 2014). However, the median and the MAD become unbounded if half of the data is replaced with outliers. MAD is computed by Equation 5.1:

$$\text{MAD}(x_i) = c \times \text{median}(\text{abs}(x_i - \text{median}(x_i))) \qquad (5.1)$$

$$c = 1.4826$$

The constant c makes the MAD consistent with the standard deviation when the observations x_i are normally distributed. The fingerprint threshold is calculated using Eq 5.2:

$$\text{Fingerprint Upper Limit} = \text{Median} + 2 \times \text{MAD} \qquad (5.2)$$

Figure 5.2 shows an example of fingerprint threshold values for the data from Figure 5.1:

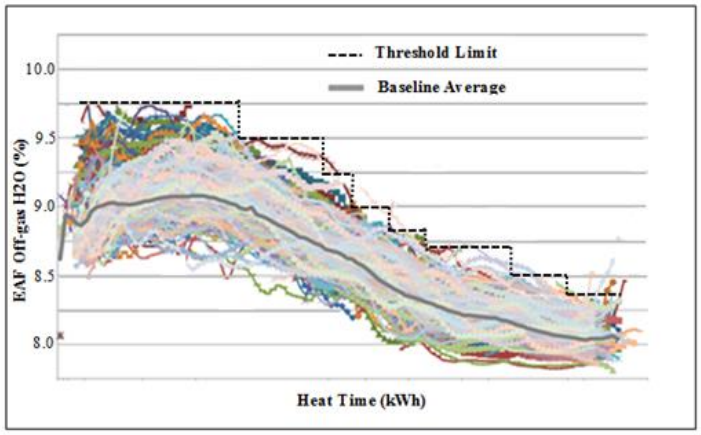

Figure 5.2: Fingerprint Threshold Values for the Data from Figure 5.1.

The fingerprint thresholds provide the basis to detect abnormal water vapor by comparing the measured water vapor of the current heat to the fingerprint threshold value in each bin. Specifically, if the value of the measured water vapor of the current heat is higher than the fingerprint threshold value, then there is a statistical condition that there is abnormal water vapor in the EAF. Figure 5.3 shows an example for a heat (red) with water vapor significantly above the fingerprint and also for an extended time.

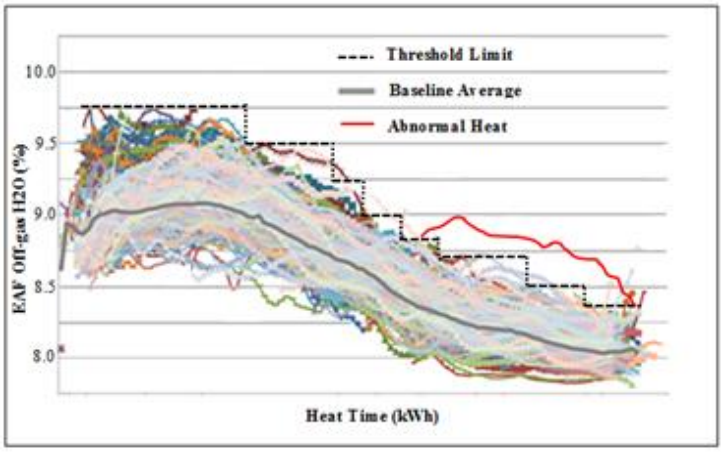

Figure 5.3: Heat with Abnormal Water Vapor Compared to the Fingerprint Threshold Limits Computed for Normal Operation.

5.2.2 EAF 1 Statistical Fingerprinting Results

Figures 5.4 and 5.5 show the off-gas water vapor for charge 1 and charge 2 for EAF 1, respectively. The maximum range of variability for the off-gas water vapor is 4% for both charges. Such variability range is typically not wide, and therefore the fingerprinting method can be tested on this furnace. The set of heats for EAF 1 that are used to compute the threshold limits and validate the model are determined in Chapter 4. Moreover, the testing heats that are used to validate the model are two trial heats where water is intentionally injected into the furnace. Figure 5.4 shows the off-gas water vapor, the bins, and the threshold limits for charge 1. The two bins shown in Figure 5.4 are determined manually, where in each bin the water vapor dynamic is similar between the heats. The fingerprint threshold value for each bin is computed using Equations 5.1 and 5.2. In Figure 5.4, trial 1 heat (red) shows a heat where water is intentionally injected into the furnace by increasing the flow of electrode spray water by approximately 60 liters per minute from the beginning of the charge 1 melting until the end. The red curve in Figure 5.4 is above the fingerprint threshold during the entire charge, and hence in this case this method is capable of detecting the additional injected water. Trial 2 heat (purple) shows the off-gas water vapor for a heat where water is intentionally injected by increasing the electrode spray water by approximately 30 liters per minute from the beginning of the charge 1 melting until the end. The purple curve in Figure 5.4 passes the threshold limit slightly, and then it drops back below the limit hence this method is not capable of detecting trial 2 for the entire charge melting. The heat time shown in Figures 5.4-5.5 is kWh/Charged ton which is the energy clock used by EAF 1 to pace the heat.

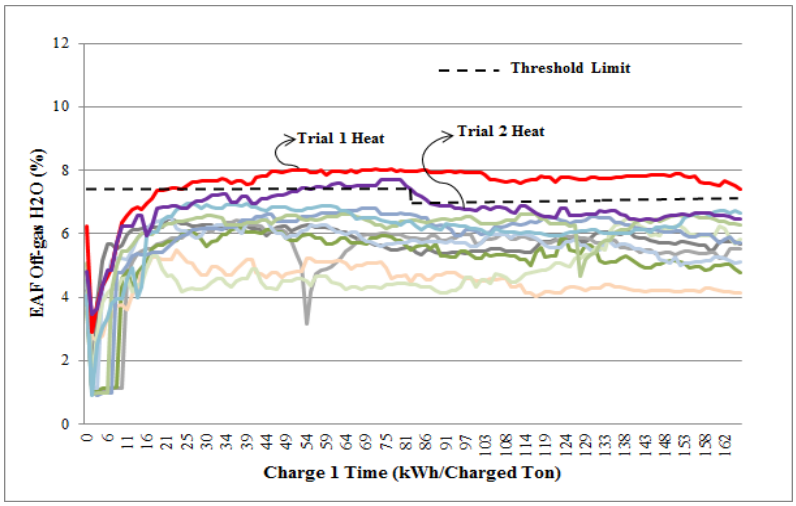

Figure 5.4: Fingerprinting Method during Charge 1 for EAF 1.

Similarly, Figure 5.5 shows 7 bins that are determined manually, where in each bin the dynamic of the water vapor measurement is similar between the heats, and the threshold value for each bin represents the upper limit to differentiate normal and abnormal operation. Equations 5.1 and 5.2 are used to calculate the threshold limits. Trial 1 heat (red) in Figure 5.5 shows the off-gas water vapor for a heat where water is intentionally injected into the furnace by increasing the electrode spray water by approximately 60 liters per minute throughout charge 2 melting. The red curve passes through the fingerprint threshold from the beginning of the melting and drops twice below the threshold limit; hence the fingerprinting method is capable of detecting the additional injection in this case.

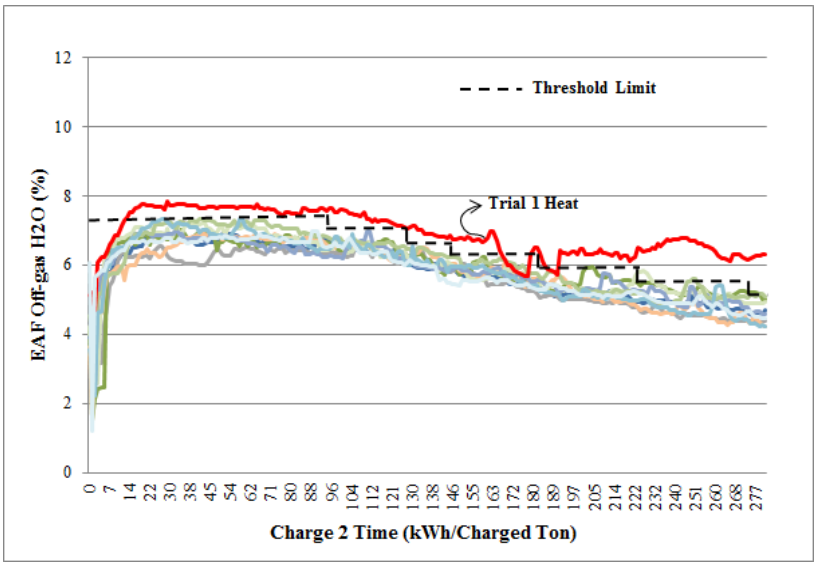

Figure 5.5: Fingerprinting Method during Charge 2 for EAF 1.

The off-gas concentration measurement of water vapor provided adequate information for EAF 1 to distinguish between normal heats and heats where significant additional water is injected into the furnace. The fingerprinting method is capable of detecting trial 1 heat where 60 liters per minute is injected, but the method is incapable of detecting the 30 liters per minute trial.

5.2.3 EAF 2 Statistical Fingerprinting Results

Figure 5.6 shows the off-gas water vapor for the training heats for EAF 2. The maximum range of variability for the off-gas water vapor is 4% for the entire heat. Such variability range is typically not wide, and therefore fingerprinting method can be tested on this furnace. Similar to EAF 1, the set of heats for EAF 2 that are used to compute the threshold limits and validate the model are determined in Chapter 4. Moreover, the testing heats that are used to validate the model are two trial heats where water is intentionally injected into the furnace. Figure 5.6 shows the off-gas water vapor, the bins, and the threshold limits. The eight bins shown in Figure 5.6 are determined manually, where in each bin the dynamic of the water vapor measurement is similar between the heats, and the bin value is calculated using Equations 5.1 and 5.2. The heat

time shown in Figures 5.6-5.7 is DRI Fed (%) which is the energy clock used by EAF 2 to pace the heat.

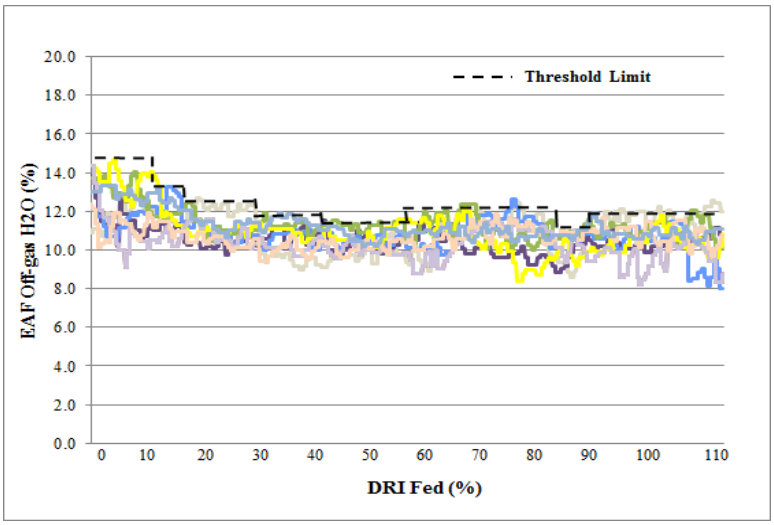

Figure 5.6: Fingerprinting Method (Normal Heats) for EAF 2.

In this furnace, water is injected for few minutes by increasing the electrode spray water flow rate. In Figure 5.7, trial 1 heat (red) shows a heat where the electrode spray water is increased by 30 liters per minute for 6 minutes. The red curve in Figure 5.7 increases and passes the fingerprint threshold in the beginning, but it drops towards the end of the trial period, and hence in this case this method is incapable of detecting the additional injected water for the entire trial period. Trial 2 heat (purple) shows the off-gas water vapor for a heat where the electrode spray water is increased by 60 liters per minute for also 10 minutes. The purple curve in Figure 5.7 passes through the fingerprint threshold from the beginning of the trial period until end, but it drops once below the threshold limit hence the fingerprinting method is capable of detecting the additional water increase in trial 2.

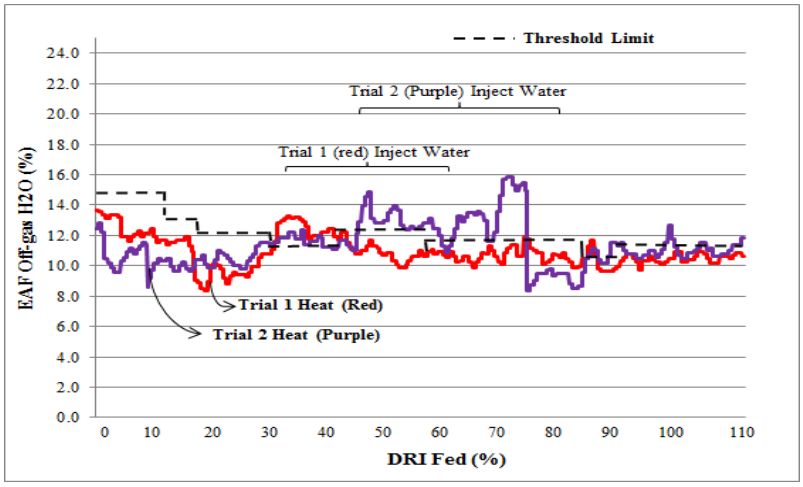

Figure 5.7: Fingerprinting Method (Trial Heats) for EAF 2.

Similar to EAF 1, the off-gas water vapor in EAF 2 provides adequate information to distinguish between normal heats and heats where water is intentionally injected. It is evident from both industrial furnaces that the fingerprinting method worked well with larger injection of water flow rates (i.e. 60 liters per minute) but the method did not detect the smaller leaks (i.e. 30 liters per minute) for the entire trial period.

The fingerprinting method is simplistic in nature; however, there are disadvantages with this method: it is incapable of detecting smaller leaks (e.g. 30 liters per minute), and it only works if the process is well controlled where in most electric arc furnaces is not the case. Therefore, the next section investigates machine learning methods to overcome the problems with the fingerprinting method. Machine learning methods explored in this work are artificial neural network and multiway projection to latent structures which use process variables to calculate expected off-gas water vapor leaving the electric arc furnace.

5.3 Artificial Neural Network (ANN)

5.3.1 Introduction

Artificial neural network (ANN) is a machine learning method that is used widely for classification and regression. ANN is used for classification if the learning technique used is unsupervised and it is used for regression if the learning technique is supervised. In this work, supervised learning is used to build the ANN model to predict EAF off-gas water vapor based on different process variables. The steps followed to build the ANN model to predict the EAF off-gas water vapor are: collect and preprocess the data, create the network, configure the network, train the network, and validate the network. The set of process variables and heats used to build and validate the ANN model are determined in Chapter 4. The selected set of heats is divided into training and testing heats. The training heats are used to train the ANN model, and the testing heats are used to validate the ANN model, and they include heats where water is intentionally injected into the furnace to test if the ANN model can recognize the water addition.

5.3.2 Data Preprocessing and Artificial Neural Network Configuration

The first step is to collect and preprocess the network inputs and outputs. Such a step allows the inputs and outputs to fall into a normalized range. The preprocessing functions performed in this work normalize the inputs and the output to fall in the range of [-1, 1], and remove inputs and outputs that are constants. The next steps completed are to create and then configure the network. The network created for this work is a multi-layer feed forward networks. Typically, the first layer is the input layer; then one or more hidden layers are constructed, and finally an output layer is used. Different types of transfer functions can be used in the hidden and output layers. In this work, nonlinear transfer functions are used in the hidden layers because they allow the network to learn nonlinear relationships between the input and the output vectors, and linear transfer

functions are used in the output layer because they are often used for function fitting problems (Nielsen, 2015).

5.3.3 Artificial Neural Network Training

Artificial neural networks are organized in layers. Layers consist of interconnected neurons, and the connections have weights. The first layer in the network has the input neurons which communicate with the hidden layers where the processing occurs by tuning the connections weights, and then the hidden layers communicate with the output layers. Neural networks weights (w_{ij}) are modified so that a specific set of inputs x_i lead to an output f(x). The networks used in this work are feedforward neural networks where the output from one layer is used as an input to the next layer; hence there are no network loops. Figure 5.8 shows a general architecture of ANN in this work:

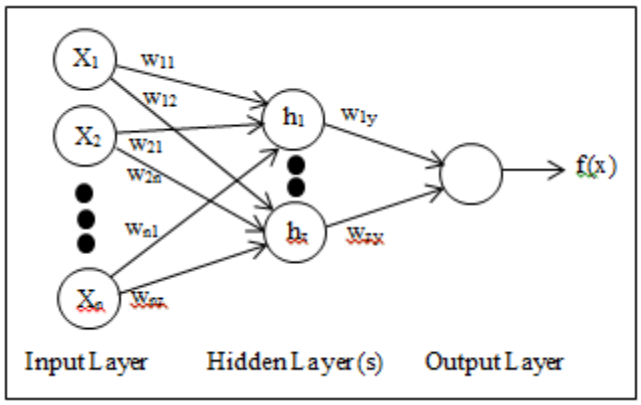

Figure 5.8: General Architecture of ANN Constructed in this Work.

