2016

A Feasibility Study for the Automated Monitoring and Control of Mine Water Discharges

Christopher R. Vass

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Vass, Christopher R., "A Feasibility Study for the Automated Monitoring and Control of Mine Water Discharges" (2016). *Graduate Theses, Dissertations, and Problem Reports*. 6869.
<https://researchrepository.wvu.edu/etd/6869>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

A Feasibility Study for the Automated Monitoring and Control of Mine Water Discharges

Christopher R. Vass

Thesis submitted
to the Statler College of Engineering and Mineral Resources
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science in
Mining Engineering

Aaron Noble, Ph.D., Chair
Keith Heasley, Ph.D.
Mark Sindelar, Ph.D.

Department of Mining Engineering

Morgantown, West Virginia
2016

Keywords: Adaptive Neuro Fuzzy Inference System, Acid Mine Drainage, Fuzzy Logic

Copyright 2016 Christopher R. Vass

Abstract

A Feasibility Study for the Automated Monitoring and Control of Mine Water Discharges

Christopher R. Vass

The chemical treatment of mine-influenced waters is a longstanding environmental challenge for many coal operators, particularly in Central Appalachia. Mining conditions in this region present several unique obstacles to meeting NPDES effluent limits. Outlets that discharge effluent are often located in remote areas with challenging terrain where conditions do not facilitate the implementation of large-scale commercial treatment systems. Furthermore, maintenance of these systems is often laborious, expensive, and time consuming. Many large mining complexes discharge water from numerous outlets, while using environmental technicians to assess the water quality and treatment process multiple times per day. Unfortunately, this treatment method when combined with the lower limits associated with increased regulatory scrutiny can lead to the discharge of non-compliant water off of the mine permit. As an alternative solution, this thesis describes the ongoing research and development of automated protocols for the treatment and monitoring of mine water discharges. In particular, the current work highlights machine learning algorithms as a potential solution for pH control.

In this research, a bench-scale treatment system was constructed. This system simulates a series of ponds such as those found in use by Central Appalachian coal companies to treat acid mine drainage. The bench-scale system was first characterized to determine the volumetric flow rates and resident time distributions at varying flow rates and reactor configurations. Next, data collection was conducted using the bench scale system to generate training data by introducing multilevel random perturbations to the alkaline and acidic water flow rates. A fuzzy controller was then implemented in this system to administer alkaline material with the goal of automating the chemical treatment process. Finally, the performance of machine learning algorithms in predicting future water quality was evaluated to identify the critical input variables required to build these algorithms. Results indicate the machine learning controllers are viable alternatives to the manual control used by many Appalachian coal producers.

Acknowledgments

I would first like to express my sincere gratitude to my committee chair, Dr. Aaron Noble for his expert guidance, understanding, support, and encouragement. While Dr. Noble is technically my advisor, his charisma, wisdom, and astute analytical reasoning have placed his role much closer to being a mentor and friend. Given his spirit of adventure in regard to research and genuine excitement in the betterment of others, Dr. Noble has established himself as a role model to both his students and peers.

Next, I would like to thank my committee members Dr. Keith Heasley and Dr. Mark Sindelar for their time, advice, and insightful comments. They have provided a wealth of knowledge, which I will never forget.

Both Fernando Campos and Lucas Santos assisted with the generation of much of the data presented in this thesis. Their hard work and painstaking effort helped to make this project successful. Fernando and Lucas are both owed a debt of gratitude for their help and support.

This work would not have been accomplished if not for my lovely wife Rachel. Her encouragement, quiet patience, and unwavering love were undeniably the bedrock upon which the last eight years of my life have been built. Rachel's tolerance for my occasional crude temperament is a testament in itself of her unyielding devotion and love.

Finally, I would like to thank my father Casey. Dad has taught me the value of a solid work ethic and provided an example of a high ethical standard which has been the foundation of the many accomplishments I have achieved. Without his guidance and support I would undoubtedly be far less of a man than I am today.

CONTENTS

1	Introduction	1
1.1	Background	1
1.2	Motivation	3
1.3	Objectives	4
1.4	Organization	5
2	Literature Review	6
2.1	Introduction	6
2.2	Review of Mine Water Discharges	7
2.2.1	NPDES & the Clean Water Act	7
2.2.2	Consent Decrees	9
2.2.3	Description and Causes of AMD	10
2.2.4	Effects of AMD on the Environment	12
2.3	Treatment of AMD	13
2.3.1	Passive AMD Treatment	13
	Aerobic Wetlands	14
	Anaerobic (Compost) Wetlands	14
	Anoxic Limestone Drains	15
2.3.2	Active AMD Treatment	16
	Chemical Treatment	16
2.4	pH Control Systems	17
2.4.1	General Characteristics of Control Systems	18
2.4.2	Artificial Neural Networks	20

2.4.3	Fuzzy Logic	22
2.4.4	Fuzzy Controllers / Mamdani Controllers	24
2.4.5	Adaptive Neuro-Fuzzy Inference Systems	25
2.4.6	Comparison of ANFIS Controllers	26
2.4.7	Sugeno Type ANFIS Controller	28
	Theory	28
	Applications	30
	Development and Training	31
2.4.8	Hybrid ANFIS Controllers	32
2.4.9	Model Predictive Controllers	33
2.5	Summary	33
3	Design and Characterization of Bench-Scale System	35
3.1	System Design	35
3.1.1	Reactors	35
3.1.2	Sensor Package	40
3.1.3	Data Acquisition	42
3.1.4	Grounding with Floating Signal Source	42
3.1.5	Software	42
3.2	Characterization of Bench Scale System	44
3.2.1	Calibration of Sensors	44
3.2.2	Calibration of Pumps	45
3.2.3	RTD Testing - Theory	46
3.2.4	RTD Test - Procedure	47
3.2.5	RTD Test - Results and Discussion	48
4	Experimental Methods	52
4.1	Goals of Experimental Program	52
4.2	ANFIS Control Tests	53
4.2.1	Test Procedures	53

4.2.2	ANFIS Controller Membership Functions and Rules	56
4.2.3	Test Matrix	59
4.2.4	Analytical Methods	60
4.3	Prediction Tests	60
4.3.1	General Testing Goals	60
4.3.2	Data Acquisition	61
4.3.3	Data Processing	63
4.3.4	Data Analysis	65
5	Results and Discussion	67
5.1	Results of Controller Tests	67
5.1.1	Test 1 - Control Under Steady-State Conditions	67
5.1.2	Test 2 - Control With Varying Flow Rate	68
5.1.3	Test 3 - Control with Changing Set-Points	70
5.1.4	Test 4 - Control with a Surge in Flow	72
5.1.5	Test 5 - Control with a Change in pH	74
5.1.6	Test 6 - Control with Change in Flow Dynamics	78
5.1.7	Test 7 - Control with Multiple Perturbations	78
5.1.8	Summary of Control Tests	80
5.2	Predictive Accuracy of ANFIS	80
5.2.1	Data Tables	83
5.2.2	Discussion	88
6	Conclusions and Recommendations	91
6.1	Conclusions	91
6.2	Recommendations / Future Work	92
	Bibliography	94

LIST OF FIGURES

2.1	Streams with orange staining from the iron oxide precipitate “yellow boy.”	12
2.2	Ideal titration curves of monoprotic and polyprotic acids after Mc-Murry and Fay (1998).	18
2.3	Example of response characteristics of a control system after Nise (2007).	19
2.4	Example of neural network with 4 input, 3 hidden, and 2 output nodes after Kim (2014).	21
2.5	Example of membership functions used by Matlab after Jang et al. (1998).	23
2.6	Example of membership function for the volume of a glass of water.	24
2.7	Layers of ANFIS architecture. Layer 1: input fuzzification; Layers 2 and 3: ANN hidden layers; Layer 4 defuzzification; and Layer 5: crisp output	27
3.1	Bench-scale system built to simulate a mine water treatment structure.	36
3.2	Installation of bulkhead fittings.	38
3.3	Process flow diagram of bench-scale system.	39
3.4	Baffle located in reactors to inhibit the flow path.	40
3.5	Demonstration of reduction in noise from sensor data using bias resistors.	43
3.6	Typical $C(t)$ and $E(t)$ curves after (Fogler, 2005).	47
3.7	Raw data collected during pulse experiment.	48
3.8	Normalized $C(t)$ curve for pulse test conducted on bench-scale treatment system.	49
3.9	$E(t)$ curve for pulse test conducted on bench-scale treatment system.	50
3.10	Ideal mixed reactor $E(t)$ curve versus imperfect reactor $E(t)$ curves, after Fogler (2005).	51

4.1	Simplified representation of Mamdani controller.	53
4.2	Simulink block diagram used to evaluate the ANFIS controller.	55
4.3	Membership functions used to develop ANFIS controller.	58
4.4	Example of Sugeno model architecture.	62
4.5	Simulink block diagram used to collect training data.	62
4.6	Example of training data set	64
5.1	Test 1 results: control under steady-state conditions.	69
5.2	Test 2 results: control with varying flow rates.	71
5.3	Test 3 results: control with changing set-points.	73
5.4	Test 4 results: control while a surge in flow is experienced by the system.	75
5.5	Test 5 results: control with change in feed pH.	77
5.6	Test 6 results: control during change in reactor residence time.	79
5.7	Test 7 results: control during multiple perturbations.	81
5.8	Example of ANFIS training showing results from both acceptable (a) and poor (b) predictions.	84
5.9	Variation in pH_2 for all data sets used in the predictive analysis	85
5.10	DS-19 results comparing RMSE to the delay in predicting all data sets	86
5.11	Histogram showing the frequency of pH values for the predicting data set categories.	89
5.12	Boxplot of training error RMSE for all data sets. (DS-05 not shown due to high RMSE value)	90

LIST OF TABLES

2.1	Technology-based limits for discharge water (Skousen, 2003).	8
2.2	List of consent decrees by publicly traded coal companies between 2008 - 2014.	10
3.1	Measured volumes of reactors.	37
3.2	Average pump rates for the feed and treatment pumps used in the bench-scale system.	45
3.3	Comparison of residence times by experiment.	51
4.1	Summary of different tests ran using ANFIS controller.	54
4.2	Rule matrices for variables used in development of ANFIS controller.	59
4.3	Summary of variables changed during course of the ANFIS controller tests.	60
5.1	Table summarizing results of control tests	82
5.2	Data sets used for prediction with a statistical analysis of the pH data used in prediction	85
5.3	Error values of the most and least accurate predicting training data sets when used to predict the most variable checking data.	87

NOMENCLATURE

t_m	Mean Residence Time
τ	Mean Hydraulic Retention Time
Al	Alumnium
ALD	Anoxic Limestone Drain
AMD	Acid Mine Drainage
ANN	Adaptive Neural Network
ARD	Acid Rock Drainage
CAPP	Central Appalachia
CWA	Clean Water Act of 1972
D	Time Delay
DAQ	Data Acquisition Device
DS-XX	Data Set Test Number
EC	Electrical Conductivity
EPA	Evnironmental Protection Agency
Fe	Iron
GPM	Gallons per Minute
LOLIMOT	Local Linear Model Tree
MIW	Mine Influenced Water
Mn	Manganese
MPC	Model Predictive Control

NPDES National Pollution Discharge Elimination System

OS Prior Time Offset

PID Proportional Integral Derivative

RMSE Root Mean Squared Error

TMDL Total Maximum Daily Load

TSS Total Suspended Solids

CHAPTER 1

INTRODUCTION

1.1 Background

The National Pollution Discharge Elimination System (NPDES), per Section 402 of the Clean Water Act, requires permitting of all point source discharges. This program broadly applies to coal mining and mandates that all water, whether it be process water or storm water, must leave through a monitored NPDES outlet. Discharge points also denoted as outlets, outfalls, or monitoring points are typically assigned a rigorous set of required quality parameters, as described in the NPDES permit. Water leaving these discharge points is often referred to as effluent, and must meet or exceed these permitted parameters to satisfy an accepted level of compliance. Iron (Fe), manganese (Mn), aluminum (Al), total suspended solids (TSS), and pH are some common effluent limits monitored by the state environmental agencies and coal companies. However, additional parameters may be assigned on a case-by-case basis. Additionally, in recent years, these limits have been reduced from more liberal technology based limits to more rigorous anti-degradation limits, thus increasing the difficulty in maintaining compliant discharges. Often, raw mine water is deficient in meeting these strict effluent limits, and as a result, various treatment processes are required to bring the water into compliance prior to release into receiving streams. Numerous compounding factors including mining and geologic conditions, effluent limit parameters, and the type of contaminants present in the water can complicate the treatment process and cause the monitoring and treatment program to constitute a significant operating cost for the mine operator.

Acid Mine Drainage (AMD) constitutes one of the most significant and widespread water quality challenges in the Appalachian coal fields. AMD refers to a general lowering of water pH that occurs when water and sulfide minerals interact within oxidizing conditions in coal and metal mining. Acid Rock Drainage (ARD) is a term that applies more generally to non-mining causes such as highway construction and other large-scale excavations. It is functionally equivalent to AMD in terms of the

underlying causes and potential problems. Today, “mine influenced waters” (MIW) is becoming more popular as the general term referring to all types of mining-induced water issues. Due to the longstanding concern over AMD, substantial research over the last 40 years has been conducted with investigators studying various causal mechanisms and treatment strategies. This research has produced several pragmatic results, which provide numerous methods to predict, prevent, and treat AMD.

While the availability of various treatment options exist, coal operators still face significant constraints that limit their ability to mitigate AMD. Often, a typical treatment system will include the addition of alkaline material and coagulants. Unfortunately, various complicating factors, such as retention time, geographic location, topography, hydrology, and water chemistry limit the application of a standardized approach to water treatment (Akcil and Koldas, 2006). In surface mining, impacted water may only be present for a short time period during the mining process. Sites with a short treatment time will not warrant the same type of treatment system used at a location requiring perpetual treatment. Additionally, location can limit the types of treatment available (Gazea et al., 1996; Johnson and Hallberg, 2003). As an example, an outlet located on an on-bench sediment structure has less working area than a large pond at the base of a valley fill. Finally, water chemistry also dictates the treatment process available to mine operators. Water containing solids collected through turbulent flow requires different chemical applications than water with a low pH and high iron content. Ultimately, the overall system design must consider all of these factors simultaneously to effectively maintain environmental compliance.

In Central Appalachia (CAPP), typical AMD treatment systems utilize a series of ponds or discrete cells within a single large pond as reactor vessels to accommodate large flow volumes and relatively slow reaction rates. These earthen structures are oriented so that runoff from the site must flow sequentially through the ponds to reach the NPDES permitted outlet. Chemical treatment is added at the inflow to the system to provide sufficient retention time for the neutralization reactions and settling of solids to occur. Pond curtains or baffles may be added to the system to prevent short-circuiting and insufficient reaction time. Adjustments to the pH induces chemical reactions causing metals to precipitate and fall out of suspension from the water before exiting the system.

Due to the steep topography and isolated locations of mines in CAPP, NPDES outlets are often located in areas with limited access to utilities. Large mining complexes, comprised of both surface and underground mines, often have hundreds of outlets spread over thousands of acres. To mitigate these challenges, mines generally use environmental technicians to monitor the outfalls and administer chemical treatment. These technicians are often only able test the water chemistry of problematic outlets

once or twice daily. They then adjust the dosage of chemical agents as necessary to meet conditions at the time of testing. Perturbations in the environmental and flow conditions may be recorded. This data is generally not updated in real-time, and the treatment system is not adjusted until new measurements are made the next day. This lack of real-time control is inflexible and unresponsive to changing inflow conditions. As an example, a large rain event may lead to increased flow, reduced retention time in the ponds, poor treatment efficiencies, and a change in water contaminants resulting in the discharge of non-compliant water in-between technician visits. Overall, the extended lag in process control, at best, may lead to drastically over-designed treatment systems and, at worst, may cause companies to violate permitted effluent limits. Due to the aforementioned reasons, mining companies will greatly benefit from the automation of these treatment systems with real-time water monitoring as well as continuous treatment.

1.2 Motivation

Several coal producers, in CAPP have entered into consent decrees with the United States Environmental Protection Agency. The consequences of the consent decrees expose companies to daily fines for noncompliance events. Fines can range from \$1,000 to \$9,000 per day, depending on the severity and frequency of noncompliance (USEPA, 2008, 2009, 2011a, 2014, 2015). These large monetary penalties will require the industry to change the methods currently used to supervise mine water discharges in order to prevent costly regulatory penalties. Development of real-time monitoring and treatment technologies will benefit the CAPP coal industry through lower regulatory costs while enhancing the quality of discharged water.

Advancements in the chemical, electrical, and computer engineering fields have led to the development and widespread implementation of real-time data acquisition technologies for various industrial applications. By adapting these mature technologies from these other fields, new solutions and tools may provide mine operators a more efficient means of managing the chemical treatment of problematic outlets. Robust sensor suites and programmable logic controls can be deployed at the treatment sites to automate the chemical dosing based on real-time conditions. Additionally, an automated system can capture, archive, and distribute real-time data to mine personnel. When coupled with an automated alert mechanism, this data system can drastically reduce response time by rapidly initiating corrective actions. Regrettably, these technologies remain relatively unproven for mining environmental applications.

1.3 Objectives

Given the aforementioned opportunities, a research program was initiated to test the application of advanced logic controllers in a simulated mining environmental setting. The three primary objectives of this research include:

- Design, construct and characterize a bench-scale AMD treatment system in a controlled environment to simulate conditions as seen in water treatment structures used in CAPP coal mines.
- Devise, develop and implement an advanced logic controller that will autonomously administer doses of alkaline material to treat simulated AMD using the bench-scale system.
- Assess the performance of machine learning algorithms in predicting the future discharged water quality when given current inflow and treatment conditions.

The bench-scale system was designed to replicate an AMD treatment system found in use at many CAPP coal mines. Normally, mine water flows through a series of ponds before exiting through the permitted outlet. If the water does not meet effluent limits, chemical treatment (alkaline material) may be added at the inlet of the pond system to raise pH and precipitate metal oxides. The bench-scale system replicates this process through the use of multiple buckets that represent discrete cells within a pond structure. Pumps supply simulated AMD and an alkaline treatment chemical, which control flow rates and water quality within the system. This bench-scale system is fully instrumented with multiple water quality sensors and an advanced logic controller.