Each layer in the neural network has a weight matrix (w^l) and a bias vector (b^l). The weights in the weight matrix are weights connecting to the l^{th} layer of neurons and each neuron in the l^{th} layer has one bias. Equation 5.3 defines the activation vector a^l , where σ is the activation function and z^l is the weighted input to the neurons in layer l:

$$a^l = \sigma(w^l\ a^{l-1} + b^l) = \sigma(z^l) \tag{5.3}$$

ANN finds the model parameters (i.e. weights and biases) using a learning algorithm. The learning algorithm used in this work is the feed forward back propagation algorithm. The feed forward back propagation network is an algorithm where errors propagate backward from the output layer during training. Back propagation is a gradient method to minimize the total squared error of the output computed by the neural network (Nielsen, 2015). Training a network by back propagation involves three stages: the feed forward of the input training pattern, the backward propagation of the associated error, and the adjustment of weights (Nielsen, 2015). The output errors determine measures of hidden layer output errors. The process of adjusting the set of weights between the layers and recalculating the output continues until the overall error falls below a given limit. Once training is completed, the ANN can be used to compute off-gas water vapor for new input variables.

Equations 5.4-5.8 summarize the back propagation algorithm (Nielsen, 2015). Mathematically, the objective of the algorithm is to find the network weights (w) and biases (b) to minimize the quadratic cost function $C\,(w,b)$.

$$C(w,b) = \frac{1}{2n}\sum_x \|f(x) - a\|^2 \tag{5.4}$$

In order to minimize the cost function, the back propagation algorithm calculates the partial derivatives $\partial C/\partial w$ and $\partial C/\partial b$ of the cost function with respect to any weight w or bias b in the artificial neural network (Nielsen, 2015). But to compute the partial derivatives, the algorithm introduces an intermediate error variable δ^L which is the error associated with layer L. The algorithm provides a way to compute the δ^L for every layer, and then it relates the errors to the partial derivatives $\partial C/\partial w$ and $\partial C/\partial b$. Equation 5.5 computes the error in the output layer. The first term in Equation 5.5 ($\nabla_a C$) computes the derivative of the cost function C with respect to the output activations (a). The second term ($\sigma'(z^L)$) calculates the rate of change of the activation function (σ) with respect to z^L :

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \tag{5.5}$$

Equation 5.6 calculates the error δ^L in terms of the error in the next layer δ^{L+1}. $(w^{l+1})^T$ is the transpose of the weight matrix for the (l + *1*) layer and multiplies it by the error in the (l + *1*) layer and then taking the Hadamard product $\odot\, \sigma'(z^l)$. This equation allows for the error to move backward through the neural network and hence for a mean for the error at the output of the l^{th} layer to be computed (Nielsen, 2015).

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \tag{5.6}$$

Equation 5.7 calculates how fast the quadratic cost function (*C*) is changing with respect to any bias (*b*) in the network, where the error variable δ and the bias *b* are evaluated at the same neuron in the network.

$$\frac{\partial C}{\partial b} = \delta \tag{5.7}$$

Equation 5.8 calculates how fast the quadratic cost function (*C*) is changing with respect to any weight (*w*) in the neural network, where a_{in} is the activation of the neuron input to the weight *w*, and δ_{out} is the neuron output from the weight *w* in the network.

$$\frac{\partial C}{\partial w} = a_{in}\delta_{out} \tag{5.8}$$

In this work, Equations 5.4-5.8 are used to train the neural network and to find the optimum weights and biases. During training of the EAF neural network, the parameters that are monitored are the magnitude of the gradient of performance and the number of validation checks. The magnitude of the gradient and the number of validation checks are used to terminate the training. The gradient becomes very small as the training reaches a minimum of the performance. The training of the model is stopped if the magnitude of the gradient is less than 1e-5. The number of validation checks is the number successive iterations that the validation performance fails to decrease, and for this work the number of validation checks used is 6. The number of epochs used to develop the models is

1000. The network is used to calculate the expected water in the off-gas after the network is trained and validated.

There is uncertainty in the neural network predictions due to inaccuracies in the training data and limitation of the neural network model. There is noise in the training set because noise is inherent in all industrial real data. Therefore, a reliable measure of confidence interval is crucial in this work. The method used to construct the confidence intervals for the neural network predictions is the bootstrap method (Nielsen, 2015). This method involves creating many bootstrap samples by resampling randomly. The bootstrap estimate of the standard error is giving by (Efron & Tibshirani 1993). For this work, 10 resamples are created to estimate the standard error and then to calculate the 95% confidence interval.

5.3.4 EAF 1 Artificial Neural Network (ANN) Results

The input and output variables for EAF 1 are preprocessed prior to building the ANN model. The preprocessing function used in this work normalizes the input and output variables to fall in the range of [-1, 1]; then the ANN model is constructed. The ANN model built for EAF 1 is a three layer feed forward network consisting of an input layer, a hidden layer, and an output layer. The output layer has a single node because the model is used to predict one output variable (off-gas water vapor) from multiple inputs. All nodes in the input layer are connected to all nodes in the hidden layer. Similarly, all nodes in the hidden layer are connected to the node in the output layer. The input variables used to train the ANN model for EAF 1 are determined based on the MPLS model developed in Chapter 4. The MPLS model predicts that the input variables that are highly correlated with the off-gas water vapor are EAF total fuel flow, EAF total main oxygen Flow, EAF total injected carbon flow, off-gas CO, off-gas CO_2, off-gas O_2, EAF fume system damper position, and EAF pressure. Electrode water flow is excluded from the ANN model because the electrode flow is constant from the beginning of the heat until the end. Eight nodes used in each of the input layer and hidden layer. More nodes were tried in the hidden layer but no improvement was observed in the results. A log-

sigmoid transfer function is used in the hidden layer to capture the nonlinearity of the water vapor behavior, and a linear transfer function is used in the output layer. The heats used to train and test the ANN model for EAF 1 are determined by the MPCA model developed in Chapter 4. The number of training heats is 15 heats, and the number of testing heats is 9 heats. The testing heats include 7 normal heats and 2 trial heats, where additional water is injected into the furnace from the beginning of the heat until the end. The R^2 of the ANN model built with the training heats is 90%.

Figure 5.9 shows two figures. The top figure shows the ANN prediction drawn in solid line and the confidence interval in dashed line with 95% confidence level for a normal heat with three charges and a refining period. The top figure shows that the ANN model provides a narrow prediction range as a consequence of the multiple layer and nodes network constructed for this model. The bottom figure shows a comparison between the measured EAF off-gas water vapor drawn in solid line and the off-gas water vapor predicted by the ANN model drawn in dashed line for a normal testing heat. There are dips in off-gas water vapor because whenever the operator charges the furnace with a new scrap bucket, the off-gas analyzer stops sampling and starts purging the probe and the sample line. The bottom figure clearly shows that the predicted water vapor closely follows the measured water vapor throughout the heat.

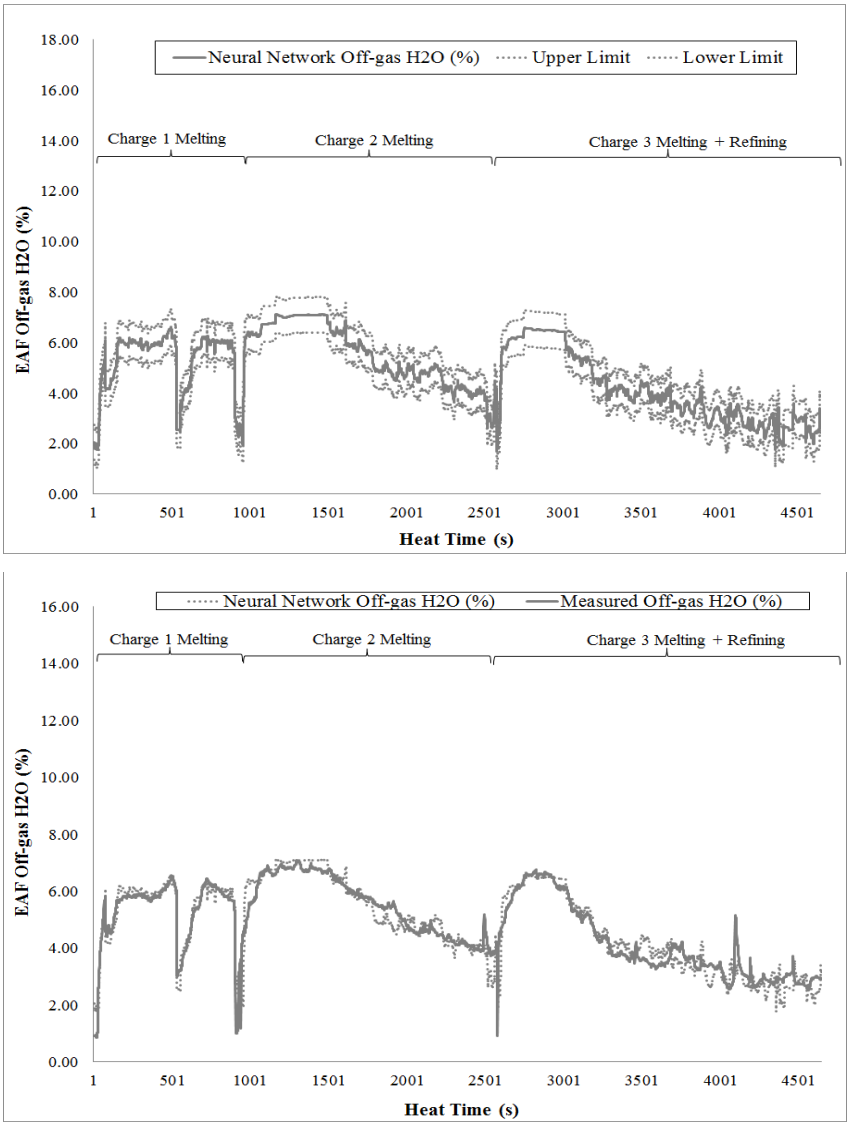

Figure 5.9: Top Figure – ANN Model Prediction (solid line) and the 95% Confidence Interval (dashed lines). Bottom Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and ANN Prediction (dashed line). Both Figures are for the Same Normal Testing Heat for EAF 1.

Figure 5.10 shows two figures. The top figure shows the error (%) which is calculated by subtracting the calculated EAF off-gas water vapor from the measured water vapor. The figure clearly shows that the error is below the 1% threshold for most of the heat, which indicates that the predicted water vapor closely followed the measured water vapor. The bottom figure shows the errors are roughly normally distributed and approximately centered at 0.

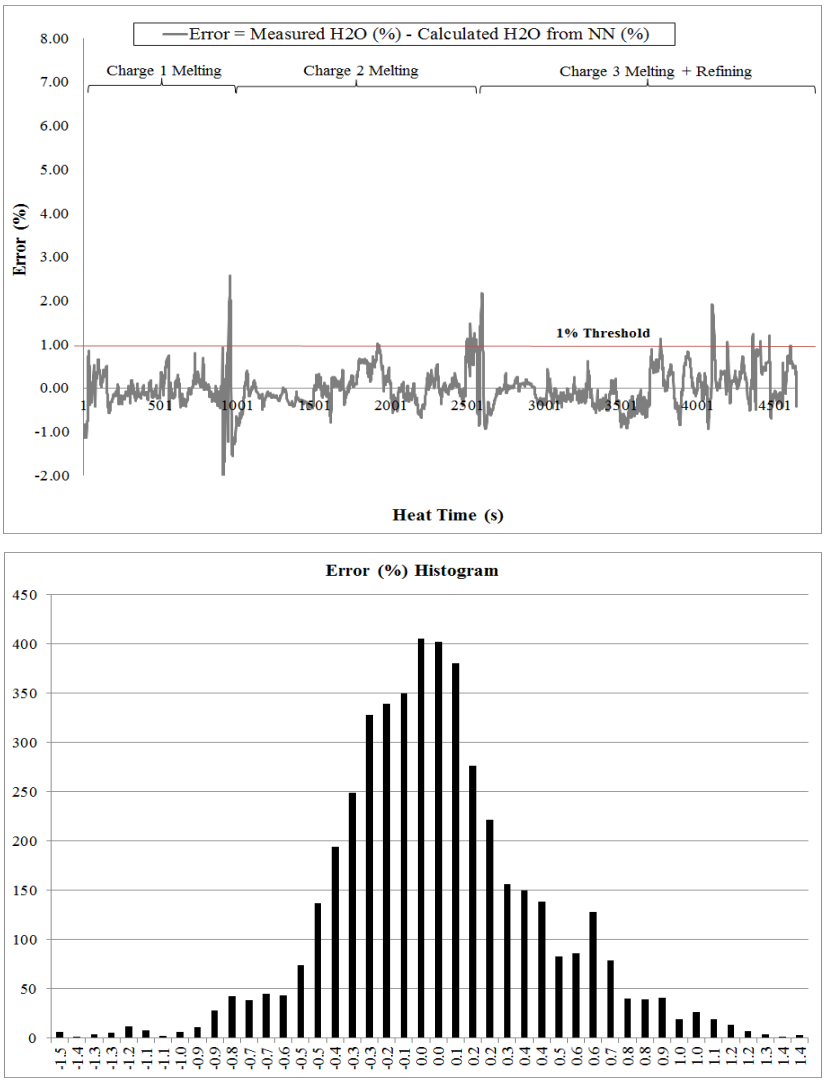

Figure 5.10: Top Figure – Error (%) Between Measured and the Calculated EAF Off-gas H_2O. Bottom Figure – Normal Distribution of the Error (%). Both Figures are for the Same Normal Testing Heat for EAF 1.

Figure 5.11 shows the ANN prediction drawn in solid line and the 95% in dashed line for a trial testing heat with four charges and a refining period. The figure shows that the ANN model provides a narrow prediction range as a consequence of the multiple layers and nodes network constructed for this model.

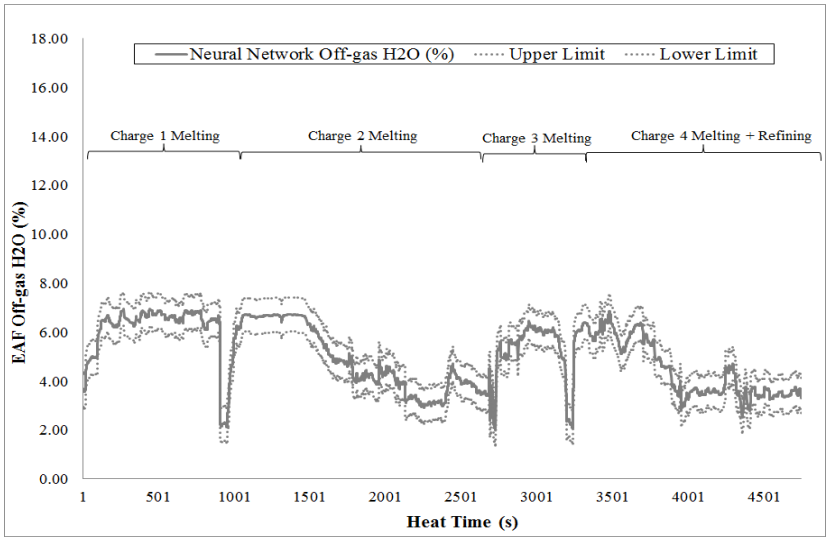

Figure 5.11: ANN Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 1 Testing Heat for EAF 1.

Figure 5.12 shows two figures for trial 1 testing heat. The top figure shows a comparison between the EAF off-gas water vapor measured by the off-gas analyzer and the off-gas water vapor calculated by the ANN model. Trial 1 was conducted by increasing the electrode spray water by a total of 60 liters per minute from the beginning of the heat until the end. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor. The bottom figure shows the error (%) between the measured and the calculated EAF off-gas water vapor. The measured water vapor is constantly higher than the calculated water vapor because the ANN model does not include water leaks as an input to the model. In this trial, water is intentionally added into the furnace to observe if a difference can be observed between the measured and the calculated water vapor. The bottom figure shows a continuous difference of greater than 1% between measured and predicted from the beginning of the heat until tapping. The error threshold is a tuning parameter, where in this case 1% is selected because it captures the water leak for both EAF 1 trial heats.