To control the pH in the bench-scale system, an autonomous logic controller was designed and implemented. Testing with this system demonstrates the feasibility of utilizing a similar controller in a full scale application. This controller can adapt to changing flow conditions and maintain a given pH set-point to meet the demands of an industrial setting.

Given the plurality of data generated during this project, a secondary objective was to assess the feasibility of using machine learning algorithms to predict future conditions from current data. Machine learning algorithms rely on large data sets, and in this project, several machine learning algorithms will be tested to determine the characteristics of input data sets that will produce robust predictive models.

1.4 Organization

Chapter 1 includes a description of background information and the motivation for this project as well as objectives and organization. This information provides an overview of the current practices and regulatory environment in CAPP as well as the proposed objectives to be completed in this thesis.

Chapter 2 provides a review of the history of AMD treatment research and regulatory requirements that must be met by coal companies operating in the CAPP region. Additionally, a review of the state-of-the-art practices in machine learning theory as it applies to control systems is discussed. This review provides a basis from which an advanced control algorithm can be created and used to regulate a nonlinear process such as pH adjustment.

Chapter 3 describes the development of a bench-scale AMD treatment system. This system has the capability to simulate AMD treatment ponds used by coal operators to treat water before it is discharged through a NPDES outlet. Next, the process of characterizing the system volumetric flows and flow patterns is presented. Finally, this characterization is compared to chemical reactor engineering principles which operate under ideal conditions to show the limitations of an ideal verses common treatment systems.

Chapter 4 introduces the methods and procedures used to build an advanced machine learning algorithm to automate the treatment of low pH AMD water. Additionally, the mechanisms used in an experimental testing regime is reviewed on a machine learning algorithm that can predict the future water pH given current inflow and treatment conditions. This system is used to identify the necessary characteristics a training data set needs to reliability develop an accurate machine learning controller.

Chapter 5 reviews the results acquired from operating a fuzzy controller with the bench-scale system. The outcome from multiple tests show how the controller overcomes a series of disturbances introduced into the laboratory-scale treatment system, including variable flow rates, changing set-points, and changes in water chemistry. Next, the results from the experimental testing procedures used to identify the necessary characteristics a training data-set requires are presented. These results indicate training data sets need specific qualities to accurately built advanced machine learning algorithms to control a nonlinear process.

Chapter 6 summarizes the key findings of this work and introduces opportunities for further research and development.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Currently, the treatment of mine water discharges is largely controlled through the manual application of chemicals to treatment ponds. This type of treatment is necessary to comply with effluent limits that are established during the permitting process. Increased regulatory scrutiny has driven the traditional use of technology based-parameters to a more stringent standard based on anti-degradation guidelines. Additionally, the United States Environmental Protection Agency (EPA) has levied enforcement action upon CAPP coal companies through the use of consent decrees, which order large civil penalties in addition to the development of environmental management systems.

The chemical and physical process which cause AMD have been well researched and documented over the past 40 years (Hill, 1969; Hoehn and Sizemore, 1977; Hedin et al., 1994; Gazea et al., 1996; Ziemkiewicz, 1998; Skousen et al., 2000; Ziemkiewicz, P. F.; Skousen, J. G.; Simmons, 2003; Johnson and Hallberg, 2005; Akcil and Koldas, 2006; Kalin et al., 2006). This research has developed several innovative methods to treat AMD. Unfortunately, the numerous factors and site-specific nature of AMD generation prevents the development of a standardized treatment strategy. When possible, passive forms of AMD treatment are implemented to lower the treatment cost. However, when this type of treatment is not feasible, active treatment options must be utilized. Active treatment relies heavily on the use of manpower to monitor and control the quality of discharged water. For these reasons, automation of the monitoring and treatment process can help eliminate non-compliant water discharges by reducing the need to rely on intermittent manual manipulation of treatment chemicals.

Several advances in the fields of computer, electrical and chemical engineering have provided a variety of models used to control water treatment processes, including the

control of pH (McAvoy et al., 1972; Gustafsson and Waller, 1992; Henson and Seborg, 1994; Ghee et al., 2002; Yu and Gomm, 2003; Singh et al., 2009; Ibrahim, 2010; Navghare et al., 2011; Petchinathan et al., 2014). The use of Proportional-Integral-Derivative (PID) controls, fuzzy logic, and Adaptive Neural Networks (ANN) have shown value by simplifying complex operations in daily routines. Regrettably, the CAPP region has been slow to adopt these technologies in environmental applications. These mature control technologies have been used in processes similar to AMD treatment. Therefore, the potential exists to adopt these technologies for use in mine water treatment. In spite of these opportunities, several challenges prohibit the direct implementation of these technologies, including, limited access to utilities at outlet sites, highly variable flow rates, and site specific treatment requirements.

2.2 Review of Mine Water Discharges

Mining companies are required by law to only discharge water within the permitted effluent limits as described in the mining permit (Heishman and Mclusky, 2012). This requirement implies that any water on the permitted area of the mine becomes the responsibility of the mine operator. All water, including rainfall, process water, and any naturally occurring water features flowing through the property must be discharged in accordance to permitted limits. Should the operator fail to meet these permitted limits, they can receive an environmental violation as well as fines associated with the discharge of noncompliant water from permit limits. Recently, the discharged water quality for mine sites has come under additional regulatory scrutiny, thereby lowering the permitted effluent limits and further adding to the penalties mining companies already incurred. This heightened sense of enforcement coupled with reduced effluent parameters suggests new technologies as well as improved practices are required to ensure faultless compliance with future permitted limits.

2.2.1 NPDES & the Clean Water Act

Mine operators must meet water quality standards as prescribed in the Clean Water Act of 1972 (CWA). This act is the basis for regulating the discharge of pollutants into the waters of the United States. Waters of the United States, for the purpose of the CWA and as applied to mining in CAPP, may be defined as all lakes, rivers, streams and tributaries which exhibit a bed, banks and high water marks (USEPA, 2004).

The NPDES, per Section 402 of the CWA, outlines permitting of all point source discharges from coal mines. This program dictates that all water leaving a permitted boundary, whether it be process water or storm water, must leave through a monitored NPDES outlet. These permitted NPDES outfalls must meet a rigorous set of effluent

Table 2.1 – Technology-based limits for discharge water (Skousen, 2003).

Effluent Parameter	Avg. Monthly Concentration (mg/L)	Max, Daily Concentration (mg/L)
Fe	3.0	6.0
Mn	2.0	4.0
TSS	35.0	70.0
pH	6.0 (min)	9.0 (max)

limits to maintain compliance. Common effluent limits monitored by coal companies include but are not limited to iron, manganese, aluminum, total suspended solids, and pH (Skousen et al., 2000). Raw mine water is often deficient in meeting these effluent limits, and as a result, various treatment processes are required to bring the water into compliance prior to environmental release. Depending on the mining and geologic conditions, water monitoring and treatment programs can constitute a significant operating cost for the mine operator.

Mining permits receive limits to discharge effluent based on one of three criteria: technology-based, water quality-based, and best professional judgment (Skousen, 2003). Historically, mining permits received limits in accordance to technology-based standards as seen in Table 2.1. Recently, mining permits have been assigned effluent limits using the more rigorous water quality-based method using Total Maximum Daily Load (TMDL). A TMDL is a load limit or maximum amount of of pollutant a waterbody can receive; furthermore, the use of TMDL limits can be viewed as an implementation of a pollution reduction plan (Skousen, 2003). Additionally, the CWA requires states to develop lists of impaired streams that do not meet water quality standards based on a designated water use category, even after point sources of pollution have been controlled and pollution control technologies have been installed under the NPDES program (Skousen, 2003). NPDES outlet limits upstream of impaired waters are assigned effluent limits based on the TMDL of the impaired waterways. Typically, these limits are stricter than the technology-based limits. Furthermore, when NPDES permits that previously had technology-based limits are renewed, the more stringent method of assigning limits may be applied to the preexisting outlets. This practice adversely affects coal companies when discharged water that was previously in compliance, now requires enhanced treatment methods to meet these new limits.

The CWA also mandates that a state must develop an anti-degradation policy, and this policy can also play an obscure role in determining NPDES limits (Heishman and Mclusky, 2012). The purpose of the anti-degradation policy is to prevent streams that

are achieving water quality criteria and meeting designated uses from significantly declining in quality (Heishman and Mclusky, 2012). This policy assigns four tiers of protection to waterways within the state. The tiers are ranked in order from lowest to highest level of protection by levels 1, 2, 2.5 and 3. Tier 3 waters are typically found in protected areas such as national forests and state parks and therefore cannot be degraded by outlets. Tier 2.5 protected waters are designated as naturally reproducing trout streams and other streams with a high biological and aquatic life score. The default level of protection is Tier 2, and these waters typically exhibit better water quality than the water quality standards established for that stream (Skousen, 2003). Tiers may also be assigned on a pollutant basis so streams can belong to multiple tiers based on the levels of pollutants within the receiving stream. Each tier has set limits on the amount of degradation that is allowable and NPDES limits are adjusted accordingly.

Mines must monitor the discharge of water at NPDES outlets on a random basis, with a minimum frequency of twice per month. In practice, an independent contractor normally conducts this sampling plan, and the schedule for sample gathering is unknown to the coal operator. This blind testing is done to prevent any bias in regard to the sampling plan. Permits usually dictate all outlets that discharge water to be tested for water quality parameters at least twice per month; however, large storm events and problematic outlets may dictate additional sampling. Typically, an environmental technician will visit an outlet to acquire a sample if the monitoring point is discharging water. This sample is then labeled and stored in a refrigerated environment for preservation. After all samples have been collected, they are transported to an independent lab for testing. Sample test results verify that the sample meets daily and monthly effluent parameters as described by the individual mining permit. Historically, results from these tests take several weeks to return from the lab to the mine environmental department. With the increased regulatory scrutiny mining companies are facing, this lead time has been reduced to approximately 24 hours. Nevertheless, the information from these results only inform the operator of a past excursion. Had the operator known of the excursion in water quality parameters prior to sampling, a proactive plan could be implemented to avoid the discharge of water outside the permitted limits, before sampling is conducted.

2.2.2 Consent Decrees

Between 2008 and 2014, every major publicly-traded CAPP coal company entered into a consent decree(s) with the EPA (USEPA, 2011b, 2015, 2009, 2008, 2014, 2011a). These settlements were required to resolve alleged violations of the CWA for exceedances in water quality associated with NPDES permit limits for manganese, total suspended solids, selenium, aluminum, pH, chloride, and iron. Table 2.2 lists the

Table 2.2 – List of consent decrees by publicly traded coal companies between 2008 - 2014.

Company	Date Signed	Civil Penalty (\$)
Massey Energy Company	04/09/2008	20,000,000
Patriot Coal Corporation	02/05/2009	6,500,000
Arch Coal, Inc.	03/01/2011	4,000,000
CONSOL Energy	03/14/2011	5,500,000
Alpha Natural Resources	03/05/2014	27,500,000
Arch Coal Inc. (ICG)	08/06/2015	2,000,000

company, the date the consent decree was signed, and the civil penalty attached to the consent decree.

The consent decrees mandate each company to establish an Environmental Management System (EMS) to track and alert the agency of any violations. Additionally, companies also face steep penalties for any excursion from permitted limits. These penalties can exceed \$9,000 per day for consecutive violations at outlets with persistent noncompliance issues. This elevated enforcement category means any outlet that exceeds effluent parameters three or more times within a rolling 12 month period will be subjected to these heightened fines.

While the allegations from the EPA appear to imply gross negligence from the CAPP coal industry, the number of excursions is representative of only a small percentage of overall compliance. For example, according to Kitts (Alpha Natural Resources, 2014), in 2013, Alpha Natural Resources had a combined total water quality compliance rate of 99.8%. Additionally, Kitts reported that this compliance rate is based on more than 665,000 chances to miss a daily or monthly average. To withdraw from the consent decree, a company's subsidiary must achieve a 100% compliance level for three years. As with most natural processes, the ability to attain perfect compliance ranges from unattainable to impractical at best using the current practices.

2.2.3 Description and Causes of AMD

In Appalachia, AMD constitutes one of the most significant and widespread water quality challenges that must be treated per NPDES guidelines. AMD is a general lowering of water pH that occurs when water and sulfide minerals interact within oxidizing conditions in coal and metal mining. The sulfide materials contained in the geologic formation being mined are typically liberated through the mining process. This liberation increases the quantity of exposed sulfide mineral surface area leading to an elevated reaction rate. For example, blasting operations liberate these minerals and increase the contact area resulting in enhanced exposure to the oxidizing agents air and water. Likewise, underground mining operations create voids that are filled

with air and water, which increases reaction rates at the perimeter of the openings. In undisturbed deposits, generation of acidic material is minimal and is a function of natural erosion via weathering. Given the slow generation rate, this minimal release from in situ deposits does not cause irreparable harm to the subsequent receiving streams. The mining conditions causing the generation of AMD have led to substantial research over the last 40 years, with investigators studying various causal mechanisms and treatment strategies (Hill, 1969; USEPA, 1994; Gazea et al., 1996; Ziemkiewicz, 1998; Skousen et al., 2000; Kleinmann, 2001; Johnson and Hallberg, 2003; Akcil and Koldas, 2006; Kalin et al., 2006). This research has produced several pragmatic results, which provide numerous methods to predict, prevent, and treat AMD.

Within the CAPP coal-mining region, pyrite and marcasite (both forms of FeS_2) are the predominant sulfides that cause AMD (Hill, 1969; Hoehn and Sizemore, 1977; Ziemkiewicz, 1998; Skousen et al., 2000; USEPA, 2000; Kleinmann, 2001; Johnson and Hallberg, 2003). Reaction 2.1 is indicative of this oxidation process:

where either molecular oxygen or ferric iron is acting as the oxidant in the reaction (USEPA, 1994). Here, the sulfur is oxidized to form hydrogen ions and sulfate, which are the products required for sulfuric acid.

Additionally, the soluble iron byproduct (Fe^{2+}) is left in solution and has the capability to react further as described in Reaction 2.2:

where ferrous iron materials are converted to ferric ions slowly at low pH values (USEPA, 1994).

The presence of certain types of bacteria also plays an important role in the generation of AMD. For example, at pH values less than 3.5 when in the presence of the iron bacterium *Thiobacillus ferrooxidans*, Reaction 2.3 may occur:

where the presence of the aforementioned bacteria will allow the pyrite to be dissolved if it is in contact with the ferric ion (USEPA, 1994).

Finally, the ferric iron precipitates from the AMD as hydrated iron oxide as shown in Reaction 2.4:

Figure 2.1 – Streams with orange staining from the iron oxide precipitate “yellow boy.”

where the hydrated iron oxide precipitates into an orange deposit on stream bottoms, commonly referred to as “yellow boy” (USEPA, 1994). This hydrated iron oxide precipitate is commonly seen in CAPP streams with high flows of acidic mine water and is indicative of AMD pollution. Figure 2.1 shows water features, which have been impaired by the precipitation of iron oxides.

Akcil and Koldas (2006) have investigated the primary factors that determine the rate of AMD generation. This study as well as others mentioned above have concluded that numerous parameters including water temperature, water pH, oxygen concentration, degree of saturation, bacteria content, and the presence of alkaline material in the host rock, all significantly affect AMD generation. Since these parameters can vary considerably from site to site and even within the same site, the ability to apply a comprehensive treatment solution is severely limited. As a result, AMD treatment must be carried out using site specific plans.

2.2.4 Effects of AMD on the Environment

Acid mine drainage has been described by many researchers as a severe environmental problem facing mining operations (Hoehn and Sizemore, 1977; USEPA, 1994; Akcil and Koldas, 2006; Kalin et al., 2006). This longstanding problem causes many detrimental effects to streams adjacent to mining operations as well as water tables and downstream tributaries accepting inflows of AMD. Regrettably, the mechanisms causing AMD can still perpetuate the generation of acidic water many years after a mine is reclaimed and underground workings are sealed (Hill, 1969). The environmental

problems caused by AMD have been well documented and include impairment in growth and reproduction rates for aquatic plant and animal life. Additionally, AMD may infiltrate into the water table and in extreme cases contaminate drinking water supplies for residents who depend on wells for potable water (Jiménez et al., 2009). Due to the severe and longstanding impacts of AMD, new technologies and methods are required to effectively treat and prevent the generation of AMD with minimal disruption and cost to the mining process.

2.3 Treatment of AMD

Despite the availability of various treatment options, coal operators face significant constraints that limit the number of effective options to mitigate AMD. While a typical treatment system includes the addition of an alkaline material and flocculants, various complicating factors, such as retention time, geographic location, topography, hydrology, and water chemistry limit the application of a standardized approach to water treatment (Akcil and Koldas, 2006). Treatment time-scales vary in length from a few months to indefinitely. Sites with a short treatment time will not benefit from the same economies of scale as a perpetual treatment site. Additionally, location in terms of topography, can limit the types of treatment available (Gazea et al., 1996; Johnson and Hallberg, 2003). For example, an on-bench sediment structure with a NPDES outfall will typically have less working area than a large pond at the base of a valley fill. Finally, the mine water chemistry drives the treatment process available to mine operators. Water containing solids from turbulent flow requires different chemical applications than water with a low pH and high iron content. Given the aforementioned constraints, the overall AMD treatment system design must consider all of these factors simultaneously. Finally, water treatment at coal operations can be classified as active or passive. Passive treatment systems utilize man-made structures to enhance the quality of water based on the contents of the structure. Alternatively, active treatment requires the addition of chemicals or treatment plants to bring water up to permitted limits.

2.3.1 Passive AMD Treatment

Passive treatment systems improve the quality of AMD through the use of naturally occurring geochemical and biological processes (Gazea et al., 1996; Skousen et al., 1997). These systems were first proposed by Huntsman et al (1978), as well as, Weidner and Lang (1982) observing improved AMD water quality after it passed through natural sphagnum bogs in Ohio and West Virginia. Today, passive treatment systems consist of man-made structures that can reproduce conditions in these wetlands found in nature. These structures are similar to sediment ponds; however, they are usually shallow and contain features similar to a marshland. Often, passive treat-

ment systems are preferred due to their low cost and ease of maintenance (Gazea et al., 1996; Skousen et al., 1997). Unfortunately, these systems are not applicable to all AMD flows due to a number of limiting constraints. Because of this, different passive treatment systems are often implemented in series to one another. Various options for passive treatment include: aerobic wetlands, anaerobic wetlands, and anoxic limestone drains. Each is described in further detail below.

Aerobic Wetlands

Aerobic wetlands are man-made bogs or marshes and treat AMD by oxidation. This general increase in pH causes metals to precipitate from solution, thus improving water quality. Several researchers have described the use of aerobic wetlands to explicitly treat AMD, including: Hedin et al. (1994), Gazea et al. (1996), Skousen et al. (1997), Ziemkiewicz et al. (2003). Aerobic wetlands are typically constructed to approximate the conditions of wetlands found in nature. The basins are constructed of an impermeable layer to prevent seepage and the depth is generally shallow ($< 2'$). Additionally, a substrate such as soil is placed in the basin to promote the growth of vegetation. Finally, the effectiveness of the wetland can be improved through the incorporation of structure to promote aeration, such as waterfalls or steps between cells.

This type of passive treatment is most effective when treating alkaline water. The aerobic wetlands ameliorate the MIW by providing a sufficient residence time to allow oxidation of the metal constituents to form and precipitate their respective metal hydroxides. The effectiveness of this type of treatment is dependent upon the initial dissolved metal concentration, dissolved oxygen content, pH, net alkalinity of feed water, presence of certain microbial species, and retention time in the wetland (Skousen et al., 1997). The plants within these structures maintain the substrate of the designed wetland, while also providing structure to inhibit flow, promote microbial growth, and increase residence time. Frequently, this type of treatment system is preceded by an anoxic limestone drain, which increases the alkalinity of water flowing into the structure.

Anaerobic (Compost) Wetlands

Anaerobic wetland ponds are used to add alkalinity to acidic water. Unlike aerobic wetlands, anaerobic or compost wetlands are constructed for the treatment of net acidic water. These systems are constructed to promote flow of water under the surface and through the substrate, in an anaerobic environment. Generally, this substrate

consists of mushroom compost, horse or cow manure, hay bales, peat, wood chips or sawdust (Gazea et al., 1996). The design of the anaerobic system promotes water flow through this substrate to stimulate chemical and microbial reduction reactions that precipitate metals and neutralize acidity within the waters described by Skousen et al. (1997). Additionally, limestone may be placed at the bottom of the wetland to generate alkaline material through dissolution. As a standard design, these wetlands consist of a cell or multiple cells with under-drains at the bottom of the structure allowing the water to pass through the substrate and out of the structure. These particular structures generate alkalinity in one of two ways as described by Hedin et al (1994). The first method of alkaline generation is through bacteria using the organic substrate as a carbon source. These reactions are described by Skousen et al. (1997) as follows:

where sulfate is converted to hydrogen sulfide and bicarbonate. Alternatively, limestone within the wetland can react with acidity in the water as follows:

where limestone reacts with acidic water to produce free calcium and bicarbonate. The anaerobic wetlands are more efficient at treating acidic water, with high Fe levels, and high dissolved oxygen content; however, like their aerobic counterparts, these systems are most successful when used to treat small AMD flows (Skousen et al., 2000). More recently, anaerobic wetlands have seen a resurgence within CAPP due to the addition of selenium as an effluent limit to many permits. Several coal operators are installing anaerobic wetlands to promote anaerobic bacteria which reduce the quantity of selenium in water (CH2M HILL, 2010).

Anoxic Limestone Drains

Anoxic Limestone Drains (ALD) are excavated channels or beds that are filled with baseball-sized limestone and covered to prevent the channel from interacting with the open atmosphere. These structures incorporate a plastic layer between the limestone and earthen channel and backfill to prevent sediment from clogging the channel. Water treated with an ALD must meet specific requirements (Hedin et al., 1994; Gazea et al., 1996; Skousen et al., 1997; Ziemkiewicz, P. F.; Skousen, J. G.; Simmons, 2003). For example, ALDs are only effective when the water feeding the ALD is low

in Fe^{3+} , has a relatively high acidity, and low dissolved oxygen content. Additionally, the water feeding an ALD should be net acidic to gain a benefit from passing through the structure. Next, the water must have low levels of dissolved oxygen, Fe and Al. This chemistry is necessary to prevent ferric oxides from forming a scale on the surface of the limestone rocks. Should such a scale form, the effectiveness of the ALD will be reduced due to the decreased limestone surface area contacting the AMD. In more severe cases, aluminum oxides may form plugging the drain with precipitate. This blockage prevents water from flowing through the drain and requires excavation and cleaning to return the drain to a functioning system. For these reasons, ALD's are typically used to treat water seeps generated from an underground source and are ineffective when treating surface waters. Furthermore, ALDs are useful when placed ahead of an aerobic wetland when the water is net acidic and the water quality does not permit treatment with the aerobic wetland.

2.3.2 Active AMD Treatment

Active treatment refers to the direct addition of chemicals to the impaired water to cause reactions that will render the water compliant after treatment. While it is normally more costly and labor intensive than passive treatment, mine operators are compelled to use active treatment, when conditions for passive treatment are unfavorable, such as: limited area, high flow rates, short time periods where impacted water is present, and tighter regulatory effluent limits. In addition to the usage of chemicals, the presence of oxygen in the water is also an important factor for designing a treatment system. Dissolved oxygen in AMD can assist in raising the pH of the water and facilitates the precipitation of metal hydroxides. For this reason, additional chemicals or mechanical aerators are sometimes used to increase the oxygen content of the water.

Chemical Treatment

The chemical treatment of MIW is primarily governed by pH control where neutral to alkaline pH values are required to reduce the amount of dissolved metals in the acidic water. In addition to the acidity, low pH mine water also contains dissolved metals including iron, manganese, and aluminum. Metal solubility is a function of water pH, and the metal concentration can be readily controlled by manipulating the pH (often by raising) until the metals precipitate. Given the large flow volumes and relatively slow reaction rates of AMD, standard treatment systems utilize a series of ponds or discrete cells within a single large pond as reactor vessels. These earthen structures are oriented so that runoff from the site must flow successively through the ponds to reach the NPDES permitted outlet.

Chemical treatment is added at the inflow to the system to provide maximum retention time for the neutralization reactions to occur. Pond curtains or baffles may be added to the system to prevent short-circuiting and insufficient reaction time. As water flows through the ponds, the pH adjustments and chemical reactions cause metals to precipitate and fall out of suspension before exiting the system (Skousen et al., 2000; Jiménez et al., 2009).

Alkaline chemicals are used in pH control of MIW facilitate the precipitation of metals from the water. According to Skousen (1997), six primary alkaline chemicals are used to treat AMD, including: (1) limestone, (2) hydrated lime, (3) pebble quick lime, (4) soda ash, (5) caustic soda, and (6) ammonia. Additional chemicals such as flocculants, coagulants, and oxidizing agents may also be used to enhance the performance of the treatment system.

Coal mines generally employ environmental technicians to monitor the discharges and supervise the chemical treatment of mine water. The water chemistry at problematic outlets is usually tested once or twice daily. These technicians then manually adjust the dosage of chemical agents as necessary to meet conditions at the time of testing. The inability of technicians to constantly monitor the treatment system results in long time periods between the testing and adjustment procedures. This lack of real-time control is fairly inflexible and non-responsive to changing inflow conditions. Overall, the extended lag in process control, at best, may lead to drastically over-designed treatment systems and, at worst, may cause companies to violate permitted effluent limits.

2.4 pH Control Systems

Since MIW treatment relies heavily upon pH adjustment, the optimal method for pH control must be considered. Unfortunately, the pH neutralization process is highly nonlinear, and as a result, fine process control can be difficult to implement. As an example, ideal titration curves are shown in Figure 2.2 for monoprotic and polyprotic acids. The steep slope of the curves between the upper and lower pH values causes particular difficulty for control systems trying to maintain a neutral target pH. In this region, small perturbations in the system variables tend to cause rapid changes to outgoing pH. Additionally, unlike many chemical process applications, the pH of mine water is only controlled in one direction through addition of alkaline material alone, whereas other industrial process can both raise and lower pH by the addition of both acids and bases. This limitation increases the difficulty of maintaining a steady pH level when natural perturbations within the flow regime are encountered.

Several researchers in the chemical, wastewater treatment, and industrial process fields have investigated control strategies used to regulate pH. McAvoy, Hsu and

Figure 2.2 – Ideal titration curves of monoprotic and polyprotic acids after McMurry and Fay (1998).

Lowenthal (1972) presented a rigorous derivation of dynamic models for pH in stirred tank reactors. This work has led several researchers to further address the subject of pH control schemes directed toward the aforementioned industries (Gustafsson and Waller, 1992; Henson and Seborg, 1994; Zhang and Morris, 1999). This body of research demonstrates that a nonlinear control system provides increased accuracy when compared to a traditional linear control system when pH is the control variable. Unfortunately, this research also uses the model of a continuously stirred tank reactor with the addition of both acidic and basic chemicals to control pH, thus limiting the direct application of this research to uni-directional AMD treatment.

2.4.1 General Characteristics of Control Systems

When implementing a control system, one must first define the characteristic feedback of a controller. As shown in Figure 2.3, the *transient response* is the period of time it takes a controller to reach a state of equilibrium from a previous state of equilibrium following a disturbance. As seen in the figure, the controller is at equilibrium at time zero before control is started. Once the control system is initiated, a *transient response* is observed from time zero until the controller reaches a state of equilibrium. At this point, the system characteristics may be described as a *steady-state response*, which is an approximation to the commanded or desired response (Nise, 2007).

The desired response or *set-point* is shown by the horizontal line labeled *Set-Point* in Figure 2.3. This point is a set value that the controller is designed to reach and maintain by manipulating the *control variable*. The *control variable* is output from the system that the controller uses as feedback. Once the system reaches steady state, any offset between the actual and desired values of the *control variable*, is denoted the *steady-state error*. Error is usually non-zero and is a system specification, Figure 2.3 (necessarily) shows an exaggeration of error magnitude.

Figure 2.3 – Example of response characteristics of a control system after Nise (2007).

A PID controller is a very common traditional control approach. Integral, proportional, and derivative feedback is based on the past (I), present (P), and future (D) control error (Astrom and Hagglund, 2001). This type of controller is widely used in industrial process control, and its significance is evident in its broad utilization. As recently as 2001, an estimated 90% of all control loops used a type of PID control (Astrom and Hagglund, 2001). Unfortunately, PID control does not adequately manage nonlinear processes as noted by many researchers (Carvajal et al., 2000; Astrom and Hagglund, 2001; Visioli, 2001). Additionally, PID controllers can be cumbersome to properly tune and the performance of the controller degrades when operating conditions vary from tuning conditions. These reasons have led to research where alternative control schemes and tuning procedures have been implemented for a PID controller operating with nonlinear and dynamic process models.

Given the limitations of PID controllers, new research has yielded benefits to PID control via the addition of alternative control schemes. While PID alone is not well suited for nonlinear control, when coupled with recent technological advances, it does have the ability to manipulate nonlinear systems. When combined with a fuzzy controller, the PID controller is capable of automatically tuning the variables related to gain; this type of control has been implemented by many researchers allowing the PID control scheme to adapt to a nonlinear process (Misir and Malki, 1996; Carvajal et al., 2000; Visioli, 2001; Ghee et al., 2002).

2.4.2 Artificial Neural Networks

Neural networks are fundamentally a pattern recognition system inspired by biological processes and the way the human brain functions. The use of neural networks for nonlinear system identification has been proven successful by a number of researchers (Narendra and Parthasarathy, 1990; Zamarreño and Vega, 1998; Yu and Gomm, 2003). Mechanistically, these algorithms are used to estimate outcomes and predict relationships using large data sets containing known inputs and outputs. These data sets include a large number of input variables, that can exceed the limits of traditional multi-linear regressive techniques.

In general, an ANN consists of many interconnected, simple processing units known as neurons or nodes (Jain et al., 2014). Figure 2.4 shows a diagram of a simple neural network. ANN's develop mathematical relationships between the input and output variables using "hidden layers" defined by a particular mathematical function and weighting factor. The mathematical functions associated with each node are referred to as activation functions. A detailed list of activation functions can be found in research presented by Laudani (2015). The nature and value of these activation functions can be determined (i.e. "trained") using a number of nonlinear optimization expressions. The knowledge gained from neural networks is implicitly represented in patterns of interactions between these network components. The neural network approach differs from a simple regression model because there is no assumed relationship between the input and output variables. Instead, the relationship between variables is established through an iterative process (King, 1999). Often neural networks are combined with other forms of controllers, such as, Model Predictive Control (MPC) or fuzzy logic to increase the robustness of the control system.

The back-propagation method is commonly used to train ANN's. Here, a large set of inputs and outputs are provided as a framework to what a response from the network should achieve. As the ANN processes this data set, the difference between the predicted and actual outputs, or error, is used to determine the ability of the ANN to correctly predict the desired output. This error is propagated backward throughout the network to manipulate the activation functions, resulting in a more accurate output with each iteration. The theory is that once an acceptable level of error is achieved, the learning process stops and the trained activation functions can be used to predict outputs on independent data sets. When an ANN is trained using both inputs and outputs, it is referred to as supervised learning since the outcomes are already known (Jain et al., 2014). As this method indicates, the predictive ability of the ANN is dependent on the supervised data used to train the network. This limitation can be problematic when trying to apply an ANN to a new process where input-output relationships have not been established.

Figure 2.4 – Example of neural network with 4 input, 3 hidden, and 2 output nodes after Kim (2014).

ANN's have demonstrated the ability to derive highly nonlinear relationships, and they can be continually updated as more data is collected. Furthermore, this process modeling approach has been successfully implemented by a number of researchers to develop accurate model predictions in a number of application areas, including fuel cells, batteries, heat exchangers, chemical reactions, and surface water quality parameters (Yu and Gomm, 2003; Singh et al., 2009; Vasickaninova et al., 2011; Shen et al., 2013; Chen et al., 2014; Elbisy et al., 2014).

The main disadvantage to using neural networks lies in the "black box" nature of the model development. ANN's can be used to approximate many types of functions; however, the architecture of the ANN does not provide insight on the structure of the function being approximated. As a result, ANNs, by definition, exclude the inclusion of any fundamental relationships inherent to the system. This deficiency is particularly noteworthy for the issue of pH control, as a tremendous body of scientific work has identified many of the causal factors (Bhatt, 2006; Viswanathan et al., 2006; Li et al., 2010; Ahmad et al., 2010; Booth and Mead, 1969; Bridle, 2006; Schulze and Eberhard, 1992).

2.4.3 Fuzzy Logic

Fuzzy logic provides a structure that allows computers to evaluate variables that do not have a discrete definition. A classic example of this vagueness lies in the question, "Is the glass of water half full or half empty?" Some people will answer half full, while other will respond with the alternative answer, half empty. Rarely, will an individual use a definitive answer such as, the glass contains 4.31 ounces of liquid. This ability to qualitatively assign values to inexact conditions shows the power of human comprehension and judgment, as compared to Boolean logic which comprises the basics of computing systems.

Zadeah (1965) originally introduced fuzzy logic as a way to quantitatively assign classes of objects with continuous grades of membership that do not have sharply defined boundaries. This research used set theory notation to explain the mathematical principles behind fuzzy logic. While fuzzy logic is a commonly misunderstood theory, it simply applies mathematics as an approach to mimic the way humans think and communicate. The work conducted by Zadeah spearheaded the framework which allows computers to quickly process linguistic rules. Upon introduction, fuzzy logic was not well accepted by the academic community (McNeill and Freiburger, 1994; Zadeh, 2008); however, the use of fuzzy logic has grown in both number and variety in practical applications used in the field of mathematics and physical science (Zadeh, 1996).

The mechanics of fuzzy mathematics involve the manipulation of fuzzy variables

Figure 2.5 – Example of membership functions used by Matlab after Jang et al. (1998).

through a set of functions representing linguistic definitions which take the form of if-then rules (Hayward and Davidson, 2003). The variables used in fuzzy logic belong to different linguistic classes through the use of membership functions. Common functions used in defining the fuzzy regions are shown in Figure 2.5. The most commonly used membership function is triangular as this function is computationally simple to define and evaluate (Hayward and Davidson, 2003). These functions overlap to allow a variable to belong to more than one membership function. In the previously discussed glass of water example, the variable *volume* may belong to more than one membership function if the variable is contained by both sets. An example of the ability to belong to multiple membership functions is shown in Figure 2.6. Continuing the glass of water example, the value 1.2 ounces belongs partially to the “empty” and “partially full” membership functions. The degree of membership is established by a weighted average using the values $W1$ and $W2$.

Today, fuzzy logic is used in many aspects of our everyday lives, such as: washing machines, cameras, toasters, subways, air conditioning thermostats, financial trading programs and many other industrial control applications (Dutta, 1993). As the uses for fuzzy logic have grown dramatically, so has the definition of what fuzzy logic means. Taking a broad view, fuzzy logic has evolved into more than just a logical system. More recently, according to Zadeah (2008), fuzzy logic has many facets, including: logical applications, fuzzy-set-theoretic facets, an epistemic facet, and the

Figure 2.6 – Example of membership function for the volume of a glass of water.

relational facet.

2.4.4 Fuzzy Controllers / Mamdani Controllers

In 1974, Zadeah's work was expanded upon by E. H. Mamdani by virtue of the introduction of the first fuzzy logic controller (Mamdani, 1974). Mamdani presented a controller which manipulated the throttle and heat applied to a steam engine by the direct use of if-then rules to control the steam engine plant. This work is representative of the first attempts to apply fuzzy logic to a control scenario. As noted by Mamdani, the use of fuzzy logic allows computers to quickly process data with a minimal storage requirement; however, one drawback of this method is that the accuracy of the controller will depend on the experience of the programmer.