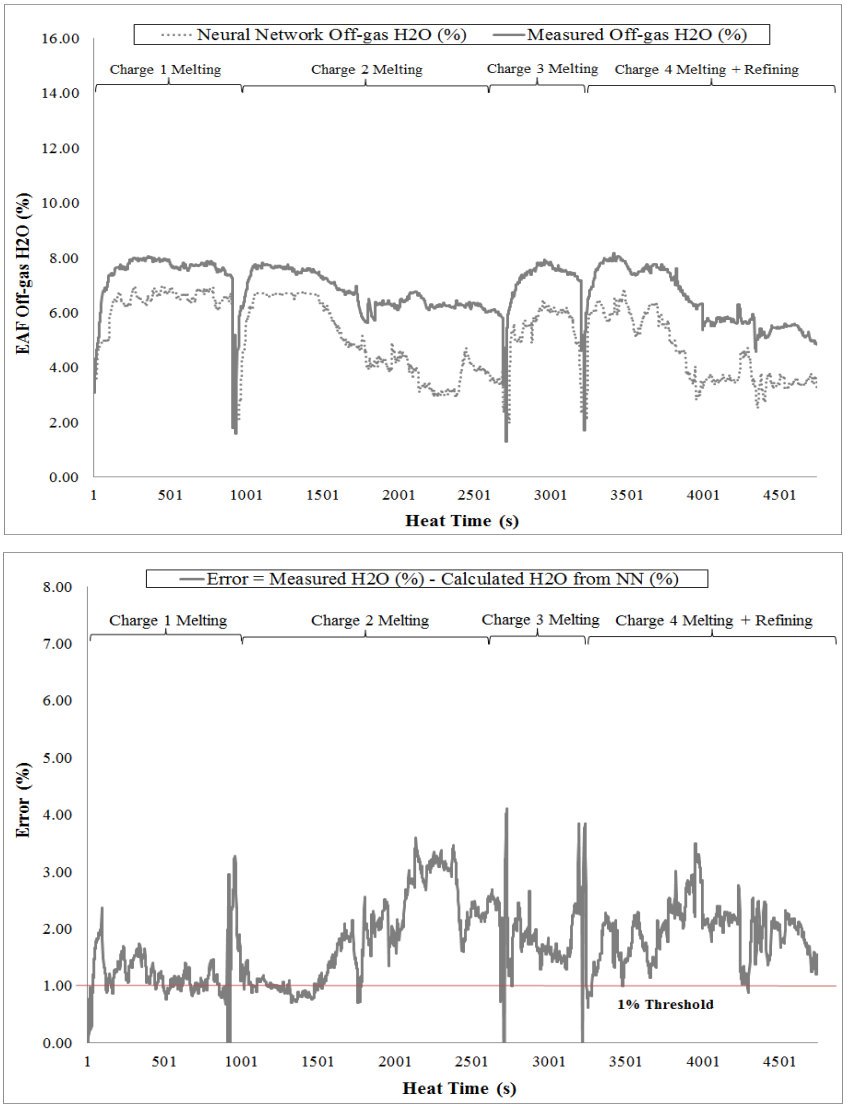

Figure 5.12: Top Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and ANN Prediction (dashed line). Bottom Figure – Error (%) Between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 1 Testing Heat for EAF 1.

Figure 5.13 shows the ANN prediction drawn in solid line and the 95% confidence interval level for a charge 1 of trial 2 testing heat, where water is intentionally injected only in the first charge. Similar to trial 1 testing heat, the figure shows that the ANN model provides a narrow prediction range as a consequence of the multiple layers and nodes network constructed for this model.

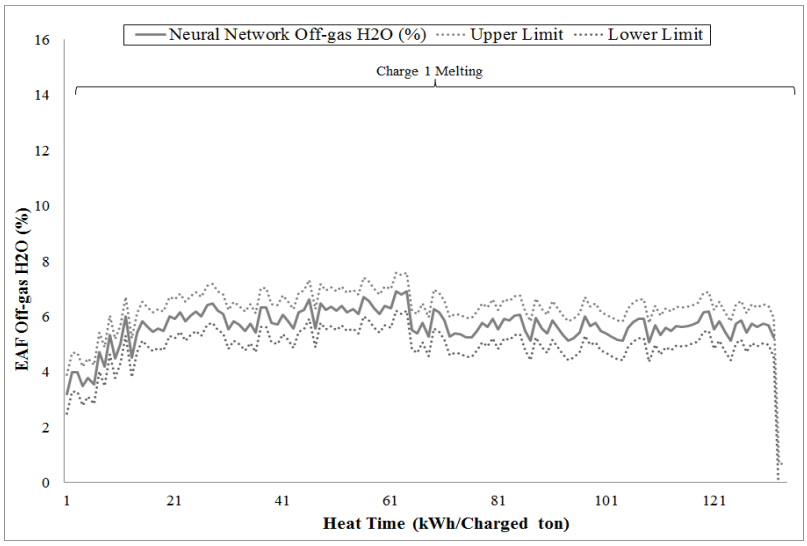

Figure 5.13: ANN Predictions (solid line) and the 95 % Confidence Interval (dashed line) for Trial 2 Testing Heat for EAF 1.

Figure 5.14 shows two figures for trial 2 testing heat. The top figure shows a comparison between the measured EAF off-gas water vapor and the off-gas water vapor (%) calculated by the ANN model. Trial 2 testing heat is conducted by increasing the electrode spray water by a total of 30 liters per minute during the entire first charge. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor. The bottom figure shows the error between the measured and the calculated EAF off-gas water vapor. The error is above the 1% for most of the charge 1 melting. However, it drops below the 1% threshold several times. Similar to trial 1 heat, the reason that the measured water vapor is constantly higher than the calculated is because the ANN model does not include water leaks as an input to the model.

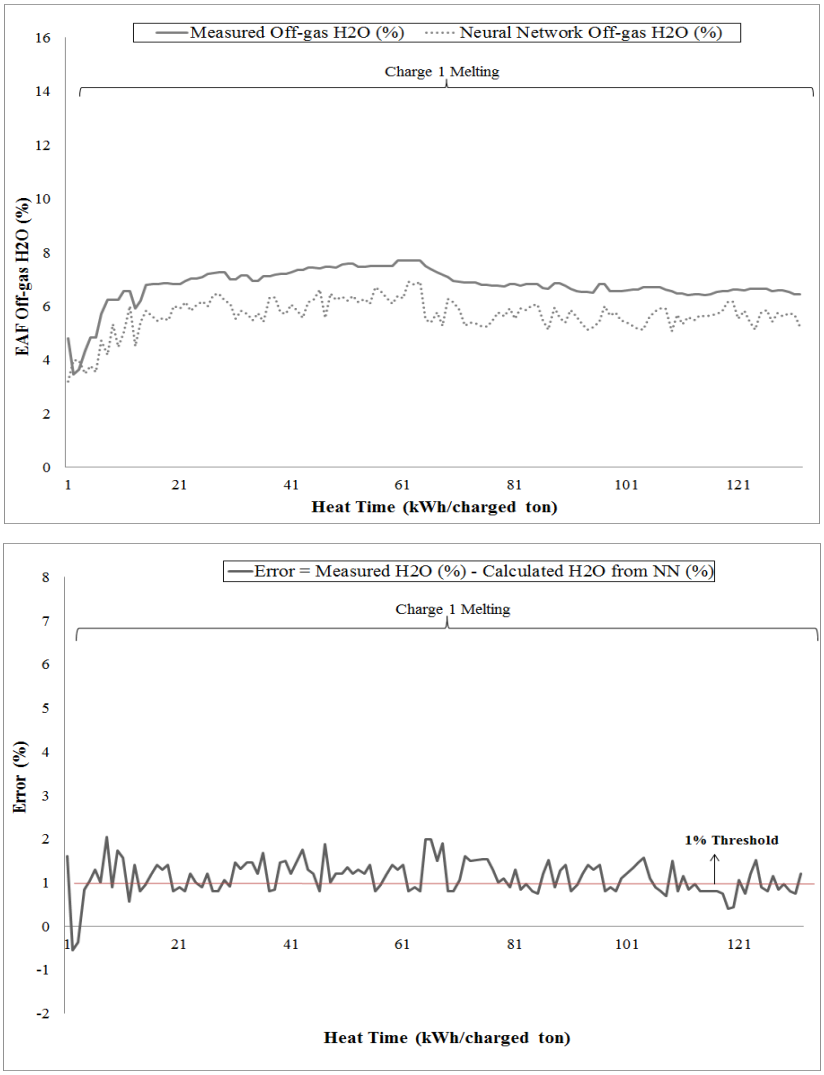

Figure 5.14: Top Figure - Comparison between Measured EAF Off-gas H2O (solid line) and ANN Prediction (dashed line). Bottom Figure – Error (%) Between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 2 Testing Heat for EAF 1.

Figure 5.15 shows the residuals from the ANN model for the normal testing heats for EAF 1. This error histogram excludes the trial testing heats. The residuals shown in Figure 5.15 are estimates of the experimental error determined by subtracting the measured off-gas water vapor from the predicted off-gas water vapor. Residuals are variability unexplained by the ANN model. Figure 5.15 shows an error distribution that is roughly normal and is centered at approximately 0 with a standard deviation of about 0.47. The significant of the residuals normal distribution being centered at 0 indicates that the model error is random.

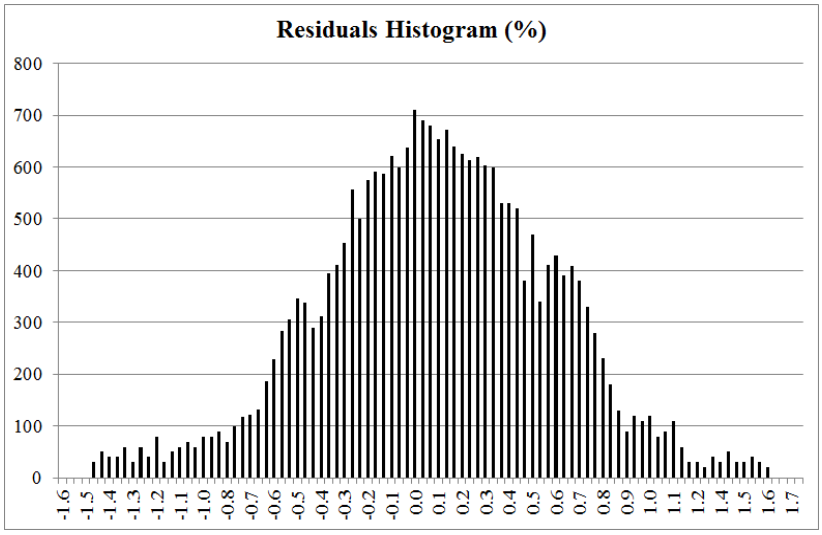

Figure 5.15: Residuals Histogram from the ANN Model for all the Normal Testing Heats for EAF 1.

5.3.5 EAF 2 Artificial Neural Network (ANN) Results

The steps followed to preprocess the input and output variables for EAF 2 and then build the ANN model are similar to EAF 1. The input variables used to train the ANN model for EAF 2 are determined from the MPLS model developed in Chapter 4. The MPLS model predicted the input variables that are highly correlated with the off-gas H_2O are EAF total fuel flow, EAF total main oxygen flow, EAF total shroud oxygen flow, DRI flow, EAF total injected carbon flow, off-gas CO, off-gas CO_2, electrode water flow, and EAF analyzer purge H_2O. The number of nodes used in the input layer and hidden layer are 9 nodes each. More nodes were tried in the hidden layer, but no improvement was observed in the performance of the model. The transfer function used in the hidden layer is a log-sigmoid to capture the nonlinearity of the water vapor behavior, and a linear transfer function is used in the output layer. The heats used to train and test the ANN for EAF 2 are determined from the MPCA model developed in Chapter 4. The number of training heats is 35 heats, and the number of testing heats is 15 heats. The testing heats include 13 normal heats and 2 trial heats, where additional water is injected into the furnace. The R^2 of the ANN model built with the training heats is 91%.

Figure 5.16 shows two figures. The top figure shows the ANN prediction drawn in solid line and the 95% confidence interval for a normal heat with DRI feeding and a refining period. The figure shows that the neural network model provides a narrow prediction range as a consequence of the multiple layers and nodes network constructed for this model. The bottom figure shows a comparison between the EAF off-gas water vapor measured by the analyzer drawn in solid line and the neural network off-gas water vapor prediction drawn in dashed line for a normal heat for EAF 2. Figure 5.16 shows that the predicted water vapor closely follows the measured water vapor throughout the heat.

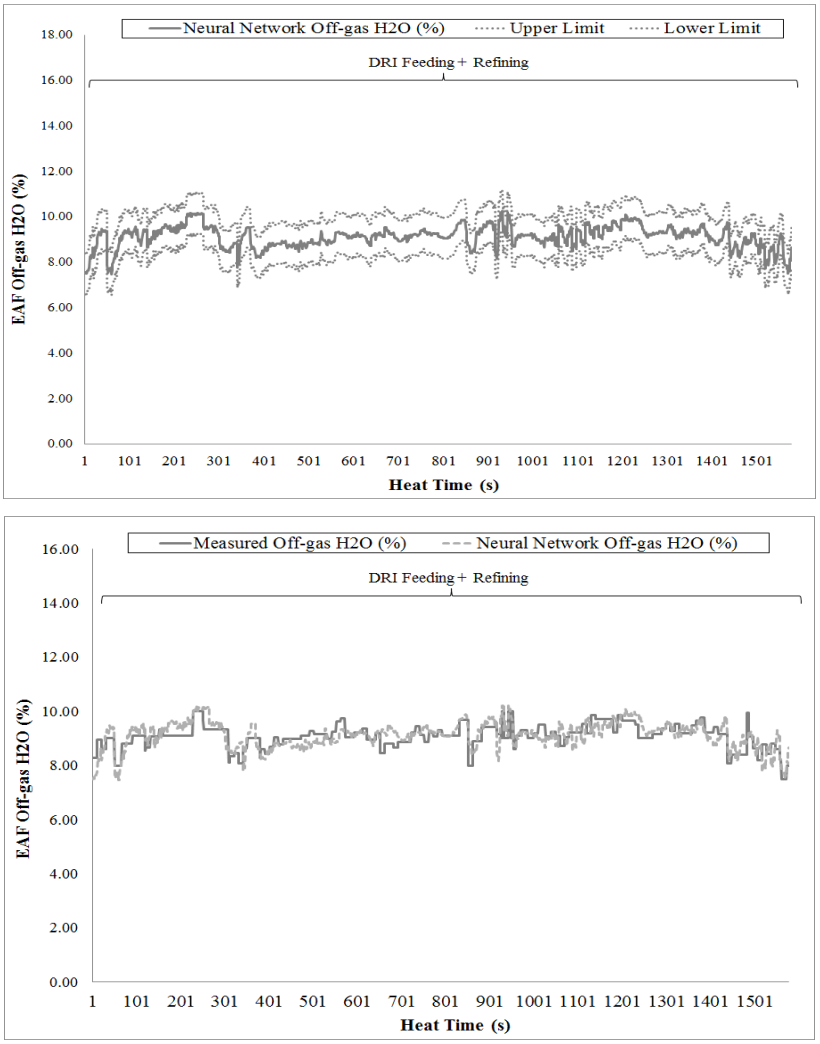

Figure 5.16: Top Figure - ANN Predictions (solid line) and the 95% Confidence Interval (dashed line). Bottom Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and ANN Prediction (dashed line). Both Figures are for the Same Normal Testing Heat for EAF 2.

Figure 5.17 shows two figures. The top figure shows the error calculated by subtracting the calculated EAF off-gas water vapor from the measured water vapor. The figure clearly shows that the error is below the 1% threshold for most of the heat, a finding which indicates that the predicted water vapor closely followed the measured water vapor. The bottom figure shows the errors are roughly normally distributed and centered at approximately 0.