Expanding on this work, Mamdani (1976) revisited the topic of using fuzzy logic for controllers in 1976. Here, Mamdani reviewed the work of several researchers utilizing linguistic controllers to automate pilot-scale plants. Kickert and Lemke (1976) used fuzzy logic to build a controller for a warm water plant. The researchers compared the results of the fuzzy controller to a PI type controller. The results indicated the fuzzy controller showed a faster step response than the digital controller with an equal amount of accuracy in regard to temperature control. Next, work by Rutherford and Bloore (1976) show two key points in the development of fuzzy logic controllers. First, the results confirmed the successful application of fuzzy logic presented by Zadeah in 1965. Second, this research demonstrated that fuzzy controllers could be easily applied in an industrial setting. The aforementioned researchers along with others allowed Mamdani to conclude that fuzzy logic is applicable to industrial plants offering difficult control schemes. Finally, the heuristic nature of fuzzy logic control follows that of the traditional PI controller, while also adding the ability to consider nonlinear

processes.

Fuzzy logic controllers provide a method to model continuous states of data as discrete numbers. This outcome is accomplished through the use of membership functions and a “rule base.” For example, the previous glass of water example may be qualitatively defined by a fuzzy set by specifying the content as “empty”, “partially-full”, “full,” or some mixture or overlap between sets. A fuzzy logic interpreter (i.e. a “fuzzification” algorithm) can then transform a quantitative measurement (e.g. 1.2 ounces of liquid) into the fuzzy set, based on the value of the respective membership functions as shown in Figure 2.6.

A fuzzy operator can then specify the appropriate action to be made, based on the conditions in the rule base associated with the fuzzy set (e.g. if the glass is “empty,” then fill with water “very quickly”). A defuzzification algorithm and the associated membership functions can then transform the fuzzy action back to a quantitative value for control purposes (e.g. pump water into the glass at a rate of 3 ounces per minute). However, the value of this approach is that the break-off between fuzzy sets need not be sharp. In the working example, an 8 ounce glass containing 1.2 ounces of liquid can exist somewhere between the “empty” and “partially-full” fuzzy sets. Given these vague classifications, fuzzy logic can be particularly useful in a variety of control applications where the input-output relationships are not specified by explicit mathematical functions (Berenji, 1992).

2.4.5 Adaptive Neuro-Fuzzy Inference Systems

A more recent adaptation of fuzzy controllers is called neuro-fuzzy modeling. Here, the membership functions and the rule base is generated from large sets of numerical data which represent the behavior of a system (Nilashi et al., 2011). This adaptation is accomplished by combining the principles of fuzzy logic with the pattern recognition abilities of a neural network.

The union of fuzzy logic controllers and ANN’s produces Adaptive Neuro-Fuzzy Inference Systems (ANFIS), a nonlinear control technique that has promise in controlling pH at mine water treatment sites. This control scheme is a combination of ANN and fuzzy logic. ANFIS provides a scientific basis for nonlinear process modeling with both empirical and analytical components. The resultant system provides substantial benefit over purely empirical approaches that disregard years of fundamental research. However, ANFIS systems do not necessarily require complete understanding of the complex chemo-physical interaction occurring within the system. By carefully respecting this tradeoff, ANFIS provides a platform to accurately model nonlinear systems and is also available in a user-friendly graphical user interface within the Matlab programming language. The ANFIS architecture was first proposed by Jang

(1993). ANFIS systems can handle imprecise and uncertain information while also exhibiting robust approximation abilities. ANFIS has been extensively utilized by researchers to build multi step-ahead prediction models (Chang et al., 2015; Jiménez et al., 2009; Wang et al., 2013). Additionally, Najah (2014) has demonstrated the effectiveness of using these new technologies in generating a prediction model for water quality parameters.

An example of an ANFIS system architecture is shown in Figure 2.7. While it appears similar to the ANN architecture, the benefit of ANFIS is the dual input approach. The fuzzy logic component permits the development of membership functions based on known fundamental system factors. For example, an environmental system may behave in different regimes based on some overarching control variable, such as temperature or pH. A fuzzy logic approach can specify to use different calculation approaches based on the fuzzy input for this control variable in the first layer. Alternatively, the ANN component permits the development of actual functional forms and mathematical relationships based on large-scale empirical data. The resultant ANFIS system thus benefits from both the widely-accepted fundamental knowledge and the preponderance of raw data from empirical experiments.

2.4.6 Comparison of ANFIS Controllers

There are two main types of ANFIS controllers, Mamdani and Sugeno, named for the following researchers (Mamdani, 1974; Takagi and Sugeno, 1985). One main difference between the controllers is how the output of the ANFIS system is derived. For a Mamdani type system, the output is assigned a membership function like the input variables. The crisp output is calculated by the center of mass of the resulting overlapped output membership functions. In contrast, the Sugeno output is derived using a simpler weighted average method to produce answers with discrete values. Another difference lies in the how the controllers are trained. Sugeno uses the systematic approach of neural networks to develop the membership functions and rule base. Alternatively, the Mamdani type controller relies on the experience and linguistic knowledge of the user to define the membership functions and rule base.

The best of these two ANFIS architectures for use in control is a debatable topic. Ozger (2009) has presented work on using both Mamdani and Takagi-Sugeno ANFIS controllers to predict stream-flow values. Ozger's research further elaborates on the differences between the two types of controllers. While both controllers are considered universal approximators (can predict any continuous function to any degree or accuracy), the Mamdani controller has a distinct advantage by incorporating the use of numerical and linguistic data. This difference is profound as the ability to use linguistic relationships will allow an expert user of a system to place that knowledge

Figure 2.7 – Layers of ANFIS architecture. Layer 1: input fuzzification; Layers 2 and 3: ANN hidden layers; Layer 4 defuzzification; and Layer 5: crisp output

within a Mamdani fuzzy system. Alternatively, the Takagi-Sugeno systems requires the user to use only numerical representations of variables. The best controller is still a matter of preference and application, as shown by various researchers (Mehta and Jain, 2009; Ozger, 2009; Kaur and Kaur, 2012; Fahmy et al., 2015).

Another comparison of ANFIS architectures was presented by Kaur and Kaur (2012). This body of work presents the results of using both a Mamdani and Sugeno type ANFIS system to control an air conditioning system using equal rule bases. Kaur describes the most basic difference between the Mamdani and Sugeno systems as the way the system determines the discrete output. Mamdani systems use fuzzy output as a means for defuzzification where Sugeno models use a weighted average method as previously discussed. Like other researchers, Kaur concludes that the Mamdani system is more easily interpreted by users since the output is based on linguistic descriptions, unlike the Sugeno architecture; however, the Sugeno architecture is more computationally efficient. In this research, an air conditioning system is modeled using temperature and humidity as input variables while the compressor speed of the air conditioning system is the output variable. Both types of architectures exhibited similar abilities in the control of the air conditioning system; however, the authors determined the Sugeno system was more adaptable, as it may be paired with other optimization techniques, such as genetic algorithms or neural networks to adapt to individual preferences or perturbations in the environment.

2.4.7 Sugeno Type ANFIS Controller

Theory

In 1985, Takagi and Sugeno (1985) presented a mathematical control tool where a fuzzy model was constructed using fuzzy implications and reasoning. This architecture uses fuzzy partitions in the input space. As a result, fuzzy sub-spaces are defined using linear input-output relationships. In other words, this architecture is able to reduce the number of piece-wise linear approximations required to define a nonlinear system. The output of this architecture is derived by the aggregation of values inferred by the fuzzy relationships that were applied to the input. This work differs from that of Mamdani, whereas controllers may be built with little experience or knowledge about the process being controlled. The motivation for this type of control lies counter to the assumption that a human operator can exhibit optimal control of a process, meaning the fuzzy architecture presented can perform better than a Mamdani type controller using linguistic relationships for input and output. Additionally, the work presented by Takagi and Sugeno demonstrates that the proposed method can define membership functions based on the root mean square of the output errors. This result is significant, as this method uses a mechanism that can obtain similar

parameters as that of a linguistic or experience-based fuzzy system. Another advantage of the Takagi-Sugeno method is the ability to identify fuzzy regions when used with data containing noise. This method has shown that when presented with input data which contains noise, the predicted parameters are equivalent to one where no noise is present.

The Sugeno structure of the ANFIS system typically consists of 5 layers, or groups of nodes. The nodes in the first layer generate the membership grade. This function is represented by the first layer in Figure 2.7 and is given by:

$$O_i^1 = \mu_{A_i}(x) = \frac{1}{1 + \left| \frac{x-c_i}{a_i} \right|^{2b_i}} \quad (2.7)$$

where:

- O_i^1 is the node function the first layer;
- μ_{A_i} is the membership function;
- x is the input to node i ;
- A_i is the linguistic label (high, medium, low, etc.) associated with the node;
- $\{a_i, b_i, c_i\}$ is the parameter set that changes the shape of the membership function, also known as premise parameters.

In the second layer of the ANFIS structure, each node is a fixed node where the output is the firing strength of a rule as a product of all the incoming signals as described by:

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y), i = 1, 2 \quad (2.8)$$

where w_i is the node output representative of the firing strength of a rule.

Next, the third layer of the ANFIS structure is also a fixed node that calculates the ratio of the i^{th} rules firing strength to the total of all the rules firing strength given by:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1, 2 \quad (2.9)$$

where $\bar{w}_i$ is the normalized firing strength.

The fourth layer computes the contribution of the i^{th} rule toward the overall output according to:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad (2.10)$$

where:

- $\bar{w}_i$ is the output of Layer 3;
- $\{p_i, q_i, r_i\}$ is the parameter set that is referred to as the consequent parameters.

Finally, the fifth layer is a single node that computes the overall (i.e. crisp or discrete) output that is a summation of contribution from each rule as follows:

$$O_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (2.11)$$

Applications

Sugeno controllers were further developed through research conducted by Jang (1996). This work presents a heuristic way to select input data for ANFIS learning. Jang further expressed the importance of the input selection especially when there are circumstances where there could be hundreds of input variables. While ANFIS is extremely powerful, developing a controller with this many variables may be computationally inefficient. For this reason, Jang introduced a heuristic method to determine the most influential variables. This method consists of the construction of multiple one-variable ANFIS models. These models are then trained for a single pass. The models with the lowest Root Mean Squared Error (RMSE) after one epoch of training then become the most applicable variables to use in training an ANFIS model. The assumption that the models with the lowest RMSE after one pass have the greatest potential for developing a lower RMSE with multiple epochs is not always true; however, Jang states it is heuristically reasonable.

The application of Sugeno controllers to a pH control process has been researched by Zhang and Morris (1999). This investigation used an ANFIS model to control pH in a continuously stirred tank reactor. This research has shown ANFIS controllers are very effective at modeling nonlinear processes having different characteristics in multiple operating regions. Additionally, both process knowledge as well as input-output data are used together to define the membership functions, allowing for a comprehensive

approach to control verses other algorithms. As an example, using only ANN's will not account for process knowledge as the control variable is determined from input-output data alone. This investigation shows the main advantage to using ANFIS to control a nonlinear process is described where multiple simple linear models can be combined to define the operating characteristics of a complex model.

A second example of Sugeno controllers used to implement pH control is described by Navghare et al. (2011). This body of work outlines the development of a Sugeno ANFIS controller for a modeled pH neutralization process in a continuously stirred tank reactor. The basis of this research assumes that if a process can be modeled well, then, in turn, it may be controlled in a complimentary fashion (Zhang and Morris, 1997). While somewhat similar to pH control for AMD, this research does differ in two aspects. First, this work is based solely on a theoretical model verses a bench or pilot scale system. Second, the modeled continuously stirred tank reactor uses a constant flow rate for addition of alkaline material to the reactor where the practical in-field use of this system within industry would most likely see variations in this flow rate. Nevertheless, this research created a controller based on two input variables, flow rate of alkaline material and the pH leaving the reactor, in addition to four rules. The researchers concluded that the ANFIS type controllers are superior to common nonlinear (PID) controllers in manipulating adjustment to the alkaline flow rate of the modeled reactor. Additionally, the researchers demonstrated the increased computational efficiency gained by using an ANFIS approach yields a controller which is even more viable when placed alongside a traditional controller.

Development and Training

To develop a Sugeno ANFIS system, three sets of data are typically required: training, testing, and checking. Optimally, each data set must contain all of the input parameters and known output values. This input and output data should cover a wide range of operating conditions, so the mathematical models can “learn” the intricacies of the input-output models as well as any effects from combining or overlapping regions of variables. The testing and training data are used in the learning process with regard to developing the ANFIS algorithm. The checking data is an independent validation set to check that the model has not over-fitted the testing and training data.

During training, multiple iterations, or epochs, are used to determine the model parameters of the data. In the course of this iterative process, the model parameters are adjusted such that the error between the training data and the model prediction is reduced for each epoch. If testing data is not considered at this stage, over-fitting will occur. Over-fitting results when the model predicts the small measurement errors

inherent to a single data set rather than the overall physical interactions fundamentally driving the model output. For each training epoch, model error for a single data set (i.e. the training data) continually decreases, as the optimizer finds better model parameters. Alternatively, the overall model error (training + testing) will pass through a minimum and begin to increase once again as overfitting occurs. To ensure that the models have the best predictive capacity, the stopping criterion for the algorithm is defined as the point where the total model error (training + testing) is minimized. Development of actual function forms and mathematical relationships is based on large-scale input data. The resultant ANFIS system thus benefits from both the widely accepted fundamental knowledge and the preponderance of raw data from continuous monitoring stations.

2.4.8 Hybrid ANFIS Controllers

More recent research has combined ANFIS models with other machine learning algorithms to improve accuracy, resulting in hybrid ANFIS models. Petchimathan et al. (2014) presented research comparing an ANFIS system with that of a Local Linear Model Tree (LOLIMOT). A LOLIMOT is a type of neural network where a nonlinear function can be modeled using multiple piece-wise linear models. Both methods were used to control the pH in a continuously stirred tank reactor where the addition of acidic material is constant and the addition of alkaline material is the controlled variable. A comparison of the two predictive algorithms shows the LOLIMOT has a slight advantage by requiring a reduced number of parameters to identify the systems; therefore, a reduced training time is needed. In regard to control of the pH system, it was determined that both approaches are valid options for modeling complex nonlinear relationships like pH control.

Recent research by Chang et al. (2015) has shown that static neural network architectures can be used in conjunction with dynamic neural networks, nonlinear autoregressive with exogenous input (NARX), to predict water quality parameters. In Chang's research, data-driven techniques are applied to the time-series data to overcome the scarcity of real-time monitoring data. This research is indicative of how ANFIS controllers have evolved from the first proposed basic control by combining this technology with other predictive algorithms to create more advanced predictive models.

Another example of paring ANFIS with other optimization techniques is presented by Lei et al. (2007). This work couples multiple ANFIS controllers with genetic algorithms to diagnose failures within rolling element bearings using vibration analysis. Vibration analysis has similarities to that of pH control where the analysis of the vibration data requires skilled operators to accurately interpret the data. Likewise, the

process of pH control using a plant requires a skilled operator make pH adjustments based on previous experience. In contrast, an inexperienced operator lacks the ability to control this nonlinear process. Lei has effectively shown the use of genetic algorithms can optimize the weights ANFIS systems use in developing the crisp output. These researchers have shown how combining multiple algorithms is an advantageous method which can be used to ameliorate the disadvantages of using a single ANFIS controller.

2.4.9 Model Predictive Controllers

Another advanced control technique commonly used by industry is model predictive control. MPC has three basic operating regions defined by, output prediction, control calculation, and closing the feedback loop (Andone et al., 2006). This type of control uses a mathematical model which represents the controlled system to predict the future outcomes of the system based on the current state. The future outcomes are based on a finite future time horizon. The predicted outcome is then used to determine the optimal output of the control variable. At each time step, the MPC computes another future prediction, where the control variable may be manipulated to achieve a desired outcome. These types of controllers are often used in chemical and refining process where the control scheme involves large, multivariate processes and where mathematical models governing the system are well known (Stewart et al., 2010). Additionally, MPC may be integrated to include fuzzy logic based modeling methodologies as described by Andone (2006).

There are many advantages associated with MPC. These include: the ability to handle constraints and uncertainties, adaptation with slow moving process with time delays, the ability to handle time-varying system dynamics, and the incorporation of cost functions to achieve multiple objectives (Afram and Janabi-Sharifi, 2014). A disadvantage of MPC is the length of time required to develop an accurate controller (Zhang and Morris, 1999). Likewise, this type of control can be computationally expensive when used to control a complex processing systems.

2.5 Summary

This section has reviewed a multitude of topics relating to pH treatment control schemes and mine water discharges in the CAPP region. Recently, increased regulatory reviews have led to the lowering of effluent limits defined in the NPDES permitting process. Lower effluent limits have impaired the ability of coal operators to discharge compliant water from mine permits. These exceedances from permitted limits have resulted in regulatory action through consent decrees, which established civil penalties and increased fines for further excursions. The increased cost associ-

ated with discharging noncompliant water has elevated the expenses associated with treating and maintaining NPDES outfalls. Additionally, CAPP mine operators have been slow to adopt new technologies prior to the expansion of water quality enforcement. Past practices used by CAPP mines have traditionally been labor intensive with intermittent monitoring and control at treatment sites.

Many researchers have studied AMD over the last 40 years. This research has resulted in numerous methods to predict, prevent and treat AMD. Two major groups of methods used to treat AMD include passive and active treatment systems. Passive treatment systems use man-made structures resembling wetlands to increase the quality of mine influenced waters. This approach is typically more cost effective; however, limitations often prohibit the use of passive treatment. Alternatively, active treatment uses the addition of chemicals to improve water quality, but this practice is expensive and often lacks consistent control of treatment chemicals. Additionally, multiple constraining factors prohibit a systematic approach to water treatment that yields flawless compliance including, topography, limited access to utilities, remote locations, dissimilar drainage areas, and diverse water chemistry properties. These factors indicate an opportunity exists to use advanced machine learning algorithms to provide a reliable alternative to the manual control strategies currently in use.

Several technologies have been developed by other industries to manipulate nonlinear processes, including pH control. Regrettably, these forms of automation remain unproven in applications similar to an AMD treatment site, even though real-time control of AMD treatment has many environmental and economic benefits. The ability to implement a robust, real-time controller for chemical treatment of AMD has many benefits.