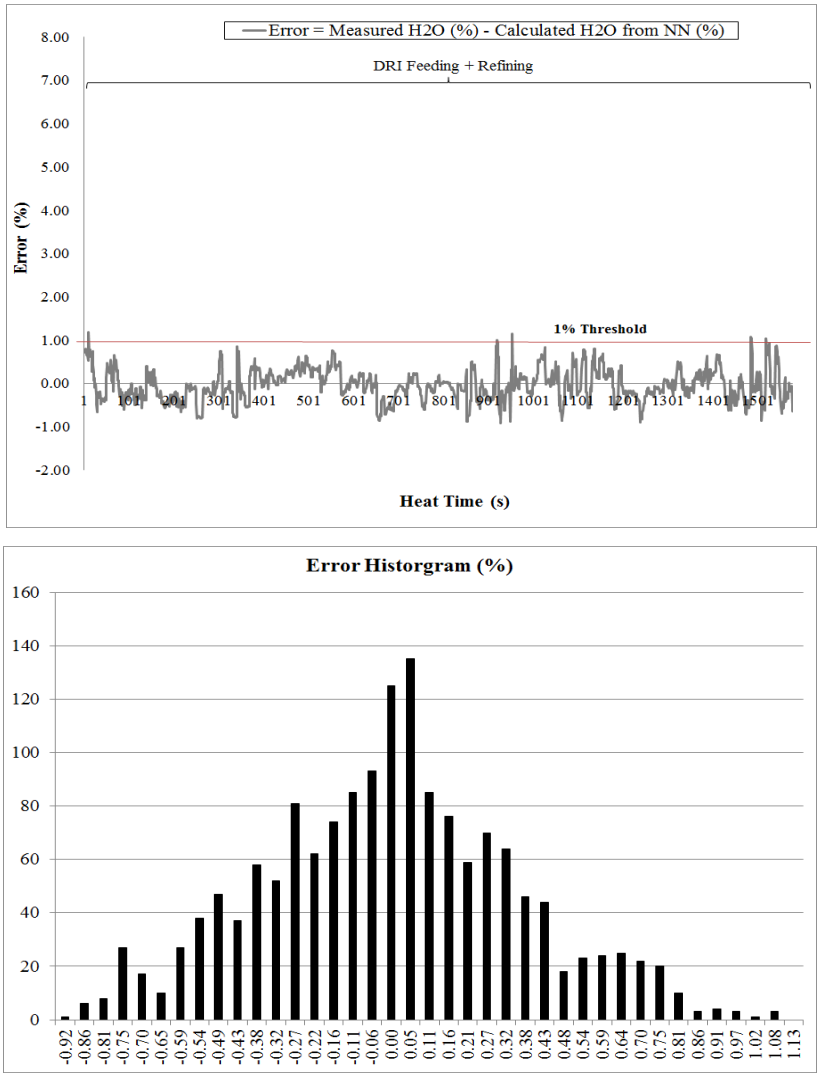

Figure 5.17: Top Figure – Error (%) Between Measured and the Calculated EAF Off-gas H_2O. Bottom Figure – Normal Distribution of the Error (%). Both Figures are for the Same Normal Testing Heat for EAF 2.

Figure 5.18 shows the ANN model prediction drawn in solid line and the 95% confidence interval for trial 1 heat with DRI feeding and a refining period. The figure shows that the ANN model provides a narrow prediction range as a consequence of the multiple layers and nodes network constructed for this model.

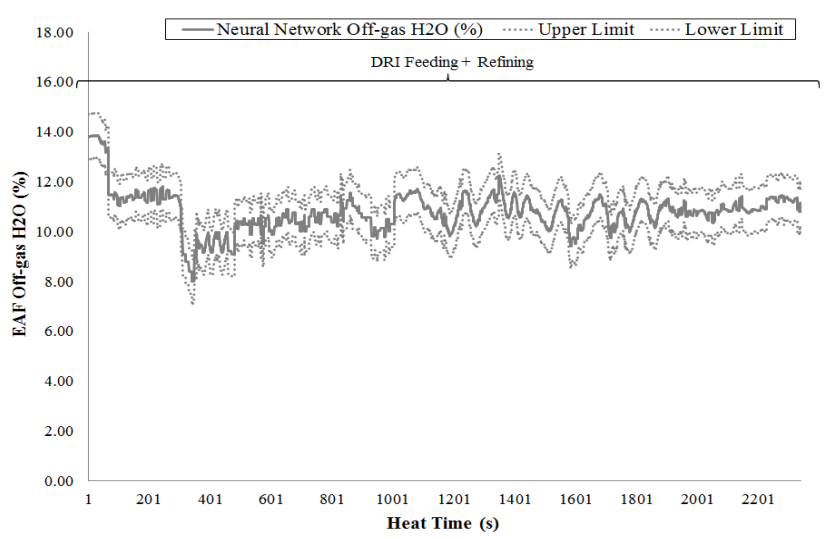

Figure 5.18: ANN Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 1 Testing Heat for EAF 2.

Figure 5.19 shows two figures for trial 1 heat for EAF 2. The top figure between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by the ANN model. Trial 1 heat is conducted by increasing the electrode spray water by a total of 30 liters per minute for approximately 6 minutes. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor by more than 1-3% during the trial test. The bottom figure shows the error between the measured and the calculated EAF off-gas water vapor. As in trials heat from EAF 1, the reason that the measured water vapor is higher than the calculated is because the ANN model does not include water leaks as an input to the model.

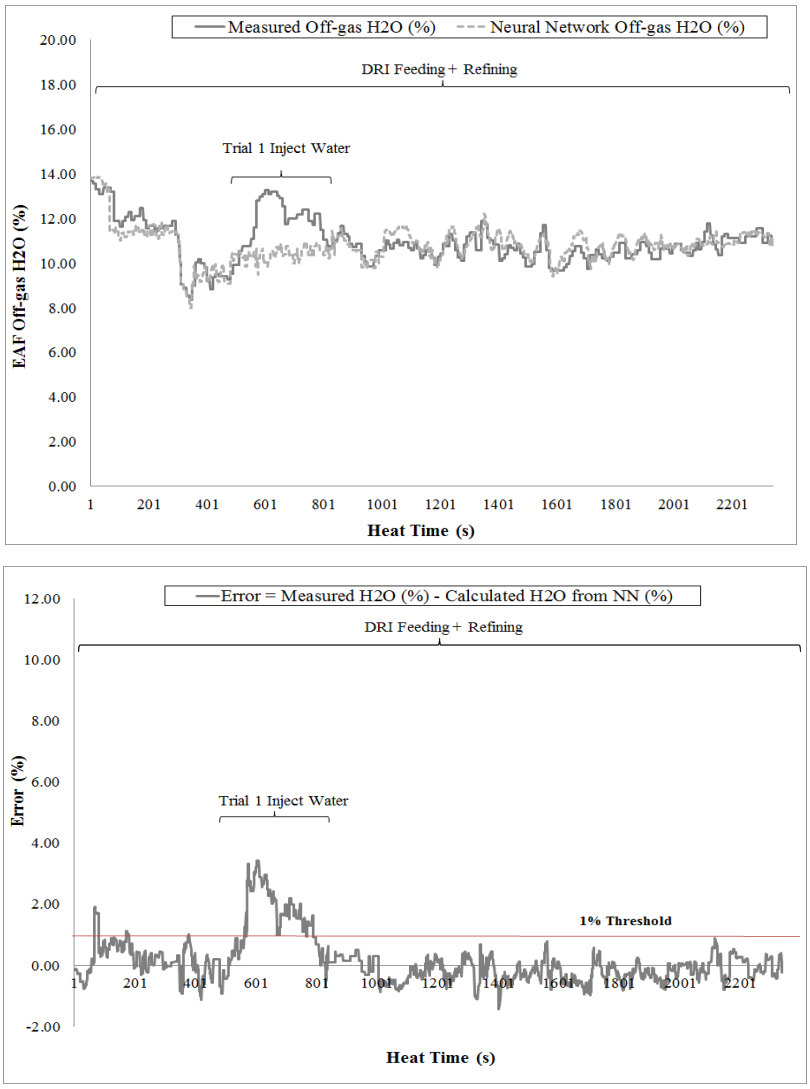

Figure 5.19: Top Figure - Comparison Between Measured EAF Off-gas H_2O (solid line) and ANN Prediction (dashed line). Bottom Figure – Error Between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 1 Testing Heat for EAF 2.

Figure 5.20 shows the ANN prediction drawn in solid line and the 95% confidence interval in dashed line for trial 2 testing heat with DRI feeding and a refining period. The figure shows that the ANN model provides a narrow prediction range as a consequence of the multiple layers and nodes network constructed for this model.

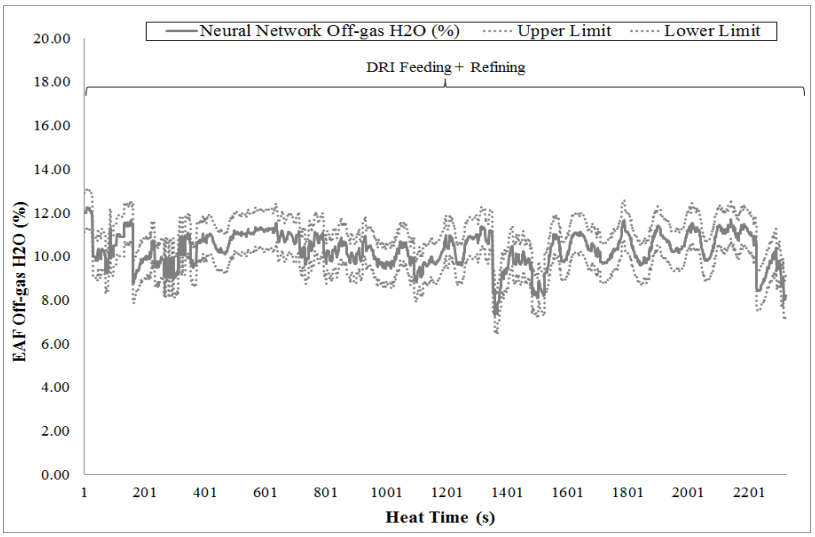

Figure 5.20: ANN Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 2 Testing Heat for EAF 2.

Figure 5.21 shows two figures for trial 2 testing heat for EAF 2. The top figure shows a comparison between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by the ANN model. Trial 2 heat is conducted by increasing the electrode spray water by a total of 60 liters per minute for approximately 10 minutes. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor by more than 1-4% during the trial test. The bottom figure shows the error between the measured and the calculated EAF off-gas water vapor. As in Trials Heat from EAF 1, the reason that the measured water vapor is higher than the calculated is because the ANN model does not include water leaks as an input to the model.

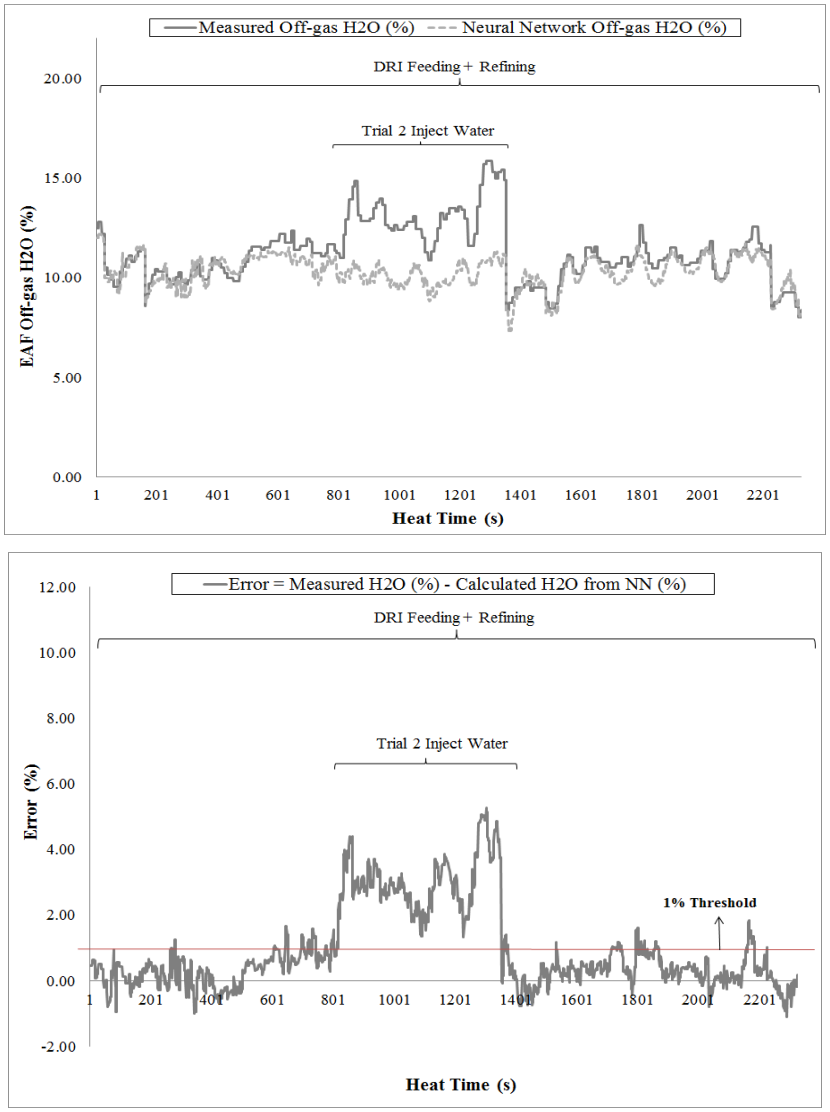

Figure 5.21: Top Figure - Comparison Between Measured EAF Off-gas H_2O (solid line) and ANN Prediction (dashed line). Bottom Figure – Error (%) Between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 2 Testing Heat for EAF 2.

Figure 5.22 shows the residuals from the ANN model for the normal testing heats for EAF 2. This error histogram excludes the trial testing heats. The residuals shown in Figure 5.22 are estimates of the experimental error determined by subtracting the measured off-gas water vapor from the predicted off-gas water vapor. Figure 5.22 shows an error distribution that is roughly normal and is centered at approximately 0 with a standard deviation of approximately 0.44. The significance of the residuals normal distribution being centered at 0 indicates that the model error is random.

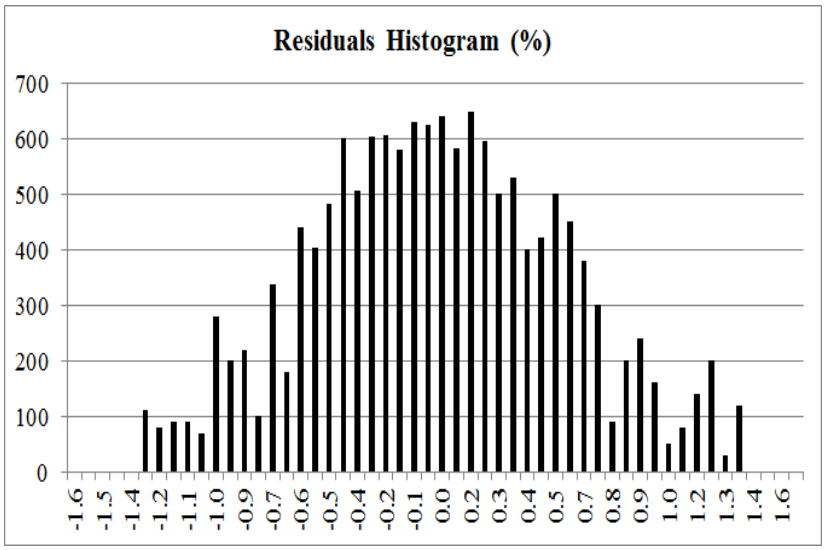

Figure 5.22: Residuals Histogram from the ANN Model for all the Normal Testing Heats for EAF 2.

5.4 Multiway Projection to Latent Structures (MPLS)

5.4.1 Method

The second machine learning method used to predict water vapor leaving the furnace is multiway projection to latent structures (MPLS). The algorithm of the method is described in details in Chapter 4. As in artificial neural network, confidence intervals are necessary due to the noise inherent in the industrial data; hence the bootstrap method is used to construct the 95% confidence intervals (Dunn, 2014). For this work, 10 resamples are created to estimate the standard error and then to calculate the 95% confidence interval.

5.4.2 EAF 1 MPLS Results

The input variables used to train the MPLS model for EAF 1 are determined in the variables selection model in Chapter 4: EAF total fuel flow, EAF total main oxygen Flow, EAF total injected carbon flow, off-gas CO, off-gas CO2, off-gas O2, EAF fume system damper position, and EAF pressure. The number of training heats is 15 heats, and

the number of testing heats is 9 heats. The testing heats include 7 normal heats and 2 trial heats where additional water is injected into the furnace from the beginning of the heat until the end. The numbers of principal components obtained are 7 components using cross validation. The calculated model R^2 is 78%. The first two components explain approximately 61% of the variability in the 15 heats. The remaining 5 components in the MPLS model explain approximately 17 % of the variability in the dataset.

Figure 5.23 shows two figures. The top figure shows the MPLS prediction drawn in solid line and the confidence interval in dashed line with 95% confidence level for a normal heat with three charges and a refining period. The figure shows that the MPLS model provides a narrow prediction range. The bottom figure shows a comparison between the EAF off-gas water vapor measured by the analyzer drawn in solid line and the MPLS off-gas water vapor prediction drawn in dashed line for a normal heat for EAF 1. The bottom figure shows that the predicted water vapor follows the measured water vapor trend throughout the heat. The heat time is normalized from seconds to %.

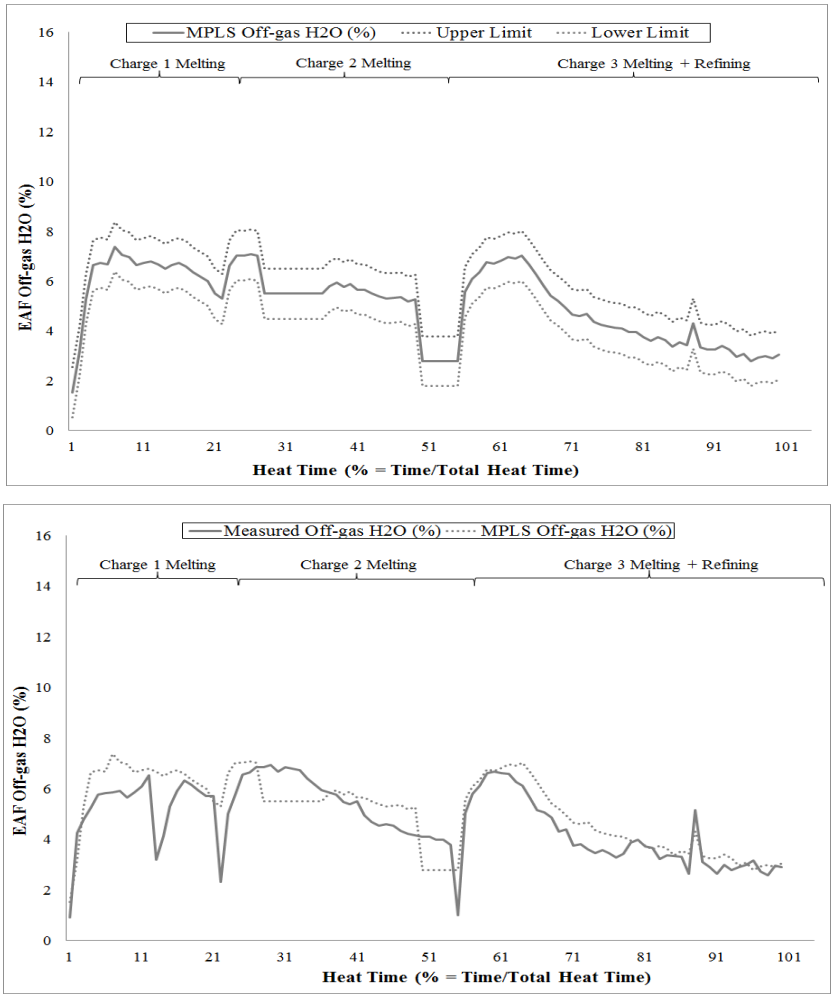

Figure 5.23: Top Figure - MPLS Predictions (solid line) and the 95% Confidence Interval (dashed line). Bottom Figure - Comparison Between Measured EAF off-gas H_2O (solid line) and MPLS Prediction (dashed line). Both Plots are for the same Normal Testing Heat for EAF 1.

Figure 5.24 shows two figures. The top figure shows the error (%) which is calculated by subtracting the calculated EAF off-gas water vapor from the measured water vapor. The figure clearly shows that the error is below the 1% threshold for most of the heat, a finding which indicates that the predicted water vapor closely followed the measured water vapor. However, there are two periods in the heat where the error is more than the 1% threshold where the ANN predictions did not have such periods. The bottom figure shows the errors are normally distributed and centered at 0.

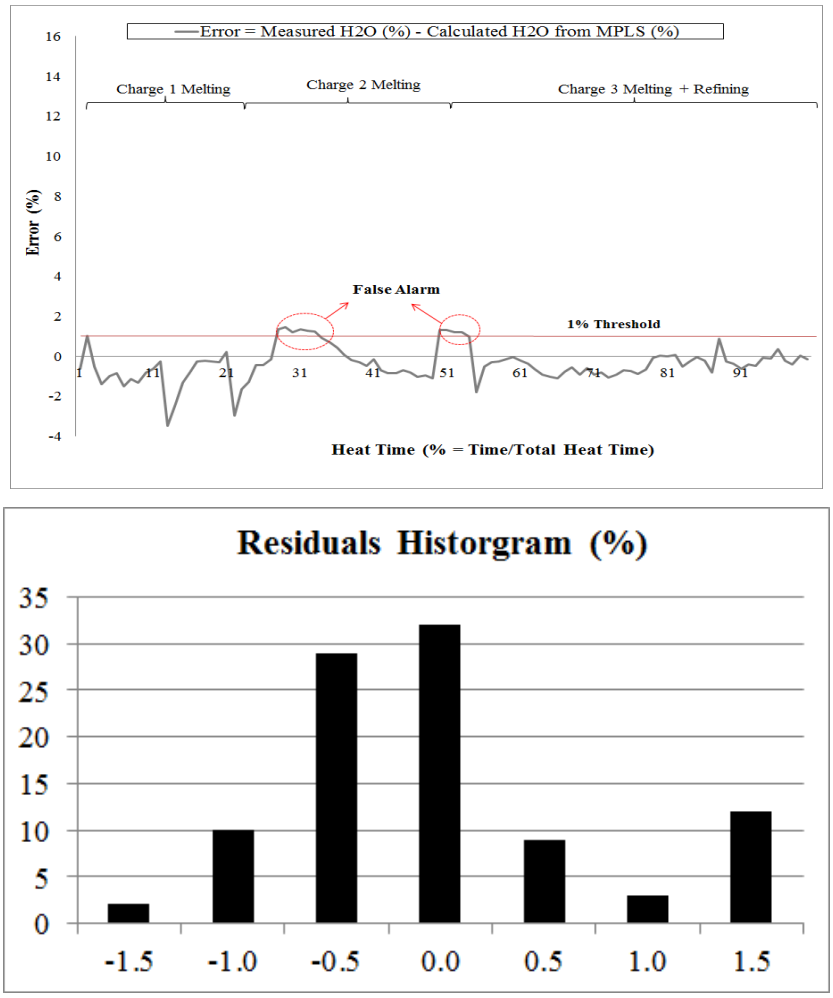

Figure 5.24: Top Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O. Bottom Figure – Normal Distribution of the Error (%). Both Figures are for the Same Normal Testing Heat for EAF 1.

Figure 5.25 shows the MPLS prediction drawn in solid line and the 95% confidence interval in dashed line for trial 1 heat with four charges and a refining period. The figure shows that the MPLS model provides a narrow prediction range.

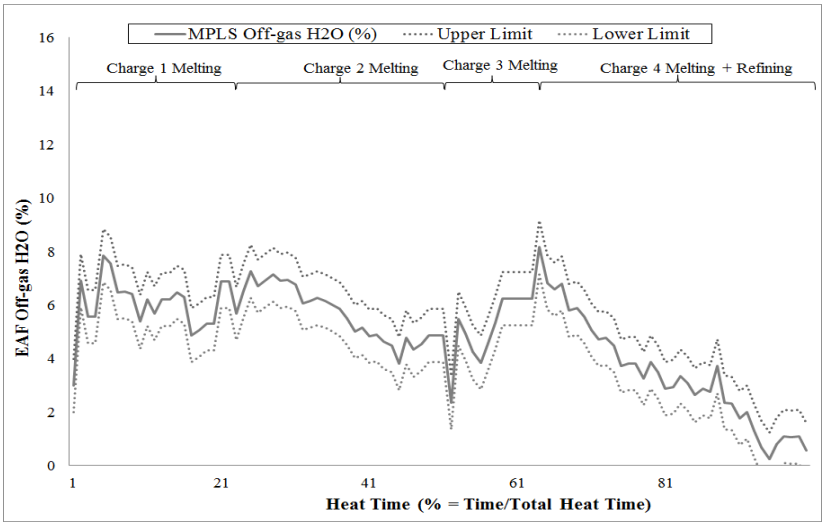

Figure 5.25: MPLS Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 1 Testing Heat for EAF 1.

Figure 5.26 shows two figures for trial 1 heat for EAF 1. The top figure shows a comparison between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by MPLS. Trial 1 heat is conducted by increasing the electrode spray water by a total of 60 liters per minute throughout the heat. The top figure shows that the measured water vapor is higher than the calculated water vapor for most of the heat. The bottom figure shows the error between the measured and the calculated EAF off-gas water vapor. The error is more than 1 % except in a period in the second charge where it drops below 1%.

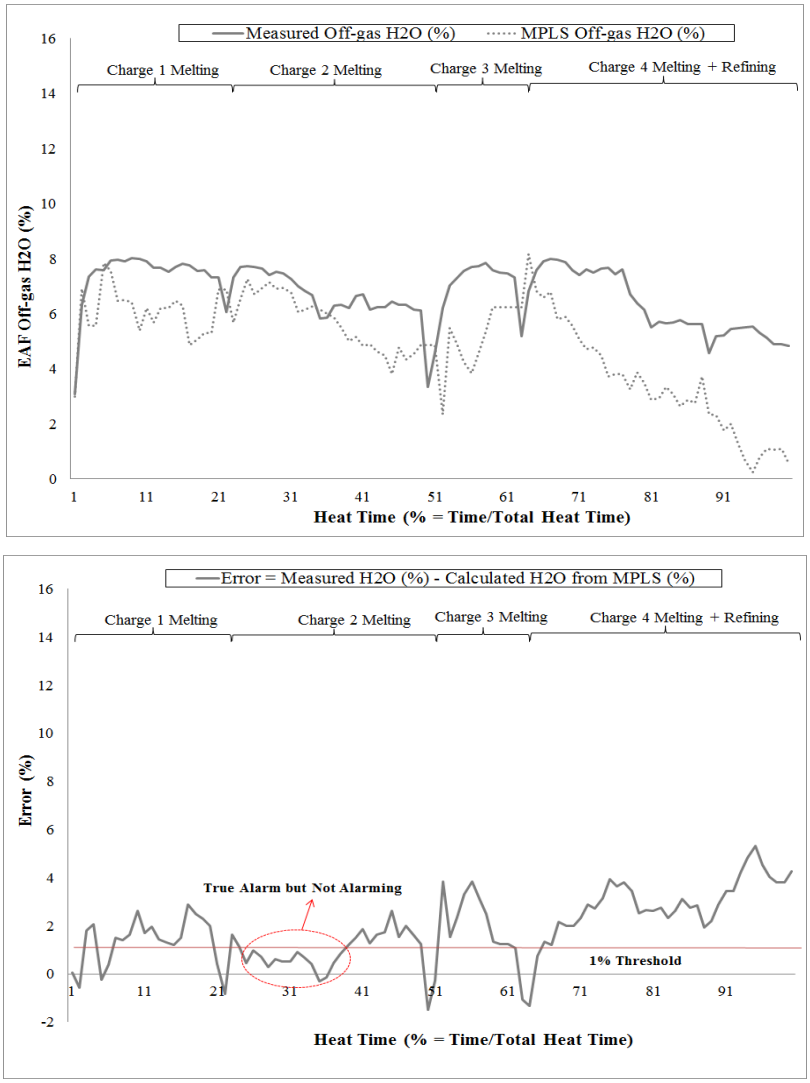

Figure 5.26: Top Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and MPLS Prediction (dashed line). Bottom Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 1 Testing Heat for EAF 1.

Figure 5.27 shows the MPLS prediction drawn in solid line and the confidence interval in dashed line with 95% confidence level for trial 2 heat. The figure shows that the MPLS model provides a narrow prediction range.

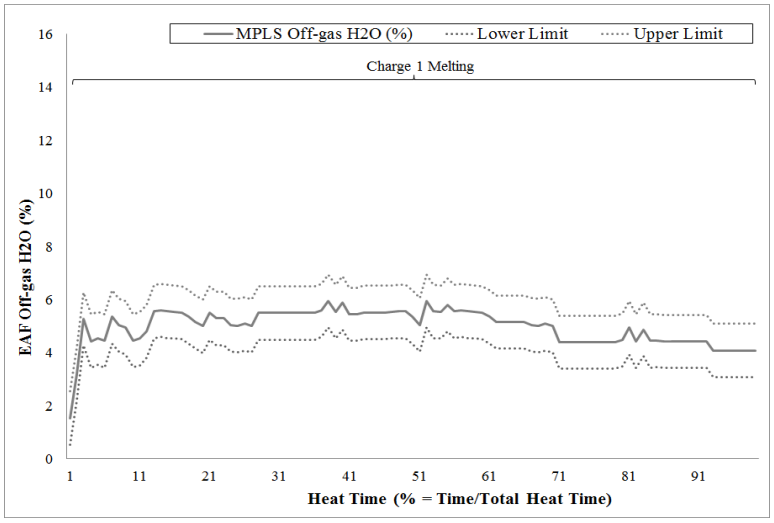

Figure 5.27: MPLS Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 2 Testing Heat for EAF 1.

Figure 5.28 shows two figures for trial 2 heat for EAF 1. The top figure shows a comparison between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by MPLS. Trial 2 heat is conducted by increasing the electrode spray water by a total of 30 liters per minute for the first charge. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor. The bottom figure shows the error (%) between the measured and the calculated EAF off-gas water vapor. As in trial 1 Heat, the reason that the measured water vapor is constantly higher than the calculated is because the MPLS model does not include water leaks or abnormal sources of water as an input to the model.

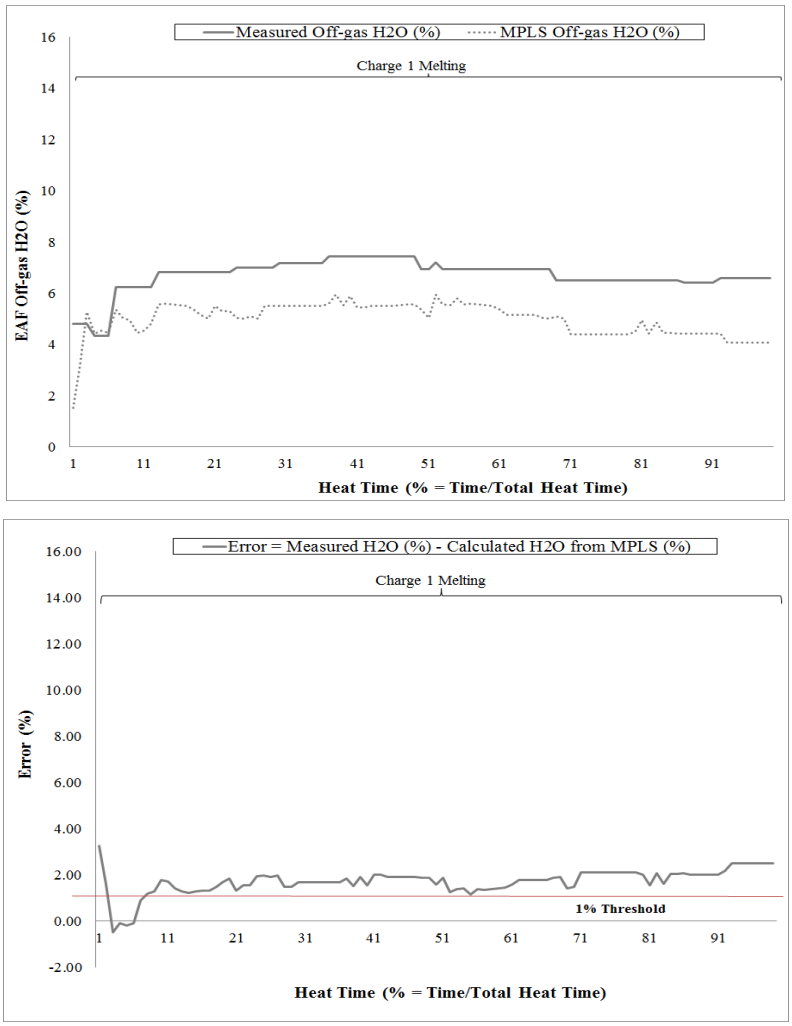

Figure 5.28: Top Figure - Comparison between Measured EAF off-gas H_2O (solid line) and MPLS prediction (dashed line). Bottom Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 2 Testing Heat for EAF 1.

Figure 5.29 shows the residuals from the MPLS model for the normal testing heats for EAF 1. This error histogram excludes the trial testing heats. The residuals shown in Figure 5.29 are estimates of the experimental error determined by subtracting the measured off-gas water vapor from the predicted off-gas water vapor. Figure 5.29 shows an error distribution that is roughly normal and is centered at approximately 0 with a standard deviation of about 0.52. The significance of the residuals normal distribution being centered at 0 indicates that the model error is random.