One advanced technology that shows promise in the control of mine water treatment systems is the combination of neural networks and fuzzy logic. The ANFIS system provides a powerful mechanism for model-based prediction, particularly for nonlinear systems. The control of pH for MIW treatment is a complex endeavor complicated by the nonlinear pH titration curve and uni-directional control. Since traditional (e.g. PID) controllers are unsuitable for this application, advanced controllers have been considered. Prior studies have shown that fuzzy logic, neural networks, and ANFIS system can be used in pH control applications; however, many of these studies were limited to modeling exercises and industrial processes. As a result, a significant research opportunity exists to assess the feasibility of these control strategies for the environmental applications.

CHAPTER 3

DESIGN AND CHARACTERIZATION OF BENCH-SCALE SYSTEM

3.1 System Design

In order to evaluate the feasibility of various control methodologies, a bench-scale mine water treatment system was first designed and constructed. To complete this objective, treatment reactors, as well as the hardware and software components needed to construct an operable system, were identified and procured. The assembled system has the ability to manipulate multiple input variables and simulate various environmental conditions. For example, the system can be manipulated to individually control the flow rate of water into the system and the flow rate of chemical treatment. In addition to the manipulated parameters, water quality sensors were placed at various positions along the flow path of this water measure to record the quality of water entering the system, after treatment, and leaving the overall system.

The bench-scale system is controlled by a personal computer that records flow rate data and measurements of water quality. This data is then used by a control algorithm on the computer to manipulate the amount of alkaline material dispensed into the treatment reactor to bring the pH of the water to a desired level. The following sections describe in detail the components used in creating this laboratory AMD treatment simulator.

3.1.1 Reactors

The bench-scale system, as seen in Figure 3.1, consists of five 5-gallon buckets (reactors) placed in series. The total working system volume is approximately 18.5 gallons, and the precise volume of each bucket is shown in Table 3.1. Each bucket decreases in elevation from the system inlet to the outlet to facilitate gravity flow, similar to AMD treatment systems used in CAPP. This design was accomplished by placing the individual reactors on concrete pavers, with the number of pavers decreasing in number as the reactors approach the outlet of the entire system.

Figure 3.1 – Bench-scale system built to simulate a mine water treatment structure.

Table 3.1 – Measured volumes of reactors.

Reactor	Volume (gal)
1	3.81
2	3.71
3	3.70
4	3.69
5	3.67
TOTAL	18.58

In order to allow flow of water between the reactors, bulkhead fittings were installed on the side of each reactor as shown in Figure 3.2. The $\frac{3}{4}$ -inch by $\frac{3}{4}$ -inch bulkhead fittings were mounted under the upper lip of each bucket. The outline of the bulkhead fitting was traced on the bucket and a hole was drilled to accommodate the fitting. Once the hole was drilled, the coupling was installed. The bulkhead coupling provided a location to plumb $\frac{3}{4}$ -inch PVC pipe to connect the series of buckets. Additionally, the use of PVC “T” couplings installed on the pipe between the buckets provided a convenient mounting location for the flow through water sensors via $\frac{1}{2}$ -inch threaded fittings inserted into the bottom of the “T” section. At the inlet to the system, pH and conductivity sensors acquire incoming water quality data, as water flows through the $\frac{3}{4}$ -inch PVC pipe. The pipe at the exit point of the first bucket and the outflow of the last bucket also contain sensors to monitor water quality. The sensors at the outflow are representative of a permitted mine water discharge. Figure 3.3 is a process flow diagram showing the general flow of material and information.

Flow of liquids into the system is handled by two peristaltic pump drives coupled with a total of three pump heads. A Masterflex I/P class pump drive unit (feed pump) with two Masterflex 77602-00 pump heads control the flow of acidic water into the system. This pump drive may be controlled through either internal or external control. External control, via personal computer, is necessary to introduce variability of water flowing into the treatment system via a random signal generator. A Masterflex L/S class pump drive (treatment pump) with a Masterflex 77200-50 pump head delivers an alkaline slurry to the treatment system. Like the feed pump, the treatment pump can be controlled externally. This feature allows the rate of alkaline chemical to be controlled by voltage supplied from a personal computer, thereby providing a desired dosage rate to treat the simulated AMD.

The largest pump, the Masterflex I/P series, has two pump heads supplying water from two independent tanks to the first bucket in series. Two supply tanks are used to accommodate various mixtures of “clean” and “contaminated” water. This configuration provides the flexibility needed to simulate water inflows from different sources

Figure 3.2 – Installation of bulkhead fittings.

Figure 3.3 – Process flow diagram of bench-scale system.

or different mining locations with varying water quality conditions. Connecting the tanks to the pump head and treatment system is a combination of reinforced nylon hose and PVC fittings. The water then flows through the remaining four buckets before the water discharges from the last bucket in series, similar to settling ponds used in the CAPP coal mining industry. This discharge is representative of a NPDES outlet as seen in the mining industry. After leaving the last bucket, water is collected in a drain, which is representative of a stream receiving mine discharge.

Water quality data is acquired at the inlet of the system, the outlet of the first bucket and the outlet of the last bucket at a frequency of once per second. The smaller chemical treatment pump is used to transport chemical from a supply tank to the first bucket in series via a nylon hose, in a similar fashion to the feed pump.

To simulate the use of pond curtains and test different flow regimes, adjustable acrylic plastic baffles were designed and implemented. These baffles inhibit the flow path to prevent water short-circuiting by dividing each bucket into two compartments joined by the open area at the bottom of the bucket. Adjustments to the baffles consist of raising or lowering the divider to three available elevations. The adjustments are 0.125 inch, 1 inch, and 2 inches in elevation from the bottom of the bucket. An example of one of the baffles is shown in Figure 3.4.

Figure 3.4 – Baffle located in reactors to inhibit the flow path.

3.1.2 Sensor Package

Initially, both bench-scale and industrial-scale sensor suites were considered for use in the project. Due to multiple constraints, the industrial-scale system was ultimately deemed inappropriate for this phase of the project, since the current efforts primarily involve the development of software algorithms. Alternatively, the bench-scale system has the ability to provide the necessary data to achieve the project objectives, while maintaining budgetary limits.

Some constraining factors associated with the use of an industrial hardware sensor system in a laboratory environment include cost, size, and placement. The cost of an industrial system is approximately twice of that of a laboratory-based system. The more robust design of the sensor components needed to withstand use in the field influences the difference in cost between the two systems. Two examples of the more robust sensors designed for long-term deployments are the Ott-Hydromet MS5 and YSI EXO1. An industrial grade sensor system is also much larger and requires a large volume of water (~ 50 gallons per reactor) to function correctly. This large volume of water was not logistically possible to accommodate in the laboratory space available. Nevertheless, the specifications for such a system were continually considered throughout the project period, as such a system may be implemented in future field studies. The current algorithm development work prioritizes adaptability and scalability of the software systems, so that the laboratory-developed algorithms

can be easily implemented into an industrial-scale sensor suite.

A multitude of laboratory-scale water quality sensors are commercially available for use; however, for the purpose of this research, pH and conductivity were selected as the primary parameters to measure and utilize for control variables. The primary effluent parameter, pH, was evaluated due to its importance in mine water treatment, as previously discussed in §2.3. Acidic mine water often contains metal hydroxides in solution, and simple pH adjustments can be used to precipitate these hydroxides from solution. Additionally, electrical conductivity measures the amount of dissolved solids with ionic bonds in water. Conductivity is useful as a measure of overall change in water quality from inlet to outlet. For this research, conductivity plays an important role in characterizing the system through residence time distribution studies. Together, these two parameters play a critical role in water quality for Appalachian coal mines, and as a result, provide suitable parameters needed to validate the initial system design.

The pH sensors used in this research are Hanna HI1001 pH monitoring electrodes paired with Hanna HI98413 transmitters. Likewise, the Hanna HI3001 conductivity monitors are utilized to read conductivity in the treatment system. The voltage and current output from a pH sensor is extremely low, therefore, the signal must be converted to a usable format. The pH transmitter converts this low output signal to a 4-20 mA signal. Additionally, the transmitter provides an interface to calibrate the pH sensors making it an essential part of the pH monitoring system.

The transmitters output a 4-20 mA signal which is commonly used in commercial instrumentation. This signal is then coupled to a 250 ohm resistor to change the signal to a voltage reference. The analog voltage signal is then converted to a digital output by the Data Acquisition Unit (DAQ) which communicates with a personal computer. Future work on this project up to and including full-scale deployment, may require the addition of other sensors to measure different water quality parameters to increase the effectiveness or scope of the automation process. Possible parameters to consider in a field implementation include but are not limited to; rain gauges, turbidity, total dissolved solids and water level.

A Kobold Magneto-Inductive flow meter is also included in this system to measure flow rates. The flow meter is a model MIK-5NAU5P L343 with a flow range from 0.2 to 4.0 gallons per minute (GPM). Unfortunately, the use of peristaltic pumps creates a pulsation of the fluid at high flow rates. This pulsation creates erratic noise in the data acquired by the flow meter. For this reason, the voltage supplied by the pump was correlated to flow rate and used as the primary measure of flow of water into the bench-scale system. Future work will include the installation of a dampening system to mitigate the electrical noise created by the peristaltic pump.

3.1.3 Data Acquisition

To collect the voltage data supplied by the sensors and transmit the signals to a personal computer, additional hardware devices are required to send and receive the sensor generated analog signals. A Measurement Computing USB-1608G DAQ is used to convert analog data from the transmitters and pumps to a personal computer. To control the chemical and AMD feed pumps, a Measurement Computing USB-3102 analog output device sends control voltages from the personal computer to the Masterflex pumps.

Each sensor was wired to a transmitter, which converted the electronic signals from the sensor to a measurable voltage. The sensor/transmitter combination correlates the measured voltage to a reference effluent parameter through the calibration process. Voltages output from the transmitters are wired to two input channels of the DAQ, which is termed a “differential measurement.” Next, personal computer records digital signals via a USB cable, sent from the DAQ. Control of the pumps is managed by the software component. A USB connection from the personal computer to the analog output unit transfers digital control signals for conversion to analog signals. Output channels on the analog output device are connected to each pump drive.

3.1.4 Grounding with Floating Signal Source

During the initial stages of testing, signals received from the hardware sensors showed a large amount of variation, identified as “noise,” as shown on the top of Figure 3.5. System inspections show that this noise was caused by the absence of grounding in the sensor package. The sensor suite uses an ungrounded or floating signal source, meaning the voltage signal is not referenced to a system ground. To compensate for this widely varying data, 47 k Ω bias resistors were installed on the positive and negative leads at the connection terminals of the DAQ. These resistors were then connected to the ground reference of the DAQ unit. The installation of the bias resistors greatly reduced the amount of noise in the measurements as seen in the comparison of two data sets in Figure 3.5.

3.1.5 Software

Data from the real-time measurements was collected and analyzed using the Matlab R2015b programming language. Additionally, Simulink, a block diagram environment for simulation and model based design was used to control the treatment system and participate as an interface between the pumps, sensors, and personal computer. For the data collection, a simple Simulink model was constructed to receive data from the DAQ unit. The pH and conductivity of the inlet, outlet of first reactor, and outlet of

Figure 3.5 – Demonstration of reduction in noise from sensor data using bias resistors.

the system were measured at a frequency of one second. This data was then stored in within a Matlab file for further analysis.

3.2 Characterization of Bench Scale System

This section describes the characterization of the bench-scale system. Multiple calibrations were performed to ensure the data provided by the pumps and sensors used in the laboratory scale treatment plant are accurate and reliable. Additionally, residence time distribution studies were conducted to quantify the flow interactions using different pumping rates and baffle positions. This information is useful to scale the chemical reactions from a bench-scale treatment system to large pond-based systems used in the CAPP region.

3.2.1 Calibration of Sensors

Calibration of pH electrodes consists of using two known solutions. The zero point is set at a pH of 7.0 while the slope is attained by using a known solution of high or low pH depending on the expected alkalinity of the media to be measured. For this research, the first electrode in the series was calibrated at a low pH of 4.0, since the AMD entering the system consists of low pH water. The second and third electrodes in series were referenced to a higher pH of 10.0 as water leaving the first reactor and system outlet was inclined to be in the neutral to alkaline range. These different solutions used in calibration ensure accuracy, as indicated by the manufacturer of the components. All calibrations were performed per the manufacturer’s instructions. Additionally, a linear correlation was established to convert the voltage from the transmitter to a pH value. Since the sensor was calibrated using only two points, this correlation results in a linear equation as follows:

$$pH = 3.50x - 3.50 \quad (3.1)$$

where x is the voltage received by the DAQ unit.

Calibration of sensors measuring electrical conductivity followed a similar method to that of pH sensors. The principle of measuring electrical conductivity (EC) is simply measuring the current flowing through the media between two poles to which voltage is applied. Solutions used for calibrating the EC sensors are 84 μ S, 5,000 μ S and 111,800 μ S. During calibration, individual sensors were immersed in the reference solution and voltage measurements were manipulated to match the manufacturer’s recommendations. A second order polynomial regression correlated the voltage to a value representing the conductivity of the solution. Equation 3.2 was implemented

within the Simulink environment to convert from voltage received at the DAQ unit to conductivity as follows:

$$EC = 1359.9x^2 + 836.29x - 2112.2 \quad (3.2)$$

where EC is conductivity measured in micro-Siemens and x is the voltage received by the DAQ unit.

3.2.2 Calibration of Pumps

To identify the respective volumetric flow rates and correlate each pumps input voltage (speed) to a flow rate, pump flows were measured at incremental intervals. The pumps were set at intervals from 10 to 100% and the water pumped was collected over a two minute period. This quantity of water pumped was measured using a graduated cylinder and a flow rate in gallons per minute was derived from the acquired data. Table 3.2 shows the average pump rates for the feed and chemical treatment pumps. This data was used to create a linear regression which correlates the voltage supplied to the pump with the flow rate.

Table 3.2 – Average pump rates for the feed and treatment pumps used in the bench-scale system.

Feed Pump			Treatment Pump		
Setting	GPM	Voltage	RPM	GPM	Voltage
10	0.22	1.39	10	0.0053	0.27
20	0.63	1.78	25	0.0159	0.80
30	0.91	2.18	50	0.0288	1.70
40	1.24	2.58	75	0.0403	2.53
50	1.59	2.98	100	0.0528	3.35
60	2.21	3.38	125	0.0660	4.20
70	2.83	3.77	150	0.0778	5.00
80	3.04	4.17	200	0.1290	6.66
90	3.42	4.57	250	0.1392	8.40
100	3.50	4.97	300	0.1571	10.00

For the feed pump, the linear regression has the form:

$$FlowRate_{FP} = 0.9912x - 1.1898 \quad (3.3)$$

where $FlowRate_{FP}$ is expressed in units of GPM and x is the voltage applied to the pump. This linear regression has a goodness of fit measured by a R^2 value of 0.9853. Additionally, the treatment pump uses a similar method to determine the flow rate of chemical slurry administered to the bench scale system as follows:

$$FlowRate_{TP} = 0.0154x + 0.0011 \quad (3.4)$$

where $FlowRate_{TP}$ is expressed in units of GPM and x if the voltage applied to the pump. Again, the correlation showed a high goodness of fit with a R^2 value of 0.9997.

3.2.3 RTD Testing - Theory

The bench-scale treatment system was characterized through the use of chemical reactor engineering. This sub-set of chemical engineering provides a basis to quantify and compare processes involving the chemical kinetics and reactor (pond) behavior of a treatment system. In an ideal chemical reactor or environmental water treatment system, the entire volume of a reactor is utilized. However, real systems are often limited by various inefficiencies such as bypassing flow and reactor dead space. Both mechanisms effectively reduce the mean residence time and may lead to poor reaction efficiencies. Depending on the reactor geometry and the nature of the flow regime, these problems may be more or less pronounced. A residence time distribution (RTD) study is one way to quantify these issues and is often used to either diagnose problems of an existing reactor or predict effluent concentrations when a new reactor is used (Fogler, 2005).

For this new bench-scale water treatment system, a RTD test was used to quantify the extent of mixing in various geometric configurations. By quantifying the RTD of the bench-scale system, comparisons and scale ratios can be used to show how well the laboratory setup matches the true residence time distribution in industrial systems. Furthermore, RTD tests can verify that different geometric arrangements are truly altering the mixing conditions as anticipated.

For the bench-scale reactor used in this research, pulse RTD tests were performed rather than step RTD tests. In a pulse test, a small quantity of a non-reactive tracer is introduced to the stream of effluent entering the reactor. The tracer concentration at the outlet of the reactor is then measured and used to determine the mean residence time (t_m) as well as the distribution of all possible residence times. These results can be normalized to produce the derived $E(t)$ curve which is given by Equation 3.5:

$$E(t) = \frac{C(t)}{\int_0^\infty C(t)} dt \quad (3.5)$$

where C is the concentration of tracer measured at the discharge at time, t . Examples of arbitrary $C(t)$ and $E(t)$ curves are shown in Figure 3.6. From this function, t_m was calculated by:

$$t_m = \int_0^{\infty} tE(t)dt \quad (3.6)$$

The mean hydraulic retention time (τ) was estimated by dividing the volume of the reactor by the volumetric feed flow. In an ideal reactor, this estimate is equal to t_m found from the RTD studies.

The main drawback to using the step test was the difficulty of injecting the indicator material into the reactor inlet. Conditions that must be met to use this test include: obtaining a reasonable entrance point to inject the tracer material, the injection must take place over a short time period in relation to the residence time and the amount of dispersion between the injection point and reactor inlet must be negligible (Fogler, 2005).

Figure 3.6 – Typical $C(t)$ and $E(t)$ curves after (Fogler, 2005).

3.2.4 RTD Test - Procedure

A pulse experiment was performed on the bench-scale pond reactor system to identify the RTD at varying flow rates and curtain settings. These tests are significant, as they can provide a basis for scaling the chemical reactions to industrial scale ponds. The indicator used for this experiment was a brine solution with a concentration of 1.06 ounces of NaCl per gallon of water. At this concentration, the solution was fully saturated with NaCl. Conductivity sensors indicated the concentration of indicator solution entering and leaving the reactor.