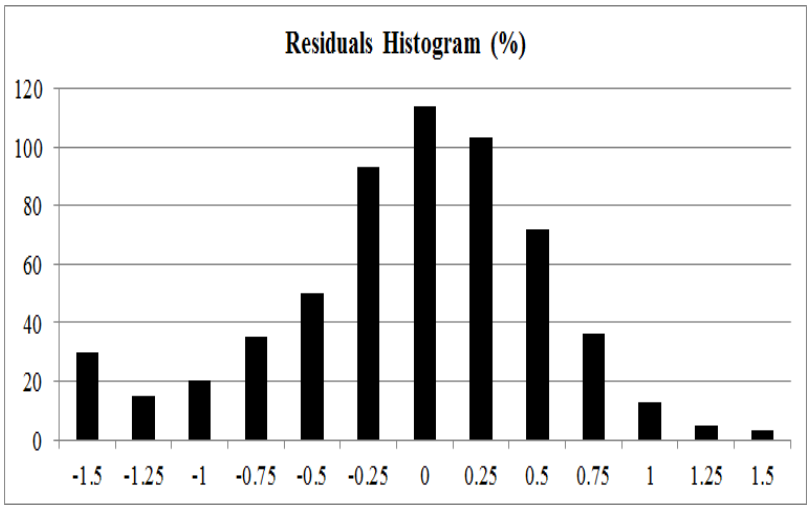

Figure 5.29: Residuals Histogram from the MPLS Model for all the Normal Testing Heats for EAF 1.

5.4.3 EAF 2 MPLS Results

The input variables used to train the MPLS model for EAF 2 are determined in the variables selection model in Chapter 4: EAF total fuel flow, EAF total main oxygen flow, EAF total shroud oxygen flow, DRI flow, EAF total injected carbon flow, off-gas CO, off-gas CO_2, electrode water Flow, and EAF analyzer purge H_2O. The number of training heats is 35 heats, and the number of testing heats is 15 heats. The testing heats include 13 normal heats and 2 trial heats where additional water is injected into the furnace. The number of principal components obtained is 7 components using cross validation. The calculated model R^2 is 80%. The first two components explain approximately 50% of the variability in the 50 heats data. The remaining 5 components in the MPLS model explained approximately 30 % of the variability in the dataset.

Figure 5.30 shows two figures. The top figure shows the MPLS prediction drawn in solid line and the 95% confidence interval in dashed line for a normal heat with DRI feeding and a refining period. The top figure shows that the MPLS model provides a narrow prediction range. The bottom figure shows a comparison between the EAF off-gas water vapor measured by the analyzer drawn in solid line and the MPLS off-gas water vapor prediction drawn in dashed line for a normal heat for EAF 2. The bottom

figure also shows that the predicted water vapor closely follows the measured water vapor throughout the heat.

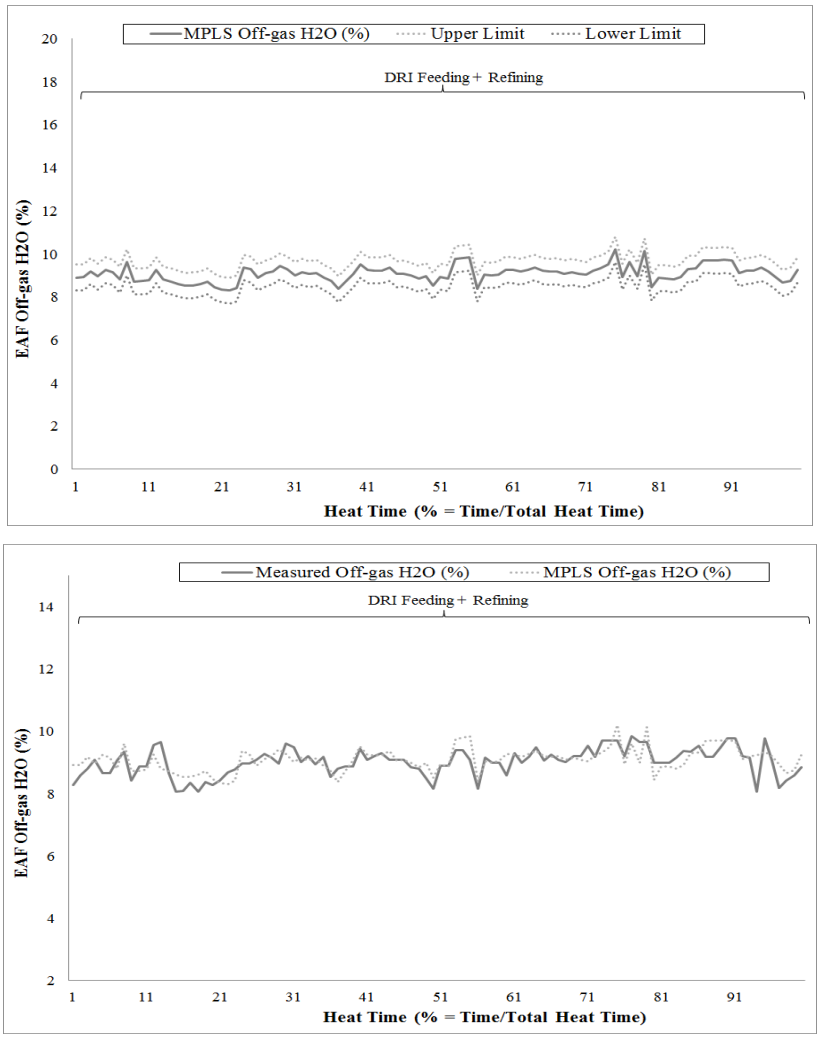

Figure 5.30: Top Figure - MPLS Predictions (solid line) and the 95% Confidence Interval (dashed line). Bottom Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and MPLS Prediction (dashed line). Both Figures are for the Same Normal Testing Heat for EAF 2.

Figure 5.31 shows two figures. The top figure shows the error which is calculated by subtracting the calculated EAF off-gas water vapor from the measured water vapor. The figure clearly shows that the error is below the 1% threshold for most of the heat, a finding which indicates that the predicted water vapor closely followed the measured water vapor. The bottom figure shows the errors are normally distributed and centered at 0.

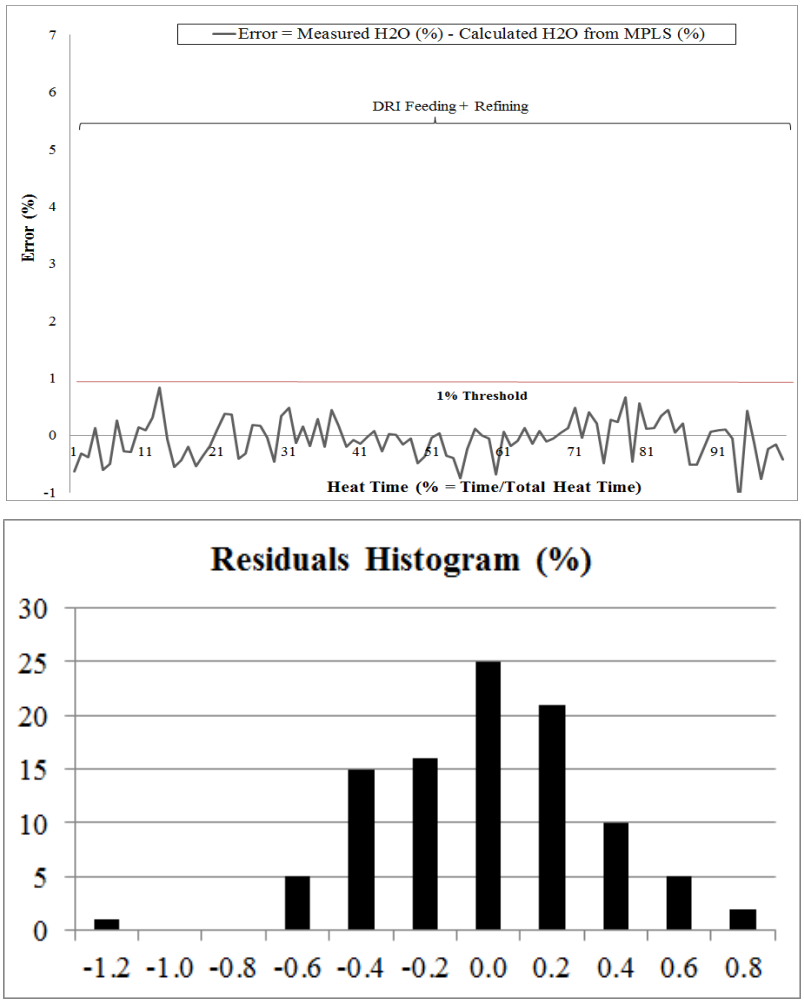

Figure 5.31: Top Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O for Normal Heat for EAF 2. Bottom Figure – Normal Distribution of the Error (%).

Figure 5.32 shows the MPLS prediction drawn in solid line and the 95% confidence interval in dashed line for trial 1 heat with DRI feeding and a refining period. The figure shows that the MPLS model provides a narrow prediction range.

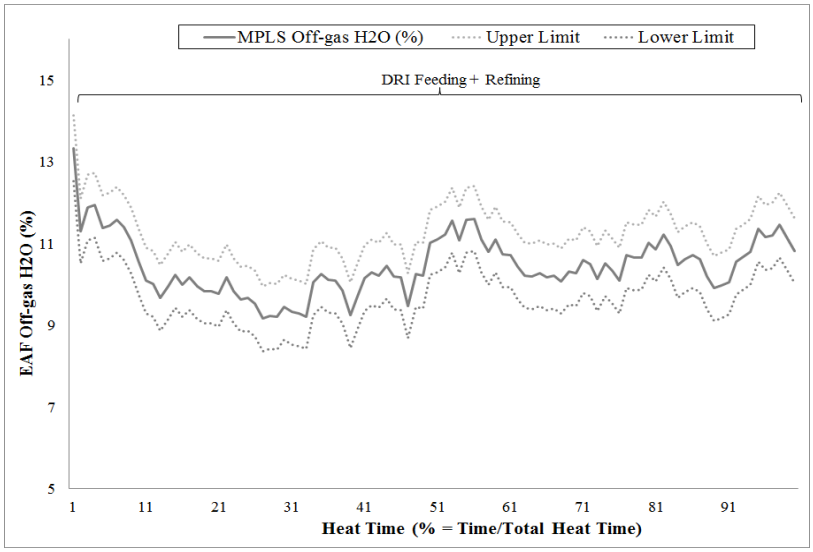

Figure 5.32: MPLS Predictions (solid line) and the 95 % Confidence Interval (dashed line) for Trial 1 Testing Heat for EAF 2.

Figure 5.33 shows two figures for trial 1 heat for EAF 2. The top figure shows a comparison between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by MPLS. Trial 1 heat is conducted by increasing the electrode spray water by a total of 30 liters per minute for approximately 6 minutes. The top figure clearly shows that the measured water vapor is higher than the calculated water vapor by more than 1% during the trial test. The bottom figure shows the error (%) between the measured and the calculated EAF off-gas water vapor. As in trial heats from EAF 1, the reason that the measured water vapor is higher than the calculated is because the MPLS model does not include water leaks as an input to the model.

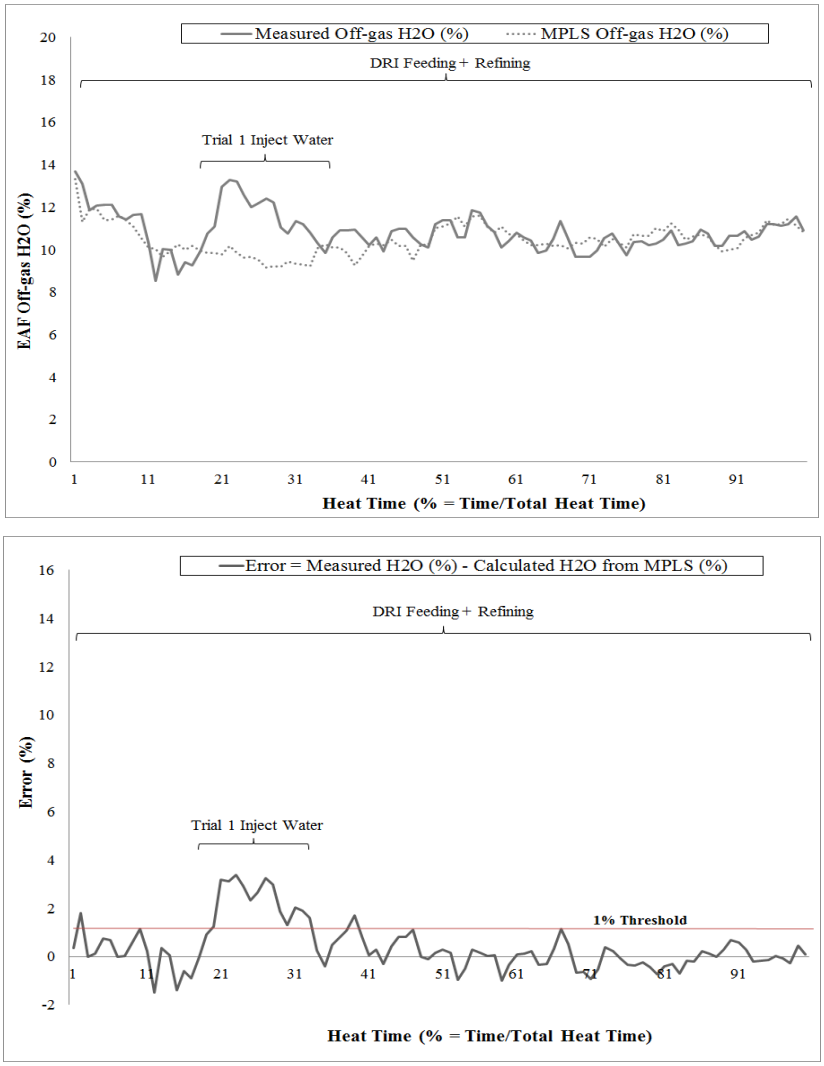

Figure 5.33: Top Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and MPLS Prediction (dashed line). Bottom Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 1 Testing Heat for EAF 2.

Figure 5.34 shows the MPLS prediction drawn in solid line and the 95% confidence interval in dashed line for trial 2 testing heat with DRI feeding and a refining period. The figure shows that the MPLS model provides a narrow prediction range.

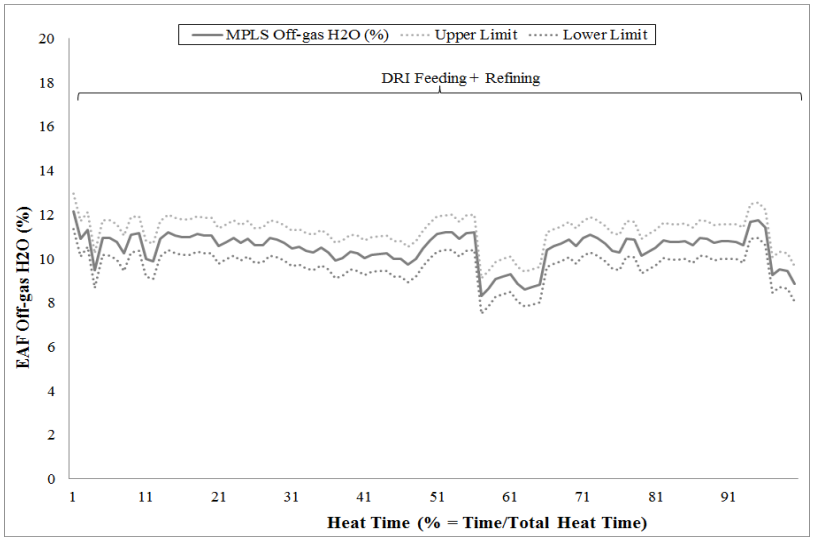

Figure 5.34: MPLS Predictions (solid line) and the 95% Confidence Interval (dashed line) for Trial 2 Testing Heat for EAF 2.

Figure 5.35 shows two figures for trial 2 heat for EAF 2. The top figure shows a comparison between the measured EAF off-gas water vapor (%) and the off-gas water vapor (%) calculated by MPLS. Trial 2 heat is conducted by increasing the electrode spray water by a total of 60 liters per minute approximately 10 minutes. The bottom figure clearly shows that the measured water vapor is higher than the calculated water vapor by more than 1% during the trial test. The bottom figure shows the error (%) between the measured and the calculated EAF off-gas water vapor. As in trial Heats from EAF 1, the reason that the measured water vapor is higher than the calculated is because the MPLS model does not include water leaks or abnormal sources of water as an input to the model.