During the tests, all of the buckets were first filled with ordinary tap water and the baffles were oriented at the desired elevation. The measured conductivity of the tap water was recorded for each test to establish an initial reading and averaged to 200 micro-Siemens. The feed pump was set to a desired flow rate, and the system was initiated and run until a flow equilibrium was obtained. The data acquisition script was started at this time, and subsequently, a syringe was used to inject 25ml

of the brine solution into the feed pipe for a duration of less than one second. The system continued pumping until the outlet conductivity was once again equal to the beginning conductivity measured. At this point, the test was stopped, and the data was saved.

Nine total iterations of this test were performed to evaluate the reactor under different conditions. Iterations involved a full factorial combination of three pump flow rates (3.0 gpm, 1.8 gpm, and 0.8 gpm) and three baffle settings (2 inch, 0125 inch bottom clearance, and no baffle), resulting in 9 total test runs. After all tests were complete, data was analyzed to determine the mean residence time, and other meaningful characterization variables of this laboratory equipment.

3.2.5 RTD Test - Results and Discussion

Figures 3.7 - 3.9 show the conductivity vs. time data collected from the experimental pulse tests. Figure 3.7 shows the raw data collected during processing. This data was first adjusted by subtracting the background conductivity of tap water and shifting the time axis to begin at the pulse injection, thus producing the normalized $C(t)$ curve shown in Figure 3.8. Lastly, this data was manipulated using Equation 3.5 to produce the $E(t)$ curve in Figure 3.9.

Figure 3.7 – Raw data collected during pulse experiment.

As shown in Figure 3.9, higher flow rates tend to induce a sharper peak in the $E(t)$ curves, while lower baffle settings tend to induce longer tails. Additionally, when no

Figure 3.8 – Normalized $C(t)$ curve for pulse test conducted on bench-scale treatment system.

baffle is present the tails become even larger due to the more randomized flow pattern. For example, the no-baffle setting, low flow rate reaches a maximum $E(t)$ peak at 1.25 minutes, compared to a peak of 2.00 minutes for the high baffle setting at the same flow rate. This qualitative difference confirms that the baffle settings are changing the flow regimes in the reactor. The high peak seen in the low baffle setting is often associated with plug flow reactors, confirming that the low baffle setting induces plug flow behavior. Additionally, it can be seen that the flow rate is influencing the reactor behavior by changing the degree of mixing within the vessel. At high flow rates, the increased rate of water entering the system facilitates quicker mixing than lower flow rates. For example, the high baffle setting, high flow rate $E(t)$ curve in Figure 3.9 is similar to the ideal mixed reactor curve shown in Figure 3.10. Furthermore, the high baffle setting, low flow rate $E(t)$ curve in Figure 3.9 is complementary to the lower left reactor curve in Figure 3.10; which is representative of a late RTD curve where the observed mean residence time, t_m , occurs later than the ideal residence time, τ . The percent difference in ideal verses observed residence times shown in Table 3.3, for Test No. 31, confirms this observation.

Accompanying the change in residence time is a change in the shape of the $E(t)$ curve for varying flow rates. This observation is substantial since a large change in flow rate indicates a change in the distribution function of the reactor. The mean residence times for the system at each setting are listed in Table 3.3. Table 3.3 also shows the ideal residence time, τ , as well as the difference between the two expressed as a percent

Figure 3.9 – $E(t)$ curve for pulse test conducted on bench-scale treatment system.

of t_m . This table shows that substantial deviations between the measured and ideal retention time are apparent for specific flow conditions, particularly low flow and low baffle positions. A comparison of the mean versus ideal residence times shows t_m of the bench-scale reactor is typically greater than τ of an ideal reactor. Two factors can explain a difference in times when comparing ideal versus non-ideal reactors. First, bypassing can occur in a reactor when all or part of the feed stream travels directly to the outlet. An idealized example of bypassing is shown in the upper right quadrant of Figure 3.10. Bypassing indicates that the total volumetric flow seen by the reactor is less than the volumetric flow introduced to the reactor, increasing the residence time. Second, time lag occurs in several of the experimental E curves as seen in the representative lower right curve in Figure 3.10. Time lag takes place when plug flow is in series with mixed flow.

These experiments were conducted in the laboratory under relatively controlled conditions. Despite this environment, the RTD and predicted mean residence time deviated, sometimes substantially, from the ideal behavior (e.g. tests 30, 34, and 38). Even small changes to reactor conditions (flow and baffle settings) result in large changes in the flow regime and residence time distribution. While these deviations were easy to measure and analyze in the laboratory, similar disturbances may be difficult or impossible to consider in the field. Full RTD tests are rarely performed for NPDES treatment ponds, and they certainly cannot be performed for all flow and baffle conditions. Since small changes in these variables produced measurable differ-

Table 3.3 – Comparison of residence times by experiment.

Test No.	Feed Flow Rate (gpm)	Mean Residence Time, t_m (min.)	Ideal Residence Time, τ (min.)	Difference (% of t_m)
High Baffle Setting				
37	3.02	1.25	1.26	-0.91%
34	1.79	2.31	2.13	8.27%
31	0.79	5.12	4.82	5.82%
Low Baffle Setting				
36	3.03	1.31	1.26	4.03%
33	1.73	2.19	2.20	-0.55%
30	0.81	5.32	4.70	11.60%
No Baffle				
38	2.87	1.48	1.33	10.32%
35	1.83	2.06	2.08	-1.05%
32	0.81	4.88	4.70	3.63%

ences in flow regime in the laboratory scale, even larger deviations are anticipated in the full scale where non-ideal geometry and unsteady-state flow are unavoidable. This result suggests that, a generic flow regime characterization is practically unobtainable in the field, thus negating the benefits of any traditional model-based control algorithm. Alternatively, this type of inconsistency is ideally suited for a machine learning application where the observation of patterns in the data is used to develop an algorithm verses the assumption of a steady state.

Figure 3.10 – Ideal mixed reactor $E(t)$ curve versus imperfect reactor $E(t)$ curves, after Fogler (2005).

CHAPTER 4

EXPERIMENTAL METHODS

4.1 Goals of Experimental Program

An experimental program was devised to test two facets used in the development of an ANFIS controller. The first phase of this experimental program tested the ability of a Mamdani-style ANFIS controller to maintain a given pH set-point under varying environmental conditions. In this program, disturbances were introduced into the treatment system and the ANFIS controller dictated the flow rate of alkaline material required to keep the pH at a desired level and the results were measured. Figure 4.1 shows a generalized form of the Mamdani control scheme and system disturbances. The controlled response to these changes was monitored to verify that the pH of the desired set-point was maintained within an acceptable level of deviation. The acceptable tolerance for deviation is considered to be plus or minus 0.5 pH units. This range was chosen based on the broad technical based limits used for NPDES outlet pH parameters as shown in Table 2.1. While typical NPDES permits allow a much broader pH range (6 to 9), the tighter specification was used for this study, since the laboratory conditions are considered much more favorable than the actual treatment system. Stated differently, a controller that cannot meet the tighter constraints in the lab will not be suitable for field implementation.

During the second experimental program, numerous sets of training data were used to develop a Sugeno type ANFIS model that can predict the outgoing pH based on current environmental and treatment conditions. Over 20 training data sets were collected using the random perturbation methods described by Zhang and Morris (1997). Each training data set was then used to build an ANFIS model, and each ANFIS model was then used to predict the pH of treated effluent of independent checking data sets at multiple time steps in the future. Since each ANFIS model was built from a distinct data set, the outcome of this test matrix will identify what qualities a successful training data set must possess to accommodate accurate pH prediction. Results from these tests when combined with the RTD behavior can be

Figure 4.1 – Simplified representation of Mamdani controller.

used to design pilot or industrial scale systems. This robust predictive model will be invaluable for ANFIS and MPC designs.

4.2 ANFIS Control Tests

For the Mamdani control campaign, a total of seven tests were conducted to record the responses of the ANFIS controller to various disturbances, including the flow regime, incoming pH, and pH set-point. Table 4.1 lists each test and the corresponding disturbance used to investigate the system response. Furthermore, the disturbances encountered in each test are correlated to a similar disturbance that could be encountered in an industrial setting. The initial test was conducted under steady-state conditions to develop a baseline response for the controller. Here the flow, pH, set-point and baffle positions were held constant for the duration of the test.

Next, the flow rate of simulated AMD introduced to the system was randomly varied. The consecutive tests employed a combination of changes in system variables (as shown in Table 4.1) to confirm the controller is able to maintain a desired set-point when multiple perturbations are encountered by the treatment system.

4.2.1 Test Procedures

Prior to testing, a Simulink model was created to both collect data and implement an ANFIS controller for the tests. Figure 4.2 shows the block diagram used during this phase of testing. In this figure, all input blocks are magenta. Additionally, green blocks are used to generate a random frequency and magnitude for voltage of the feed

Table 4.1 – Summary of different tests ran using ANFIS controller.

Test No.	Description of Test	Field Condition Simulated
1	Tested ANFIS Control under steady-state conditions, all variables were held constant as the controller tried to achieve a set-point of 7	Normal field operations under steady state conditions
2	Flow rate to the system was randomly varied while all other variables remained constant	Normal field conditions where flow varies over time
3	Flow rate to the system was randomly varied while the set-point was changed from 6 to 8, and then back to 6	Simulates condition where a higher pH is required and the controllers ability to respond to a change in the set-point; such an increasing could model increasing the pH to reduce manganese
4	Flow rate was doubled without ANFIS controller seeing a change in the pump voltage	Simulates condition where flow measurement is not available as large change in flow rate occurs, eg. large rain event that damaged flow recording device
5	Changed pH of feed stream, reduced from 3.4 to 2.6	Simulates condition of changing pH, change in feed water may lead to increased/decreased pH seen by treatment system
6	Removed baffles in reactors for a period of 500 seconds.	Simulates condition where a pond baffle fails due to environmental circumstance or vandalism
7	Combination of perturbations to variables in the system	More accurately simulates conditions where multiple variables change at once, for example, a large intermittent rain event, along with a change to feed pH

Figure 4.2 – Simulink block diagram used to evaluate the ANFIS controller.

pump. Next, the blue blocks represent the analog output signals sent to the feed and treatment pumps. For tests using constant feed flow rates, a single step function was generated by the blue blocks versus the random voltage used by the green blocks for variable flow rates. The black blocks represent the output data from the treatment system which is saved to a Matlab file. Variables which are saved include the pH of all of the sensors, voltages from the feed and treatment pumps, voltage from the flow sensor, and the desired set-point. Mechanistically, the system uses voltage as the control signal; however, this signal can be correlated to the flow rates supplied by the pumps using the calibration shown in §3.2.2. This method was used to report all results in GPM versus volts. Finally, the red blocks are used for the ANFIS controller. These variables are recorded at a frequency of once per second. The blocks preceding the controllers change the pH values to the *Error* and *Change in Error* as described in further detail below.

To conduct a test, the feed water reservoir was first filled with neutral tap water, and the pH was then lowered to pH 3.0 using hydrochloric acid (HCl). This value was confirmed using a handheld pH meter. Next, the alkaline slurry was prepared using five gallons of water, mixed with 5.29 ounces of Sodium Carbonate (Na_2CO_3). Then, the sensors were installed into the preexisting locations on the bench-scale treatment system. Finally, power was supplied to the transmitters and the Simulink model was started to initiate the controller and record the response of the treatment system

sensors. After a test was completed, the Simulink model autonomously archived the data set for future analysis. This data set was then evaluated to understand the performance of the ANFIS controller.

During testing, the status of the treatment system was monitored in real-time. Preliminary tests were monitored and aborted once the controller exhibited undesired characteristics. For example, if a controller could not maintain the desired set-point, the test was stopped and a new controller was designed with adjustments to the membership function and rules. Several of these controllers were able to maintain a constant set-point; however, when a change in set-point was applied the controller became unstable. Likewise, many of the preliminary controllers could maintain a constant pH set-point under steady flow conditions; nevertheless, they also became unstable when varying feed flow rates were tested in the system. A total of 36 controllers were built and tested using this trial and error procedure before a reliable controller was produced which exhibited robust control.

Once a robust controller was attained, multiple iterations were performed to determine the most effective time delay for use in the *Change in Error* calculation. The delay time was increased from 1 to 10 seconds in 1 second intervals. The controller was then tested under steady state conditions. The 5 second delay, or approximately 0.03% of a bucket t_m , used in this research was determined to be the best time delay based on a qualitative evaluation of the controller performance.

4.2.2 ANFIS Controller Membership Functions and Rules

The ANFIS controller was developed in the Matlab programming environment using the Mamdani architecture. To accomplish this development, rules and membership functions were developed for three input variables and one output variable. The input variables used for control were *Error*, *Change in Error*, and the *Feed Pump Voltage* (flow rate), while the output variable is the *Change in Voltage* (flow rate) which is sent to the treatment pump. Error was defined as follows:

$$Error = SetPoint - pH_2 \quad (4.1)$$

where *SetPoint* was the desired pH leaving the first reactor and pH_2 was the actual pH leaving the first reactor. Furthermore, the *Change in Error* was the difference between the current error and the error five time steps lagging from the current time as follows:

$$\Delta Error = Error_{t=0} - Error_{t=-5} \quad (4.2)$$

Finally, the output variable was the magnitude of the *Change in Voltage* ($\pm$) the controller will send to the treatment pump. The choice to use change in voltage allows the pump to respond to dynamic conditions. Initially, controllers were designed that used the voltage applied to the treatment pump as the output variable. These controllers exhibited unstable characteristics which led to the decision to allow the pump to work over a dynamic range and only apply an increase or decrease in voltage to control the flow rate of treatment chemical.

The choice of these variables directly contributes to the robustness of the system as the *Error* indicates the magnitude and direction of the difference between the set-point and current pH. The *Change in Error* is significant because it indicates the velocity at which the system is approaching the set-point. Finally, the feed pump voltage indicates how fast the overall system will react to a change in the distribution of treatment chemical. This value was chosen as a variable from the results of the RTD test which indicated that larger flow rates increase the degree of mixing within the vessels.

While the scientific literature does not provide direct guidelines for developing membership functions for AMD pH control, initial rules and relationships were modeled after many researchers who have built Mamdani type controllers for nonlinear systems (Ghee et al., 2002; Ibrahim, 2010; Singh et al., 2011; Kaur and Kaur, 2012). Next, the relationships were modified using the overall boundary conditions, desired control response, and knowledge of the characteristics of the bench-scale system. This method lacks a direct analytical approach; however, it is relevant given the ANFIS architecture used. In fact, one characteristic of the Mamdani controller when compared to other traditional types of control is the explicit knowledge of the system held by the user directly contributes to the accuracy of the controller. Finally, multiple iterations of testing were required to tune the parameters of the membership functions used in the tested controller.

During testing, a maximum voltage limit was prescribed to the output of the treatment pump to account for the time lag between alkaline slurry injection and the reading of the pH sensor leaving the first reactor. Without this constraint, the treatment pump would dispense an excessive amount of treatment resulting in the pH of the treatment system elevating to a range above 10. This large swing in pH values resulted in an unacceptable oscillation of the pH in the upper region of the pH scale. Furthermore, the pH leaving the system did not stabilize at a the desired set-point and instead remained above the desired pH level.

Figure 4.3 shows the memberships functions used for the final ANFIS controller. Triangular membership functions were used for the majority of the variable parameters to define the degree of membership. Additionally, S and Z shaped membership func-

Figure 4.3 – Membership functions used to develop ANFIS controller.

tions were used to define the outer extents of the variables *Error*, *Change in Error*, and *Change in Voltage* (flow rate) to treatment pump. The S and Z shaped membership functions allow the upper and lower extremities of these variables to be inclusive of extreme values which may lay outside the normal operating parameters. Furthermore, the inclusivity of these membership functions allows for a low number of overall membership functions, making the controller more computationally efficient.

For *Error*, *Change in Error*, and *Change in Voltage* (flow rate), a total of seven membership functions were used. These membership functions range in scale from Negative High (NH), Negative Medium (NM), Negative Low (NL), Zero (Z), Positive Low (PL), Positive Medium (PM), and Positive High (PH). The remaining variable, *Feed Pump Voltage*, is defined by three membership functions, Low (L), Medium (M), and High (H). These relationships are shown in Figure 4.3.

Rules were also defined based on the variables relationships to their respective membership functions. One thing to note on the rules relating to *Error* and *Change in Error* is the unbalanced distribution of NH values. This designation was necessary due to the characteristic of the treatment system only being able to control the upward movement of the pH. Should both basic and acidic chemicals be used in pH control this offset would not be necessary; however, this dual control is not expected

Table 4.2 – Rule matrices for variables used in development of ANFIS controller.

		<i>Change in Error</i>						
		NH	NM	NL	Z	PL	PM	PH
<i>Error</i>	NH	NH	NH	NH	NH	NH	NM	NL
	NM	NH	NH	NH	NH	NM	NL	Z
	NL	NH	NH	NM	NM	NL	Z	PL
	Z	NH	NH	NM	NL	Z	PL	PM
	PL	NH	NM	NL	Z	PL	PM	PM
	PM	NM	NL	Z	PL	PM	PM	PH
	PH	NL	Z	PL	PM	PM	PH	PH
N		=	Negative		H	=	High	
P		=	Positive		M	=	Medium	
Z		=	Zero		L	=	Low	

		<i>Feed Pump Voltage</i>		
		L	M	H
<i>Error</i>	NH	NM	NH	NH
	NM	NL	NM	NH
	NL	NL	NL	NM
	Z	Z	Z	Z
	PL	PL	PL	PM
	PM	PL	PM	PH
	PH	PM	PH	PH

for treatment sites in CAPP.

Table 4.2 shows the rules used for the tests of the ANFIS controller. A total of 70 rules were created to guide the controller to perform the desired response to changes in the system variables. The tables may be read as: if *Error* is NH and *Change in Error* is PH then, *Change in Voltage* to the treatment pump is NL. The same convention is used for the rules that pertain to *Error* and *Feed Pump Voltage*. To synthesize the output from the two tables, the resulting *Change in Voltage* output to the treatment pump is the centroid of the combined output membership functions.

4.2.3 Test Matrix

Table 4.3 is a summary of the quantitative conditions for each test conducted. These tests were all completed using the same ANFIS controller. The tests were designed to introduce disturbances to the control system to record how the controller will react to the change in operating conditions. Test Nos. 4, 5, and 7 were conducted in a way where the change in flow rate was not observed by the controller. In these tests, the flow rate was changed by adding water directly to the reaction vessel and the

Table 4.3 – Summary of variables changed during course of the ANFIS controller tests.

Test No. (#)	Test Duration (sec)	Feed Flow (gpm)	Feed pH (pH)	Set Point (pH)	Baffle Position (txt)	Disturbance Time (sec)
1	1,800	2.25	2.95	7	Low	N/A
2	3,600	0.75 - 3.40	3.00	7	Low	ALL
3	2,000	0.95 - 2.95	2.95	6 - 8	Low	825 - 1,372
4	2,500	2.25- 5.50	3.10	7	Low	700 - 1,400
5	3,000	2.25	2.58 - 3.41	7	Low	1,000 - 1,731
6	1,500	2.25	2.97	7	None - Low	500 - 1,000
7	3,600	0.73 - 3.22	2.5 - 3.05	7	Low	400, 1,000, 2,000

feed pump did not see an increase in voltage that would normally correspond to an increase in flow rate. This result is interesting as it shows the controller may be able to operate with fewer variables, rules and membership functions. Results from these tests are presented in further detail in Chapter 5.

The tests were conducted for time periods which are dependent upon the conditions of the test. For example, Test No. 1 was conducted for 1,800 seconds, while the flow rate of AMD, feed pH, set-point, and baffle position remained constant. Conversely, Test No. 7 lasted for 3,600 seconds. The difference in duration can be directly contributed to the increased number of disturbances applied to the treatment system in addition to the amount of time required for the transient responses to reach steady state. Recall from Chapter 3 that the mean retention time for the entire system is between 1.25 and 5.32 minutes depending on the feed flow rate.

4.2.4 Analytical Methods

The results from the seven controller test were evaluated based on a qualitative basis. Each test run was plotted and evaluated to verify the accuracy of the controller during each test. Two conditions were required for the test to be considered successful. First, the controlled reaction of the pH exiting the first reactor needs to exhibit a steady state response around the desired set-point for tests of short duration. Second, the pH leaving the bench-scale system needs to stay within an acceptable limit, which was defined as ± 0.5 pH units from the desired pH set-point.

4.3 Prediction Tests

4.3.1 General Testing Goals

To develop a Sugeno ANFIS model, a training data set is required to “teach” the controller how the system will respond to changes and perturbations in the measured

and controlled variables. After learning the mathematical intricacies of these relationships, the controller can then specify the best course of action needed to move the current outgoing condition to the desired condition in an optimal manner. To properly train the model, the training data must cover a wide range of input and output conditions, since the controller will be limited only to the information contained within that data set. Since very limited data is available for ANFIS control of AMD treatment, the conditions and characteristics of an ideal training data set have not been rigorously defined. The goal of this experimental program was to test numerous training sets, all with unique characteristics, and compare which data sets are capable of making the most accurate future predictions. Unlike the prior tests, the development of a suitable controller was beyond the scope of the current work, since a Sugeno-type ANFIS controller also requires further optimization and lengthy data collection. Nevertheless, the data generated from this study can be considered the appropriate first step in a fully-operational Sugeno controller.

For this testing campaign, input pH, feed water flow rate, treatment chemical flow rate, current outgoing pH, and prior outgoing pH are considered input variables, while future outgoing pH is considered the single output variable. An example of this model architecture is shown in Figure 4.4. The time offsets defining “prior” conditions and “future” conditions are experimental parameters that were adjusted and tested during the experimental program. Unlike the Mamdani control tests, the training data acquisition does not require a pH set point. During training data acquisition, output pH is free to trend to any possible value, since the experimental goal is fully ascertain the input-output relationships across the entire pH range. Alternatively, variation in the system is driven by randomly-applied perturbations to the model inputs, such as feed flow rate and treatment chemical flow rate.

4.3.2 Data Acquisition

A separate Simulink file was developed to generate training data sets. Figure 4.5 shows the Simulink block diagram environment used to control the treatment system during data collection. The Simulink model controls the feed pump and the chemical treatment pump while recording data from the pumps and sensors at a frequency of once per second. The different colors used in the Simulink model represent the function of the block-set. Green blocks are used to generate a random frequency and magnitude for voltage of the chemical treatment pump. Next, the blue blocks represent the analog output signals sent to the two pumps. The feed pump was controlled with a single or multiple step inputs, while the treatment pump required a simple voltage to initiate or cease the pumping action. Magenta blocks are used to gather voltage data from the sensor/transmitter units. Finally, the black blocks represent the output data which is saved to a Matlab file. Note that the primary

Figure 4.4 – Example of Sugeno model architecture.

difference between the Mandami controller Simulink file and the current file is that the current file lacks a control loop.

Training data was generated by running the bench-scale treatment system and collecting the resulting data in the Matlab/Simulink program. To prepare the bench-scale system, the supply tank was filled with water and HCl to obtain a solution with a pH of 3.0. The pH was measured using an Oakton pH 450 handheld pH meter. Five gallons of water was mixed with 5.29 ounces of sodium carbonate (Na_2CO_3) in the treatment tank to generate an alkaline treatment solution. After preparation, the Simulink program was initiated and ran from 1,000 to 7,200 seconds, depending on the test trial. Data was recorded and archived in a Matlab file.

Figure 4.5 – Simulink block diagram used to collect training data.

In total, 23 independent tests were conducted. During these data collection experiments, the control voltage to the treatment pump was randomly varied in both frequency and magnitude, producing a variable addition of treatment chemical that was held constant for a random time duration. Initial training data was collected using a constant feed flow rate, while later tests introduced a variable feed flow rate in a similar fashion as the randomly applied treatment chemical addition. These dually varying input parameters created a highly dynamic system that produced a broad range of output conditions, albeit in a variable fashion between different data sets.

An example of a training data set is shown in Figure 4.6. Here, the pH of the effluent leaving the first reactor and leaving the entire system is represented by the plot at the top of the figure. The feed pH is shown in the second plot, while the feed flow and treatment pump flow are shown in the bottom two subplots. Figure 4.6 clearly shows the dynamic reactions between the input and output variables. For example, at 2,414 seconds the flow rate delivered by the feed pump decreased instantaneously from 2.14 GPM to 0.97 GPM. Shortly after this change, the flow rate to the treatment pump increased from 0.037 GPM to 0.070 GPM at 2,352 seconds. When taken together, the flow of acidic water was reduced in half, while the flow of alkaline slurry was doubled. This change in flow rates produced a large increase in the pH of water immediately after mixing in the first reactor.

4.3.3 Data Processing

After acquiring 23 independent data sets, each set was used to generate ANFIS models that were tested using each of the other sets as checking data. This process was repeated while adjusting the future prediction time delay (D) and the prior time offset (OS). Delay was tested at 27 points, varying incrementally from 1 to 180 seconds, while prior offset was tested at a single time of 10 seconds. These values were determined from initial shakedown tests that suggested realistic predictions were obtainable with these parameters. A limited number of trials also tested 5 and 15 second offset values, but these were deemed inadequate and unnecessary for the current experimental program. Given these parameters, a total of 13,662 tests were conducted (23 ANFIS models x 22 checking data sets x 27 delay times x 1 offset times).

To prepare the data for analysis, the raw data was first trimmed at the initial and final values to ensure that measured data points were available for the delay and offset positions (i.e. a prepared data set that spanned the entire test duration would have null values for these elements). The data was then arranged so that input and output parameters were in similar positions in the data arrays. For the actual training, input parameters include: the pH entering the treatment system (pH_1), the pH leaving the first reactor (pH_2), voltage supplied to the feed pump (V_{fp}), voltage supplied to the

Figure 4.6 – Example of training data set

treatment pump (Vtp), and the pH leaving the first reactor OS seconds in the past (pH_2-OS). The single output from the model is the pH leaving the first reactor D seconds in the future ($pH_2 + D$).

The Matlab function ANFIS was used to train the ANFIS models using the pre-processed data sets. The ANFIS routine requires an initial fuzzy inference system that denotes the number and shape of the membership functions as well as the training stopping criteria. For the current testing campaign, three membership functions were used for each input variable, and the Gaussian function was specified as the type. Once again, these parameters were determined from shakedown tests and represented a suitable tradeoff between computational efficiency and training accuracy.

The training duration was set to a maximum of 25 epochs, and the stopping criteria was defined as the point where the checking data prediction error was minimized. Checking data error was selected over training data to prevent overfitting. ANFIS training algorithms are designed to adjusted model parameters such that the error between the training data and the model prediction is reduced for each epoch. As the prediction error is reduced for the training data, the model will eventually begin fitting the small measurement errors, rather than the physical interactions fundamentally driving the model output. When this condition occurs, model error for the checking data set will begin to rise. As a result, the stopping criterion for the algorithm is defined as the point where the total model error of the checking data set is minimized. Training was conducted using a hybrid method, where a combination of least squares and back-propagation gradient descent methods are used to identify membership function parameters and firing strengths.

4.3.4 Data Analysis

Each of the 621 ANFIS models (23 data sets x 27 time delays) was tested against the remaining 22 data sets to determine the predictive capability inherent to each original data set. Using all these parameters, the total training time for each model was approximately 3 hours. In this case, “predictive capability” is quantitatively defined by the root mean squared error (RMSE) between the predicted future pH values and the actual pH values in the checking data sets. RMSE is given by:

$$RMSE_{Error} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (4.3)$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, and n is the number of predictions. Data sets with the lowest RMSE were considered optimal for use in training.

To characterize the data sets, several statistical parameters were calculate for each set, including mean, variance, the number of times the data crosses a pH of 7.0,

and the number of times the data set crosses its mean value. While these values do not necessarily encompass all of the possible statistical characterizations, they do provide a reasonable method to quantify the volatility and range of the data. When compared against the RMSE values, these characteristics will provide insight on what parameters are characteristic of superior training data sets.

In addition to the raw analysis of testing each model against each checking data set, a more selective analysis highlighted only the prediction of highly variable data sets. Since highly variable data sets are inherently more difficult to predict than relatively steady data, the inclusion of this parameters provide a more sophisticated measure of prediction quality. In this analysis, the five data sets with the highest variance were isolated, and the best predictors for these data sets were identified. These models and their original training data were then further analyzed to elucidate any other factors contributing to superior predictive capability.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Results of Controller Tests

A single Mamdani ANFIS controller was used in seven different tests to verify the ability of the controller to maintain a desired set-point given different disturbances to the bench-scale treatment system. Each of these experiments test the controllers adaptability when one or more variables are changed during the course of the test. Variables changed in these tests include flow rates, feed water pH, changing set-points, and combinatory effects when multiple variables are changed in unison.

5.1.1 Test 1 - Control Under Steady-State Conditions

This control scenario marks the first step in testing the optimal ANFIS controller. Here, all variables were held constant. The goal of the controller was to maintain an acceptable pH range at the provided set-point. Variables held constant in this scenario include incoming water flow rate, feed pH, and set-point. Figure 5.1 shows the results of the ANFIS controller during a test for a duration of 1,800 seconds. In this figure, four separate plots depict variables and conditions of the bench-scale system as the controller operates. The top plot shows the pH of the water immediately after treatment as it travels between the first and second reactor as a red line. The blue line represents the pH of the water leaving the overall system while the black line indicates the manually selected set-point, or desired pH. Finally, the remaining three plots show the feed pH, flow rate into the system, and voltage sent to the treatment pump, respectively. As a note, the mean residence time for this condition is approximately 1.75 minutes (105 sec.)

The controller displays a transient response to the startup conditions for a period of 150 seconds. Next, the controller overshoots the desired set-point by 0.87 pH units at 194 seconds. In contrast, the controller also stops supplying voltage to the treatment pump at 134 seconds when the pH leaving the first reactor is 5.98, well

below the desired set-point. This lag in response to a change in the control regime illustrates the difficulty in developing a robust controller for a nonlinear system such as an AMD treatment system. Specifically, the controller must possess the ability to predict when to stop applying voltage to a pump well before the system reached a desired set-point. Furthermore, the addition multiple perturbations to the flow and chemical characteristics will compound this difficulty.

The controller appears to reach a steady-state condition at approximately 400 seconds. Of particular note, the pH leaving the first reactor displays a wide variation around the set-point oscillating from approximately plus or minus one unit of pH. This variance is due to the aforementioned delay, from when the controller stops supplying alkaline material to when the final chemical reactions take place as the water exits of the first reactor. While this wide oscillation is a typical impediment to traditional control systems, the variation is mitigated by the extended residence time of water traveling throughout the entire treatment system. Despite this oscillation, the pH of the water leaving the treatment system approaches the desired set-point and maintains a consistent pH as it leaves the treatment system.

After reaching the steady state condition at 400 seconds, the final pH leaving the system still shows a pH of 6.16. The outlet pH does not approach equilibrium at the desired set-point until 1,400 seconds. Here the outlet pH is 6.86 showing an a steady state error of -0.14 pH units. While this deviation from the desired pH is minimal, a true value of 7.0 could be reached through the implementation of an offset to the desired set-point.

5.1.2 Test 2 - Control With Varying Flow Rate

The next test measured the effectiveness of the ANFIS controller while randomly varying the flow rate of the feed water delivered to the system as shown in Figure 5.2. Here, a random number generator was used to vary both the magnitude and frequency of the voltage applied to the feed pump as represented by the cyan line. The range of flows through the bench-scale system varied from a low of 0.75 gallons per minute to a maximum of 3.40 gallons per minute.

To compensate for the varying flows, the ANFIS controller adjusted the frequency and magnitude of the voltage applied to the treatment pump supplying the alkaline slurry to the system. The magenta line in the graph clearly shows distinct regions where the controller speed changes, due to changes with the ANFIS membership functions. As an example, at 1,480 seconds the flow drops from 3.4 gallons per minute to 0.94 gallons per minute. Next, at 1,504 seconds the voltage to the treatment pump is stopped for a duration of 66 seconds to compensate for the slower addition of acidic water to the treatment system. In contrast, when flow rate was at the highest point from 1,186

Figure 5.1 – Test 1 results: control under steady-state conditions.

seconds to 1,480 seconds, the controller initiated full output to the treatment pump at a frequency of approximately every 30 seconds.

The overall pH leaving the treatment system during this experiment varied slightly from 7.10 to 6.88. This variation in pH is consistent and would be considered acceptable if this controller was implemented in the field. In this test, both the pH leaving the first reactor and the pH leaving the system stayed within the acceptable range established for pH variation. The tight control shown here can be attributed to the high starting pH in this test. Here the controller is operating in the more precise rule base area where adjustments are in the medium and low range verses the first test where the controller started with high adjustments to the change in treatment pump voltage.

Furthermore, the amount of controller manipulation needed to achieve this outcome clearly shows the difficulty in attaining a consistent neutral pH. Should this system have a manual control strategy, similar to systems used in industry, the level of control attained here would be difficult at best. By using manual control when there are sizable and frequent variations in flow, the mine operators face the risk of wasting chemical treatment products as well as under-treating which could lead to a discharge of noncompliant water.

5.1.3 Test 3 - Control with Changing Set-Points

This control scheme is similar to the previous test; however, it employs an additional change in the desired pH set-point. Initially, the test was started using the random number generator to determine the magnitude and frequency of flow rate change into the system. The initial pH of the set-point was set to a value of 6.0. The controller displayed a transient response to the initial conditions and reached a steady-state point of equilibrium at approximately 300 seconds as shown by the red line in Figure 5.3.

At 825 seconds, a step change was initiated by changing the set-point from a value of 6.0 to 8.0. To compensate for this disturbance, the controller increased the frequency and duration of the voltage applied the treatment pump. This increased dosage of alkaline material brought the pH of the water leaving the first reactor from 6.13 to 8.02 in 90 seconds, where the pH leaving the first reactor began oscillating around the desired set-point. Additionally, another step change was initiated at 1,372 second by lowering the set-point back to 6.0. Following this change, the controller stopped supplying voltage to the treatment pump for 189 seconds when the pH leaving the reactor reached a value of 6.18.

A steady state condition was again achieved at a set-point of 6.0 at approximately 1,800 seconds. This longer duration to achieve steady state is directly attributed to

Figure 5.2 – Test 2 results: control with varying flow rates.

the characteristics of the system where pH can only be increased through control while a decrease in pH is dependent of the flow rate of acidic water delivered to the treatment system. Additionally, the rate of change in pH of water leaving the outlet of the system is represented by the blue line in Figure 5.3. After the first step change at 825 seconds, the pH exiting the system reached a maximum value of 7.41 which is 948 seconds after the change in set-point value. Should this test have been allowed to continue at the set-point of 8.0 for an extended period of time, the output pH would have approached a value of 8.0 as seen in Figure 5.1.

This control test is characteristic of a change in desired pH at a AMD treatment system. Often when certain effluent parameters, like manganese, approach levels of noncompliance, a higher pH is required to precipitate these metal hydroxides from the AMD. The ability of this controller to handle multiple set-points is significant, as in-field conditions require a robust controller able to adapt to multiple set-points as environmental conditions change at a site.

5.1.4 Test 4 - Control with a Surge in Flow

In this experiment, a surge in flow was applied to the feed rate of water entering the treatment system between 700 and 1,400 seconds as shown in Figure 5.4. This increase was accomplished by using two pump heads on the feed pump. To clarify, this test is anomalous because the ANFIS controller was not able to detect this doubling in flow, as the voltage to the pump remained constant (recall that the pump voltage is the variable supplied to the Simulink program). The constant pump rate can be confirmed by the cyan line showing flow rate into the system as seen in Figure 5.4. All changes in flow were accomplished manually by manipulating valves at the manifold preceding the first reactor.