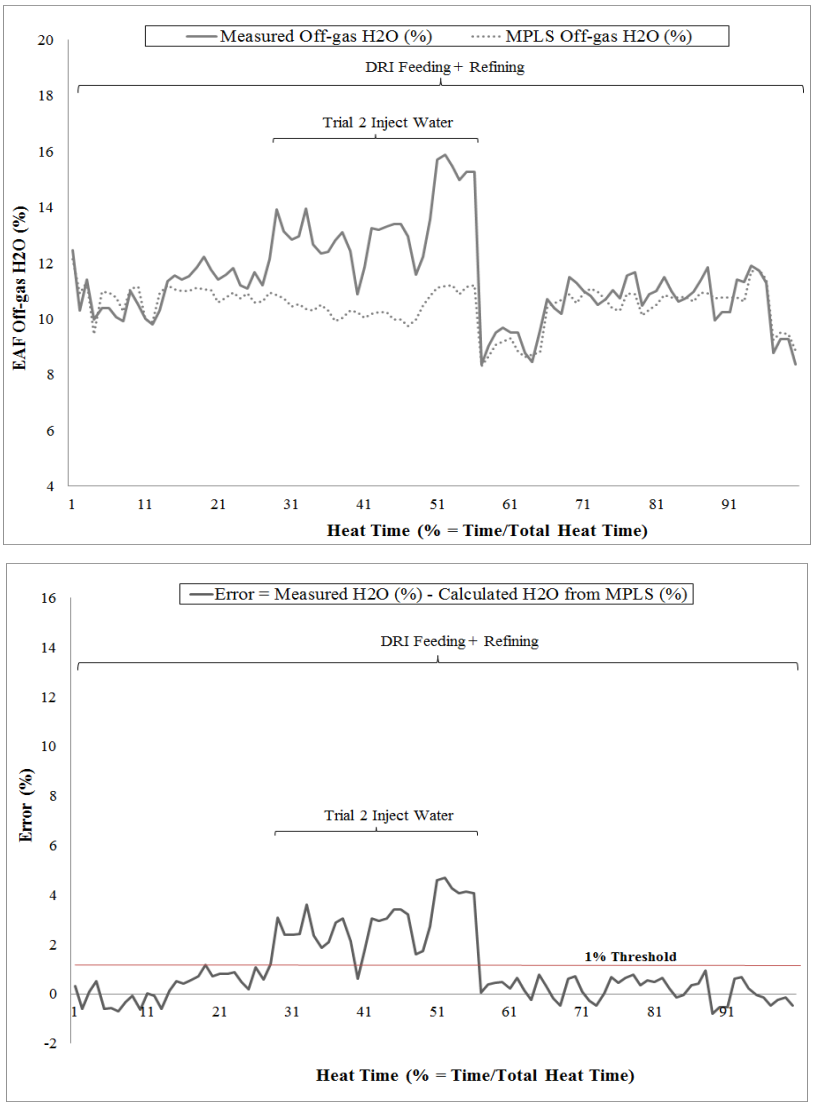

Figure 5.35: Top Figure - Comparison between Measured EAF Off-gas H_2O (solid line) and MPLS prediction (dashed line). Bottom Figure – Error (%) between Measured and the Calculated EAF Off-gas H_2O. Both Figures are for Trial 2 Testing Heat for EAF 2.

Figure 5.36 shows the residuals from the MPLS model for the normal testing heats for EAF 2. This error histogram excludes the trial testing heats. The residuals shown in Figure 5.36 are estimates of the experimental error determined by subtracting the measured off-gas water vapor from the predicted off-gas water vapor. Figure 5.36 shows an error distribution that is roughly normal and is centered at approximately 0. The significance of the residuals normal distribution being centered at 0 indicates that the model error is random.

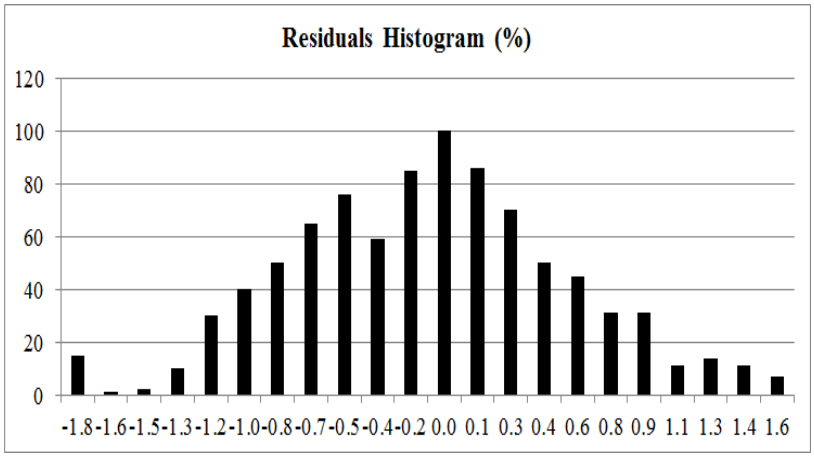

Figure 5.36: Residuals Histogram from the MPLS Model for all the Normal Testing Heats for EAF 2.

5.5 ANN and MPLS Performance Comparison

Two machine learning techniques, multiway projection to latent structures and the artificial neural network have been used in this chapter to predict the off-gas water vapor at EAF 1 and EAF 2. Table 5.1 summarizes the results for the EAF water vapor prediction performance of ANN and MPLS:

Table 5.1: Prediction Performance Comparison between ANN and MPLS

	Method	Training Heats	Testing Heats	R^2	RMSEP	Trial Heat 1 Detection	Trial Heat 2 Detection
EAF 1	MPLS	15	9 (7 Normal and 2 Trial Heats)	78	0.52	Partially Successful	Successful
	ANN	15	9 (7 Normal and 2 Trial Heats)	90	0.47	Successful	Successful
EAF 2	MPLS	35	15 (13 Heats and 2 Trial Heats)	80	0.45	Successful	Successful
	ANN	35	15 (13 Heats and 2 Trial Heats)	91	0.44	Successful	Successful

The roots mean squared prediction error (RMSEP) measures the square root of the expected squared distance between what the model predicts (y_t^p) for the water vapor and what the off-gas analyzer measures (y_t) during the heat time (t = 1 to n). RMSEP gives

the standard deviation of the model prediction error and hence is an indication of the quality of prediction. The unit for RMSEP is the same unit as the off-gas water vapor. Equation 5.9 shows how RMSEP is calculated:

$$RMSEP\ (\%) = \sqrt{\frac{\sum_{t=1}^{n}(y_t^p - y_t)^2}{n}} \tag{5.9}$$

EAF 1 training set R^2 for the ANN model is 90% and for the MPLS model 78%. The ANN RMSEP for EAF 1 testing set is 0.47%, whereas MPLS RMSEP is 0.52%; hence ANN model outperformed MPLS model for EAF 1. Moreover, MPLS did not completely detect the additional injected water into the furnace during trial 1. EAF 2 training set R^2 for the ANN model is 91% and for the MPLS model 80%. The ANN RMSEP for EAF 21 testing set is 0.44%, whereas MPLS RMSEP is 0.45%; hence ANN model also outperformed MPLS model for EAF 2 by a small margin. Both models detected successfully the trials heats, and the performances are similar for EAF 2. ANN model performed significantly better in EAF 1 where the behavior of the off-gas water vapor was non-linear, and it performed slightly better in EAF 2 where the behavior of the off-gas water vapor was less non-linear.

ANN is useful in the case of non-linear systems such as the EAF off-gas water vapor because it has the capability of capturing nonlinear and complex underlying characteristics of physical non-linear process. The method works well for large data sets, and it is a non-parametric method; thus this eliminates the error in parameter estimation. Disadvantages of the ANN method include the method being a black box model where it is difficult to extract knowledge of the weights computed and where the method is incapable of extrapolating the results. Another disadvantage of ANN is overfitting which occurs when the neural network memorizes the training heats but is incapable of generalizing to new heats. The approach used in this work to avoid overfitting is to collect more data and increase the size of the training set. Although the ANN model outperformed the MPLS model, the MPLS has attractive features that the ANN lacks. MPLS model is easier to interpret and extract knowledge of the model loadings and

scores. Moreover, due to the nature of the MPLS model, it is capable of extrapolating the results. MPLS effectively handles collinearity between variables and missing data in the training and testing heats; hence the method can predict even when some of the inputs are missing, making it an appealing model for real-time predictions. However, ANN outperformed the MPLS in EAF off-gas water vapor predictions because the behavior of the off-gas water vapor is nonlinear in both furnaces.

5.6 Fault Detection Methodology

As discussed earlier, there is always a normal level of water vapor in the freeboard off-gas inside the EAF. However, it is important to note that the absolute level of normal water vapor in the off-gas is varying throughout the heat and from heat-to-heat depending on the quality of the scrap, on burner firing rates, on post combustion at any point during the heat, on the level of the electrode sprays, and on the level of fume system suction. To be effective, a water detection system must be able to quickly and correctly distinguish between abnormal water vapor levels due to a water leak into the EAF and a normal level of water vapor due to operating practice. Two fault detection algorithms are proposed in this work. Fault Detection Method 1 is based on the fingerprinting method and Fault Detection Method 2 is based on using the difference between the measured and the predicted off-gas water vapor.

Fault Detection Method 1 is implemented when statistical fingerprinting is used. This method compares the measured off-gas water vapor against the baseline water vapor. The baseline values are calculated based on the algorithm described earlier in section 5.1. Hence, this method distinguishes between “normal” and “abnormal” water vapor conditions in the EAF freeboard and then provides the operators with alerts that clearly indicate when water vapor levels exceed normal levels. Figure 5.37 schematically illustrates Fault Detection Method 1 for triggering “Operator Alerts”. While the method is not a failsafe method, it does provide operators with valuable real-time alerts indicating the statistical probability of excessive high amounts of off-gas water vapor in the EAF. Specific threshold limits are calculated according to the

fingerprinting method described in section 5.1. When the off-gas water vapor is equal to or less than the normal fingerprint threshold is green. This indicates the statistical probability of excessive amounts of water in the EAF is low. When the indicators exceed the upper threshold limit for 5 seconds, a "Red Alert" is issued indicative that the off-gas chemistry is significantly out of the statistically normal range and there is a high probability of excess water in the EAF. Red Alerts require immediate protective action by EAF operating staff.

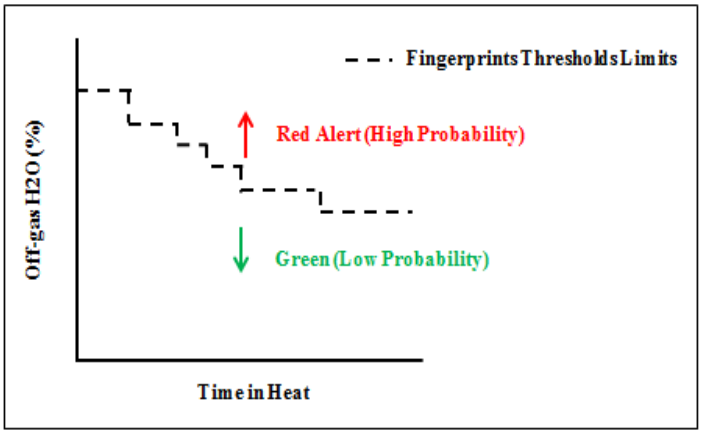

Figure 5.37: Operator Alerts Based on Fault Detection Method 1.

Fault Detection Method 2 is used when MPLS or ANN is used to predict off-gas water vapor. Figure 5.38 schematically illustrates Fault Detection Method 2. This method proposes to use the difference between the measured and the calculated off-gas water vapor in the furnace. The reason for this is that the calculated water vapor model includes all potential sources of water except a water leak and the off-gas analyzer measures the EAF off-gas in real time; therefore, this work proposes to use the difference as an indicator for a water leak. The analysis from this work indicates using a 1% as an alarm threshold. The analysis is based on minimizing the false alarm rate and capturing the artificial leak. However, this is a tuning parameter, where in other electric arc furnaces, this value may change. When the calculated difference exceeds 1% for 5 seconds, a "Red Alert" is issued to indicate that there is excessive water in the furnace not included in the empirical model. The 5 seconds timer is also a tuning parameter to minimize false alarm rate and capture the artificial leak.

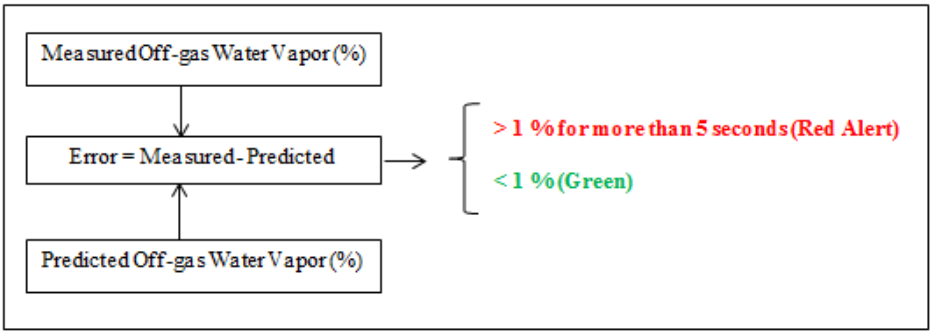

Figure 5.38: Operator Alerts Based on Fault Detection Method 2.

Chapter 6

Conclusions and Recommendations

6.1 Summary and Conclusions

In this thesis a framework for water detection in an industrial electric arc furnace is developed. The objective of the framework is to prevent water leak furnace explosions. The framework is based on a simplified mechanistic mass balance model and empirical models because the electric arc furnace is too complex to be modeled solely by first principles. The framework developed in this work is implemented on two industrial AC electric arc furnaces. The first EAF is a scrap charging furnace, and the second one is mostly a DRI furnace. Figure 6.1 shows the framework developed in this work for the water leak detection in an industrial electric arc furnace:

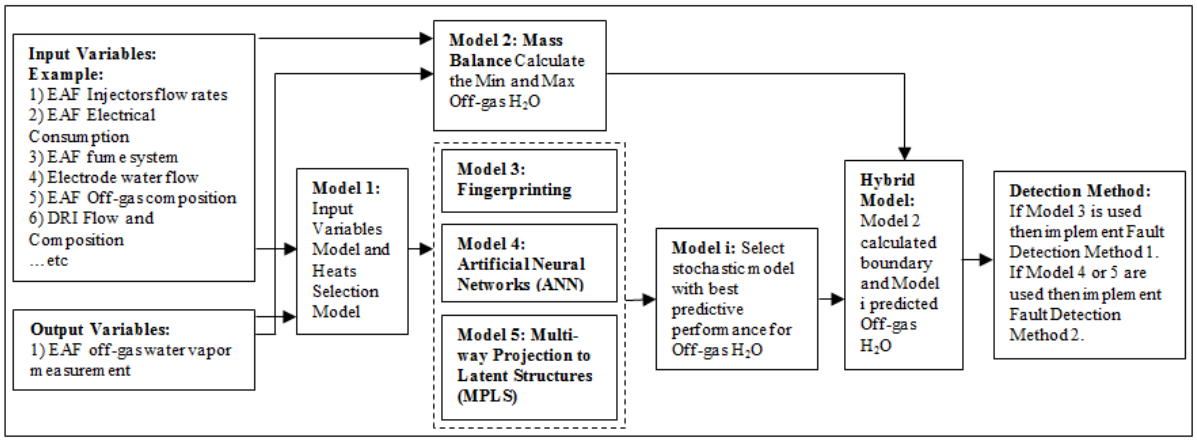

Figure 6.1: Water Detection Framework Developed in this Work.

The mechanistic model (Model 2) is completed by performing a mass balance on the furnace and it provides a boundary with upper and lower limits in real-time of expected EAF off-gas water vapor leaving the furnace. The mechanistic model has shown in both industrial EAFs that it provides a valuable on-line monitoring tool to the operator on what boundary to expect for the off-gas water vapor.

There are many input variables and historical heats in an EAF operation; hence before building the empirical component of the hybrid algorithm, a heat selection model and input variables selection model (Model 1) are constructed based on latent variable methods. The outcome of the heats selection model is heats with normal operation, where some of those heats are used to build the predictive models and the remaining are used to test the models. The outcome of the variable selection model is variables that have the highest correlation with the off-gas water vapor.

The empirical component of the algorithm is investigated by exploring three different methods: statistical fingerprinting, artificial neural network, and multiway projection to latent structures. Statistical fingerprinting method is simplistic in nature; however, there are few disadvantages with this method: it cannot detect smaller leaks (30 liters per minute), and it only works if the process is well controlled where in most electric arc furnaces this is not common. ANN and MPLS performed better, and they are capable of detecting all of the trials; trial 1,however, had a partial success in detection when MPLS was used in EAF 1. ANN outperformed the MPLS in EAF off-gas water vapor predictions because the behavior of the off-gas water vapor is nonlinear in both furnaces.

The last section of this thesis proposes two approaches to detect for a potential water leak in the furnace. Fault Detection Method 1 is implemented when statistical fingerprinting is used. This method compares the measured off-gas water vapor against the baseline water vapor. Fault Detection Method 2 is used when MPLS or ANN is used to predict off-gas water vapor. This method proposes to use the difference between the

measured and the calculated off-gas water vapor in the furnace. The reason behind this method is that the calculated water vapor model includes all potential sources of water except a water leak and the off-gas analyzer measures the EAF off-gas in real time, therefore, this work proposes to use the difference as an indicator for a water leak.

6.2 Future Work

Opportunities for future work include increasing the quantity of available measurements that would improve the performance of the off-gas water vapor predictive models. For example, measuring scrap and DRI composition and off-gas temperature can improve the accuracy of the empirical predictive models. Another interesting problem would be to develop a water detection method based on off-gas hydrogen. It was shown in Chapter 3 (Figure 3.4 for EAF 1 and Figure 3.5 for EAF 2), that off-gas hydrogen is present in the furnace freeboard. It was also shown in Chapter 4 that off-gas hydrogen can be produced from different sources such as natural gas combustion with oxygen and the water shift reaction. The water shift reaction produces less off-gas hydrogen as the temperature rises in the furnace but the reaction can produce hydrogen at low temperatures in the furnace. The objective of looking at this problem is to see if off-gas hydrogen contains water leak information that can improve the performance of the water detection framework to detect smaller leaks (i.e. below 30 liters/min leaks) in the furnace.

Another interesting problem is multi-model fusion. Multi-model fusion is the process of integrating information from various empirical models and combining them into one prediction. Botwey et al. (2014) implemented multi-model fusion to improve an early warning system for hypo-/hyperglycemic events and thus to improve patent's safety. Data fusion techniques used by the authors were based on Dempster-Shafer Evidential theory, genetic algorithm, and genetic programming, which were used to merge the complimentary performance of the prediction models. The authors had shown that the fusion schemes significantly improved the prediction performance with lower root mean square errors, lower time lags, and higher correlation. The methodology followed in this work is to predict off-gas water vapor with three empirical models and

then to compare the performance of the predictions between the different empirical models and to select the optimum model for the EAF operation. Hence an interesting problem is to investigate if multi-model fusion can be used to improve the prediction performance of the off-gas water vapor.

References

"Accident Report Detail – Accident: 202004982 and Report ID: 0317000." OSHA.com. *Occupational Safety & Health Administration*, 27 May 2007. Web. 1 Jan. 2015.

Alshawarghi, H., A.Vazquez and H. Suzuki. "Optimization Results of Applying Dynamic Control for the EAF Chemical Energy using Tenova's EFSOP® Technology at Kanto Steel – Japan." *Proceedings of the Asia Steel Conference 2015*, Yokohama, Japan. October 5-8, 2015.

Bekker, J., I. Craig, and P. Pistorius. "Modeling and Simulation of an Electric Arc Furnace Process." *The Iron and Steel Institute of Japan*, **39**, pp.23–32, 1999.

Benaicha, A., G. Mourot, K. Benothman, and J. Ragot. "Determination of principal component analysis models for sensor fault detection and isolation." *International Journal of Control, Automation, and Systems*, **11**, pp.296-305, 2013.

Burnham, A., J. F. MacGregor, and R. Viveros. "Latent Variable Multivariate Regression Modeling," *Chemometrics and Intelligent Laboratory Systems*, **48**, pp.167-180, 1999.

Burnham, A., J. F. MacGregor, and R. Viveros, "Frameworks for Latent Variable Multivariate Regression." *Journal of Chemometrics*, **10**, pp.31-45, 1996.

Boin, N. and A. Vazquez. "Utilizing Tenova Goodfellow's EFSOP® Technology to Improve EAF Performance and Enhance Safety." *Proceedings of the ABM 2011, Steelmaking Seminar – International*, Bahia, Brazil, May 15-18, 2011.

Botwey, R.H., E. Daskalaki, P. Diem, S. Mougiakakou. “Multi-Model Data Fusion to Improve an Early Warning System for Hypo-/Hyperglycemic Events.” *IEEE Engineering in Medicine and Biology Society Conference Proceedings*, pp. 4843-4846, 2014.

Cameron, A., N. Saxena, and K. Broome. “Optimizing EAF Operations by Dynamic Process Simulation.” *56th Electric Furnace Conference Proceedings*, pp.689–696, New Orleans, LA, 1998.

Chapra, S. C. and R. P. Canale. *Numerical Methods for Engineers: With Personal Computer Application*s, New York: McGraw Hill, 1985.

Chetouani, Y. “A Neural Network Approach for the Real-time Detection of Faults.” *Empirical Environmental Research and Risk Assessment*, **22**, pp.339-349, 2008.

Chryssolouris, G., M. Lee, and A. Ramsey. “Confidence Interval Prediction for Neural Network Models.” *IEEE Transactions on Neural Networks*, **7**, pp.229-232, 1996.

Daskalaki, E., K. Norgaard, T. Zuger, A. Prountzou, P. Diem, S.Mougiakakou. “An early-warning system for hypo-/hyperglycaemia events based on fusion of adaptive prediction models.” *Journal of Diabetes Science Technology*, **7**, 3, pp. 689-698, 2013.

Dunn, Kevin. *Process Improvement Using Data*. McMaster University, Hamilton, ON, 2014.

Eastment, H. T., and W. J., Krzanowski, “Cross-Validatory Choice of the Number of Components from a Principal Component Analysis,” Technometrics, 24, 73 (1982).

Efron B. and R. J. Tibshirani. *An Introduction to the Bootstrap*, Chapman and Hall, London, UK, 1993.

Fang, K., D. Mu, S. Chen, B. Wu, and F. Wu. "A prediction model based on artificial neural network for surface temperature simulation of nickel-metal hydride battery during charging." *Journal of Power Sources*, **208**, pp.378-383. 2012.

Freeman, P., R. Pandita, N. Srivastava, and G. J. Balas. "Model-Based and Data-Driven Fault Detection Performance for a Small UAV." *IEEE/ASME Transactions on Mechatronics*, **18**, (4), pp.1300-1309, 2013.

Fruehan, R. J., *The Making, Shaping and Treating of Steel*. AISE Steel Foundation, Pittsburgh, PA, 1998.

Garcia, S., T. Kourti, and J. F. MacGregor. "Model Predictive Monitoring for Batch Processes." *Industrial & Engineering Chemistry Research*, **43**, pp.5929-5941. 2003.

Garcia, S., T. Kourti, J. F. MacGregor, A. Matheos, and G. Murphy. "Troubleshooting of an Industrial Batch Process Using Multivariate Methods." *Industrial & Engineering Chemistry Research*, **42**, pp.3592-3601, 2003.

Golshan, M. and J. F. MacGregor. "Identification for the Control of Variable Trajectories in Batch Processes." *Industrial & Engineering Chemistry Research*, **52**, pp.2352–2367, 2013.

Grabowski D. and J. Walczak. "Deterministic model of electric arc furnace–a closed Form solution." *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, **32**, pp.1428-1436, 2013.

Gueguim Kana, E.B., J.K. Oloke, A. Lateef, and M.O. Adesiyan. “Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm.” *Renewable Energy*, **46**, pp.276-281, 2012.

Grieshaber, K. and F. Martinez. “ZoloSCAN Laser Diagnostic System for EAF Leak Detection and Process Optimization.” *Proceedings of Metec Conference*, Dusseldorf, Germany, June 15–19, 2015.

G. W. Chang, C. Chen, and Y. Liu. "A neural-network-based method of modeling electric arc furnace load for power engineering study." *IEEE Transactions on Power Systems*, **25**, pp. 138-146, 2010.

Heskes T., “Practical confidence and prediction intervals.” *Advances in Neural Information Processing Systems*, **9**, pp. 176–182, 1997.

Hui Z. and X. Wang.“Prediction model of arc furnace based on improved BP neural Network.” *Journal of Information Technology*, pp.127-130, 2009.

Irons, G. A., “Developments in Electric Furnace Steelmaking”, *AISTech Conference*, Charlotte, NC, 2005.

Isermann, R., “Fault Diagnosis of Diesel Engines.”*Mechanical Engineering*, 135, pp. 6-14, 2013.

Jeong H., S. Cho, D. Kim, H. Pyun, D. Ha, C. Han, M. Kang, M. Jeong, and S. Lee. “A heuristic method of variable selection based on principal component analysis and factor analysis for monitoring in a 300 kW MCFC power plant.” *International Journal of Hydrogen Energy*, **37**, pp.394-400, 2012.

Jiang Q., X. Yan, and W. Zhao, H. “Fault Detection and Diagnosis in Chemical Processes Using Sensitive Principal Component Analysis.”, *Industrial & Engineering Chemistry Research*, **52**, pp.1635-1644, 2013.

Jie, L., “A Bayesian inference based two-stage support vector regression framework for soft sensor development in batch bioprocesses.” *Computers & Chemical Engineering*,**41**, pp.134-144, 2012.

Jones, J.A.T., E. Pretorius, and H. Oltmann. *EAF Fundamentals, Notes by Process Technology Group at LWB Refractories and Nupro Corporation*, 2005.

Jones, J.A.T., “Electric Arc Furnace Design and Operations.” *AISE Electric Furnace Steelmaking Seminar*, Birmingham, AL, 2014.

Kaistha, N. and C. Moore. "Extraction of Event Times in Batch Profiles for Time Synchronization and Quality Predictions." *Industrial & Engineering Chemistry Research*, **40**, pp. 252-260, 2001.

Kassidas, A., J. F. MacGregor, and P. Taylor. "Synchronization of Batch Trajectories Using Dynamic Time Warping." *American Institute of Chemical Engineers Journal*, **44**, pp. 864-875, 1998.

Khan, M.I. and E.J. Evenson. “Results of Goodfellow EFSOP® at Birmingham Steel Corp., Seattle, WA, USA.” *ISS Tech, Iron& Steel Society*, Indianapolis, IN, 2003.

King, P.E. and M.D.Nyman. “Modeling and control of an electric arc furnace using a feedforward artificial neural network.” *Journal of Applied Physics*, **80**, pp. 1872-1877, 1996.

Kourti, T. and J. F. MacGregor. "Multivariate SPC Methods for Process and Product Monitoring." *Journal of Quality Technology*, **28**, pp. 409-428, 1996.

Kourti, T., P. Nomikos, and J. F. MacGregor. "Analysis, Monitoring and Fault Diagnosis of Batch Processes Using Multiblock and Multiway PLS." *Journal of Process Control,* **5**, pp. 277-284, 1995.

Logar, V., D. Dovzan, and I. Skrjanc."Mathematical Modeling and Experimental Validation of an Electric Arc Furnace." *The Iron and Steel Institute of Japan*, **51,** 3, pp. 382-391, 2011.

Logar, V., D. Dovzan, and I. Skrjanc. "Modeling and Validation of an Electric Arc Furnace: Part1, Heat and Mass Transfer." *The Iron and Steel Institute of Japan*, **52**, 3, pp. 402-412, 2012.

Logar, V., D. Dovzan, and I. Skrjanc. "Modeling and Validation of an Electric Arc Furnace: Thermochemistry." *The Iron and Steel Institute of Japan*, **52**, No. 3, pp. 413-423, 2012.

Lin, Ming-I, W. Groves, A. Freivalds, E. Lee, and M. Harper. "Comparison of artificial neural network (ANN) and partial least squares (PLS) regression models for predicting respiratory ventilation: an exploratory study." *European Journal of Applied Physiology*, **112**, pp.1603-1611, 2012.

Macgregor, J.F. and P. Nomikos. "Monitoring Batch Processes Using Multiway Principal Component Analysis." *American Institute of Chemical Engineers Journal*, **40**, pp. 1361-1375, 1994.

Macgregor, J. F., C. M. Jaeckle, C. Kiparissides, and M. Koutoudi. "Process Monitoring and Diagnosis by Multiblock PLS Methods." *American Institute of Chemical Engineers Journal*, **40**, (5), pp. 826-838, 1994.

Macrosty, R. and C. Swartz. "Dynamic Modeling of an Industrial Electric Arc Furnace." *Industrial & Engineering Chemistry Research*, **44**, pp. 8067-8083, 2005.

Maiolo, J., K. Nikkanen, S. Mistry, and M.I. Khan. “Operation Benefits at ArcelorMittal Contrecoeur Ouest using EFSOP® Technology.” *AISTech Conference*, Indianapolis, IN, 2011.

Mathworks, *Statistics Toolbox User's Guide for Use with MATLAB*, 2009.

Mathworks, *Neural Network Toolbox User's Guide for Use with MATLAB*, 2014.

Matson, S. and W. F. Ramirez. “Optimal Operation of an Electric Arc Furnace.” *In 57th Electric Furnace Conference Proceedings*, pp. 719–728, Pittsburgh, Pennsylvania, 1999.

Meng, M. and G.A. Irons. “Comparison of Electric Arc Models with Industrial Data,” *58th Electric Furnace Conference and 17th Technology Conference, ISS Publishers*, Warrendale, PA, pp 183-194, 2000.

Missio, M., N. Boin, and M. Khan. “Optimization Results at Ferriere Nord Using EFSOP® Technology.” *AISTech Conference*, Pittsburg, PA, 2010.

Modigell, M., A. Trabert, and P. Monheim. “A Modelling Technique for Metallurgical Processes and Its Applications.” *AISE Steel Technology*, **28**, pp. 45–47, 2001.

Nielsen, Michael, *Neural Networks and Deep Learning*. Determination Press, 2015.

Nikkanen, K., J. Maiolo, and E. Chiarullo. “Dynamic Control And Optimization of EAF Chemical Energy Using Tenova’s EFSOP® Technology.” *43rd ABM Steelmaking Seminar International*, Belo Horizonte, Brazil, 2012.

Nomikos, P., "Detection and Diagnosis of Abnormal Batch Operations Based on MultiWay Principal Component Analysis." *ISA Transactions*, **35**, pp. 259-266, 1996.

Nomikos, P. and J. F. MacGregor. "Multi-Way Partial Least Squares in Monitoring Batch Processes." *Chemometrics and Intelligent Laboratory Systems*, **30**, pp. 97-108, 1995a.

Nomikos, P. and J. F. MacGregor. "Multivariate SPC Charts for Monitoring Batch Processes." *Technometrics*, **37**, (1), pp. 41-58, 1995b.

Putnam R., Q. Mohaidat, A. Daabous, and S. Rehse. "A comparison of multivariate analysis techniques and variable selection strategies in a laser-induced breakdown spectroscopy bacterial classification." *Spectrochimica Acta*, **87**, pp.161-167, 2013.

Quiroga, P. "Leaks Prevention: Design Directives to Achieve Reliable Water Cooled Equipment." *Maintenance Guild Presentation*, Beaver Falls, PA, 2013.

"Resource Efficiency." Http://www.worldsteel.org. *World Steel Association*, 1 Mar. 2014. Web. 1 Apr. 2016.

Rodrigues, Cecilia Pereira., John Frederick MacGregor, Theodora Kourti, and McMaster University. Dept. of Chemical Engineering. Industrial Batch Data Analysis Using Latent Variable Methods. 2006.

Sheibat-Othman N., N. Laouti, J. Valour, and S. Othman. "Support Vector Machines Combined to Observers for Fault Diagnosis in Chemical Reactors." *The Canadian Journal of Chemical Engineering*, **92**, pp.685-695, 2014.

Tang, X., M. Kirschen, M. Abel, and H. Pfeifer. "Modeling of EAF Off-Gas Post Combustion in Dedusting Systems using CFD Methods." *Steel Research International*, **74**, pp. 201-210, 2003.

Tiwari, M. K., and C. Chatterjee. "Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs)." *Journal of Hydrology*, **382**, pp.20-33, 2010.

Tominaga, Y., "Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN." *Chemometrics and Intelligent Laboratory Systems*, **49**, pp.105-115, 1999.

Torabian Esfahani, M. and B. Vahidi. "A New Empirical Model of Electric Arc Furnace Based on Hidden Markov Model: A Study of Its Effects on Power System." *IEEE Transactions on Power Delivery*, **27**, 4, pp. 1893-1901, 2012.

Wieczorek T., and K. Macza. "Modeling of the AC-EAF process using computational intelligence methods." *Electrotech*, **11**, pp. 184-188, 2008.

Wold, S., "Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models," *Technometrics*, **20**, 397, 1978.

Wold, S., K. Esbensen, and P. Geladi, "Principal Component Analysis," *Chemometrics and Intelligent Laboratory Systems*, **2**, 31, 1987a.

Wold, S., P. Geladi, K. Esbensen, and J. Ohman, "Multi-Way Principal Components and PLS Analysis," *Journal of Chemometrics*, **1**, 41, 1987b.

Yu, J. A., "A Bayesian inference based two-stage support vector regression framework for soft sensor development in batch bioprocesses." *Computers & Chemical Engineering*, **41**, pp. 134–144, 2012.

Zuliani, D.J., H. Alshawarghi, V. Scipolo, M. Khan, and W. Biliski, W., "Real-Time Water Detection Technology at ArcelorMittal Coatesville." *Iron & Steel Technology*, pp. 3340-3349, 2014.