Initially, the system was started and allowed to reach a point of equilibrium at approximately 550 seconds. Next, at 700 seconds, the valves on the manifold were manipulated to allow both pump heads to supply water into the first reactor. Before the valves at the manifold were turned to allow both flows of water to enter the system, one of the pump heads was recirculating the pumped water back into the feed tank. With this doubling in H^+ ions, the system was unable to maintain the desired set-point of 7.0; however, the ANFIS controller was able to maintain a steady state condition at approximately 6.7. The frequency at which the controller triggered the treatment pump doubled during this time period to compensate for the increased quantity of acidic flow into the system, as indicated by the magenta line at the bottom of Figure 5.4.

At 1,400 seconds the valves were again manually manipulated to bring the flow back to 2.25 gallons per minute. At this point, the controller overshoots the desired set-

Figure 5.3 – Test 3 results: control with changing set-points.

point of 7.0 by 0.27 pH units. Again, the frequency with which the treatment pump is activated decreases by approximately half.

While the controller was unable to detect the increased amount of flow coming into the system, it was still able to react to the perturbation in the flow by only using the two input variables, *Error* and *Change in Error*. This result indicates the robustness of the controller lies in the use of these two variables. While the importance of the control variable *IP Pump Voltage* is questionable given the results of this test, additional testing was not conducted to verify the necessity of the *IP Pump Voltage* variable using this bench-scale system. Pump capacity is the limiting factor in performing a validation test. The maximum feed pump flow rate is approximately 3.5 GPM.

The relevancy of this test to conditions experienced in the field is critical as unexpected conditions are always present at remote locations where this type of control would be used. This test in particular indicated that should an ANFIS controller be installed in an industry setting, the controller would be able to function despite the absence of an input variable. For example, during a large rain event, a surge in flow may damage or move a flow sensor from the installed position resulting in false or nonexistent readings. Additionally, damage and vandalism to treatment sites is common in CAPP by both wildlife and the local population. For this reason, a control scheme that can remain operational in the absence of an input variable is attractive to the end user.

5.1.5 Test 5 - Control with a Change in pH

During this test, the pH of the water entering the treatment system was lowered from 3.41 to 2.58. This change was accomplished by using two water supply tanks and manually manipulating the source of flow using valves located on the manifold preceding the first reactor. The decrease in pH was representative of a increase in the quantity of acidic material entering the system by a factor of 83. Due to the large increase in acidic material entering the reactor, an adjustment to the maximum allowable voltage being sent to the LS treatment was implemented for this test alone. This adjustment raised the maximum voltage from 5.25 volts to 8.00 volts as seen in the increased range of the magenta line in Figure 5.5.

As this test started, the controller had a transient response to start-up conditions from 0 to 90 seconds. At this point, the controller overshoot the desired set-point by 1.43 pH units. Again, as in previous tests, a steady-state condition was achieved; however, the magnitude of oscillation around the desired set-point was greatly increased due to the increased output of basic treatment chemical. For comparison, the oscillation in Figure 5.1 under steady state conditions were less than one pH unit.

At 1,000 seconds, the feed water supplying the treatment system was changed to the

Figure 5.4 – Test 4 results: control while a surge in flow is experienced by the system.

lower pH feed stream. As indicated by the green line representing feed pH, the pH sensor displayed a step change at 1,009 seconds from a pH of 3.41 to 2.58. The nine second lag in time was due to the distance the water must travel in the manifold to reach the pH sensor. Additionally, the pH exiting the first reactor indicated a sharp decline at 1,084 seconds and reaches a minimum value of 2.70 at 1,130 seconds. At 1,089 seconds, five seconds after the sharp decline, the controller started to compensate for the reduced pH by increasing the voltage to the chemical treatment pump. Alkaline slurry was applied to the lower pH stream for 187 seconds until the exiting pH approached a value of 6.0.

At approximately 1,370 seconds a steady state condition was achieved by the effluent stream leaving the first reactor at a pH of approximately 6.4, indicating a steady-state error of 0.6 pH units. This pH was maintained for the remaining time the lower pH water was introduced to the system. When the feed water was changed back to the original tank containing 3.41 pH water at 1,731 seconds, the controller responded by suspending the addition of treatment chemical for a duration of 359 seconds.

The lack of treatment chemical and increase in feed pH initiated a large gain in the pH of water leaving the first reactor at 1,808 seconds. At this point the ANFIS controller resumed periodic dosing of chemical treatment at a frequency similar to the time period before the step change was initiated. Again, the controller reached a steady-state condition after 1,808 seconds with large oscillations around the set-point of 7.0.

Also, of particular interest is the pH of the effluent leaving the treatment system. As seen by the blue line in Figure 5.5, the exiting pH remains close to the desired set-point despite the wide variations in the pH of the water exiting the first reactor. This result is attributed to the use of multiple cells within the treatment systems allowing the wide variations to balance as the effluent travels through the system before reaching the final outlet.

While a large and sudden decrease in pH is uncommon in an industry setting, this experiment does provide insight into what limitations may exist when implementing this type of control in the field. The limit applied to the upper range of the treatment pump is required to prevent the controller from over-treating during the period of time between when the treatment is applied and a response is received by sensors placed directly after treatment. Furthermore, additional rules or an increase in the number of membership functions may allow this controller to govern wider ranges of pH disturbances in a more effective manner.

Figure 5.5 – Test 5 results: control with change in feed pH.

5.1.6 Test 6 - Control with Change in Flow Dynamics

In this experiment, the flow dynamics of the system was changed for 500 seconds while all other variables remained constant. To accomplish the change in the reactor flow regime, all baffles were removed from the system at 500 seconds, as shown in Figure 5.6. By removing the baffles, the vessels in the treatment system became less effective as treatment reactors due to the change in characteristics. With the baffles installed in the low setting, the reactor was more similar to a plug flow system where a larger volume is effectively used. With no baffles installed, the reactor exhibited characteristics of short circuiting and a longer residence time.

Initially, the treatment system quickly approached a steady state condition as the effluent within the reactor at the start of the experiment was close to the desired set-point. After the baffles were removed, the ANFIS controller responded by increasing the frequency at which treatment chemical was administered to the first reactor. This response is illustrated by the magenta line during the time period of 550 to 1,000 seconds.

One interesting result from this experiment was the increase in the set-point offset during the period of time the baffles were removed. Previously, before the change was initiated the pH exiting the first reactor oscillated around the set-point between the values of 6.88 to 7.17. For the extent of time when the treatment system was operating without baffles, the pH exiting the first reactor continued to oscillate, but at a higher frequency and magnitude with high and low values in the range of 7.20 to 7.05 respectively.

In an industrial setting the loss of a pond curtain is a tangible event that can arise with increased flows or improper installation of the curtain. As seen in Figure 5.6, the ANFIS controller was capable of administering chemical at an appropriate rate to overcome the change in the flow regime. This observation is significant as it further shows the robustness of the controller in a dynamic environment where bypassing or an under-utilization of treatment reactor volume is present.

5.1.7 Test 7 - Control with Multiple Perturbations

Finally, the last experiment displays the effectiveness of the controller when presented with multiple disturbances. In this test, three buckets were filled with approximately 4.8 gallons of acidic feed water. The pH of the water in the first two buckets was identical to the feed at 3.05, while the last bucket contained water with a pH of 2.5. As shown in Figure 5.7, these buckets of water were manually added to the manifold preceding the pH sensor at the inflow to the first reactor. The additional water was poured into the manifold at 1,000, 1,500, and 2,000 seconds for a duration of 80

Figure 5.6 – Test 6 results: control during change in reactor residence time.

seconds each. This disturbance was not recorded by the flow sensors installed in the bench-scale system.

The first disturbance at 1,000 second occurred as the pH of the effluent leaving the first reactor was decreasing. The increased flow of acidic water perpetuates this decline and drove the pH of the system to 4.90 at 1,027 seconds. Additionally, the controller initiated treatment at 992 seconds, eight seconds prior to the surge in flow rate. At 1,500 seconds the second surge in flow was administered to the system. This surge did not have the same dramatic affect on the system and the controller sporadically increased the voltage to the treatment pump since the first disturbance. Finally the last bucket containing water a pH of 2.5 was added at 2,000 seconds. This disturbance had a profound effect on the pH leaving the first reactor resulting in a general lowering of the effluent pH by 3.2 units. To counter this surge of acidic water, the controller applied full voltage to the treatment pump for 167 seconds. Despite these disturbances, the pH exiting the overall treatment system maintained a pH range between 6.0 and 7.0.

Again, the controller was able to compensate for unexpected disorder introduced to the system. Frequently, in an industrial setting, multiple disturbances happen at the same time. As an example, a large rain event may induce increased flows, as well as a change in the pH of the water entering treatment pond. Further tests are required to capture the response of the controller with all imaginable combinations of disturbance; however, these experiments have shown an ANFIS controller reacts in an promising manner when exposed to perturbations on a bench-scale system.

5.1.8 Summary of Control Tests

Results from the controller testing campaign indicate the Mamdani type ANFIS controller is capable of implementing control when multiple disturbances are introduced to the bench-scale treatment system. The majority of the test results exhibited excellent operating characteristics. Large drops in the feed pH caused the most difficulty for the Mamdani controller. Table 5.1 shows the overall performance and unique takeaways for each test performed with the Mamdani controller.

5.2 Predictive Accuracy of ANFIS

This second experimental program was designed to test numerous data sets with uncommon characteristics to determine which characteristics make the most precise future predictions. Results of this testing program will provide a rigorous definition of the properties an optimal training data set should possess to generate an accurate ANFIS controller. These experiments are significant since little previous research exists defining the operation of ANFIS with an AMD treatment system.

Figure 5.7 – Test 7 results: control during multiple perturbations.

Table 5.1 – Table summarizing results of control tests

Test No.	Overall Performance	Key Points
1		Controller is able to maintain a the set-point under steady state conditions. Overall pH leaving treatment system exhibits good control.
2		Controller can operate with varying flow rates with a narrow oscillation around the set-point.
3		The controller is able to maintain control at multiple set-points and varying feed rates.
4		This test indicated the robustness of the controller lies in the variables <i>Error</i> and <i>Change in Error</i>
5		Modification to pump voltage limits is necessary to accommodate a large change in feed pH values.
6		Changes in flow dynamics have a minimal impact on the controller.
7		Multiple perturbations challenge to controllers functional ability; however, acceptable control is achievable.
<p>Legend</p> <p>Better Worse</p> <p>←————→</p>		

An example of results from this training routine is shown in Figure 5.8. This example shows model predictions exhibiting both high and low RMSE values. In the figure, the training Data Set No. 27 (DS-27) is represented by the plot on the left. Known output values are indicated by circles while the results from the training are shown by the red line. Likewise, the checking data DS-23 is shown on the right of the figure. Here, the actual values from the checking data set is shown using circles, and the red line is the predicted pH generated by the ANFIS model from DS-27.

In comparing the plots in Figure 5.8, DS-23 is shown to be a better predictor than DS-22, given the respective checking data. From a qualitative standpoint, the predicted data in Figure 5.8 (a) when compared to the actual values of the checking data are congruent. Conversely, when the same information in Figure 5.8 (b) is compared, there are many gross dissimilarities. This difference in predictive accuracy is quantitatively expressed by the RMSE values for these data sets when compared to each other. The more accurate prediction has a RMSE of 0.2252, while the inaccurate prediction has a RMSE of 5.91. This large difference is qualitatively observed by the goodness of fit of the line representing the predicted data. For the full test program, this analysis was repeated for each training data set against each checking data set for multiple time steps. These time steps indicate how far into the future the data set is able to predict.

Figure 5.9 shows the variation in the pH_2 variable (the pH leaving the first reactor) used in the experimental program. Each plot has a dashed line at the neutral pH line of 7.0 for reference. An example of the variability in the data sets is qualitatively seen when comparing DS-23 to DS-12.

5.2.1 Data Tables

Table 5.2 lists descriptive statistical values of the entire population while random processes were used to generate the data. The data in this table quantifies the variability that can be visually observed in Figure 5.9. The variance within the training data covers a wide range. The training data with a high amount of variance typically covers multiple points along the pH scale. An example of this behavior is seen in Figure 5.9 in DS-27 that has the highest variance at 4.22. On the other hand, data sets with low variance exhibit a consistent trend with a more linear characteristic (e.g. DS-12). Additional qualitative measures used in Table 5.2 include the number of times the pH_2 line crosses the mean value and the frequency that pH_2 crosses the neutral pH value of 7.0.

As an example of one of the 23 ANFIS models, Figure 5.10 shows the results of the DS-19 model in predicting each of the other data sets at all time delays in the future. A similar graph could be generated for each of the 23 ANFIS models, however, DS-19

Figure 5.8 – Example of ANFIS training showing results from both acceptable (a) and poor (b) predictions.

Figure 5.9 – Variation in pH_2 for all data sets used in the predictive analysis

Table 5.2 – Data sets used for prediction with a statistical analysis of the pH data used in prediction

Data Set	Mean	Std Dev	Variance	Cross Mean	Cross 7
1	8.44	0.51	0.26	9	0
2	7.64	0.54	0.30	4	10
3	6.25	1.42	2.01	7	1
4	6.15	1.31	1.71	9	8
5	9.33	1.23	1.51	13	1
6	7.70	2.00	4.01	7	9
7	9.24	1.27	1.61	15	1
8	9.45	1.27	1.61	10	1
9	7.90	1.04	1.08	11	5
11	6.94	0.72	0.51	1	3
12	7.25	0.39	0.15	1	3
13	6.36	1.04	1.07	1	3
17	7.59	1.64	2.68	10	9
18	8.27	1.76	3.08	15	3
19	7.76	1.63	2.67	14	15
20	9.93	0.20	0.04	7	0
21	5.67	1.71	2.91	17	2
22	5.59	1.95	3.80	27	5
23	5.77	1.73	2.98	38	6
24	6.37	1.31	1.71	11	7
25	6.57	1.39	1.94	12	10
26	7.77	2.05	4.20	6	6
27	5.13	2.05	4.22	13	6

is shown here for indicative purposes. This plot shows accurate future predictions are feasible up to ten seconds into the future; however, when the delay exceeds ten seconds the RMSE values tend to increase excessively, showing the difficulty involved when predicting further outcomes further into the future. This trend is generically true for all checking data, even though the magnitude of the error varies, as some are easier and more difficult to predict when using DS-19. Similar trends were observed for all of the other 23 models. This indicates the limitation of the ANFIS predictive model is approximately 0.05% of the average mean residence time for the bench-scale system.

Figure 5.10 – DS-19 results comparing RMSE to the delay in predicting all data sets

To analyze the large number of results, the initial analysis focused on the models ability to predict the 10 most variable data sets. Table 5.3 shows the results of the best and worst performing predictive data sets listed in order of descending and ascending RMSE respectively. The checking data was ranked from most difficult to least difficult to predict, based on the amount of variance in the pH of the checking data set. Table 5.3a shows the best predictors of the highly variable data sets, listed in order of RMSE. Table 5.3b lists the same information for the worst predictors. Among the best predicting models training DS-19 appears eight times. This data

set is considered to be the most accurate predictor due to the high frequency of occurrence. Likewise, DS-18, DS-17 and DS-23 appear in this list on a frequent basis. Additionally, among the poorly performing training data, DS-05 appears the most frequently with seven occurrences. This data set is considered to worst performing training data among the group of 23 training data sets. DS-20, DS-07, DS-04 and DS-03 also appear on this list several times.

Table 5.3 – Error values of the most and least accurate predicting training data sets when used to predict the most variable checking data.

(a) Best predicting data sets								
Best Predictors								
Checking Data	Variance	Prediction Data Set (RMSE)						
27	4.22	23	(0.44)	19	(0.64)	17	(0.65)	
26	4.20	7	(0.35)	19	(0.35)	17	(0.42)	
6	4.01	19	(0.56)	17	(0.62)	18	(0.66)	
22	3.80	9	(0.28)	13	(0.31)	18	(0.32)	
18	3.08	19	(0.50)	17	(0.55)	23	(0.61)	
23	2.98	19	(0.20)	21	(0.20)	27	(0.23)	
21	2.91	19	(0.49)	17	(0.57)	18	(0.75)	
17	2.68	19	(0.41)	23	(0.48)	18	(0.48)	
19	2.67	17	(0.45)	18	(0.48)	23	(0.52)	
3	2.01	4	(0.96)	19	(1.07)	6	(1.11)	

(b) Worst predicting data sets								
Worst Predictors								
Checking Data	Variance	Prediction Data Set (RMSE)						
27	4.22	5	(106)	20	(64)	1	(27)	
26	4.20	4	(8.2)	3	(7.8)	23	(5.9)	
6	4.01	20	(83)	5	(59)	21	(22)	
22	3.80	26	(11)	4	(6.7)	3	(6.2)	
18	3.08	5	(28)	7	(8.5)	20	(2.9)	
23	2.98	4	(9.3)	3	(9.1)	1	(1.4)	
21	2.91	5	(13)	7	(12)	8	(11)	
17	2.68	5	(23)	7	(4.7)	8	(2.9)	
19	2.67	5	(23)	20	(7.5)	7	(5.4)	
3	2.01	20	(72)	5	(25)	23	(15)	

This categorization was used to define the most accurate and inaccurate models. Next, the data sets used to derive the models in these two categories (best predictors and worst predictors) were compared on a qualitative basis to determine common qualities

both sets may contain. This effort was used to isolate the desirable characteristics of the accurate data sets and the undesirable characteristics in the poorly performing training data. This comparison has yielded several results which follow.

One observation made when comparing the data sets is seen in Table 5.3, which shows the placement of training DS-23 within both the good and poor predicting categories. Upon further review of DS-03, DS-23, and DS-26, it was noted that both DS-03 and DS-26 had several values which lie in the upper region of the pH scale. Alternatively, DS-26 had very few data points in the upper range of the pH scale. This lack of values residing in the upper limits of the pH scale of the training data may be attributed to the poor performance of DS-23 to predict DS-03 and DS-26.

Figure 5.11 shows a histogram of the frequency of pH values contained in the best and worst data sets. This plot shows that the best predicting data sets have distributions of pH values which cover a wide range. Alternatively, the worst predicting data sets tend to be skewed toward the higher end of the pH scale. Additionally, the poorly performing data sets which are not skewed tend to have few common data points. This finding is seen in DS-20, DS-04 and DS-03.

Figure 5.12 shows the average RMSE for all of the predictive testing. This plot clearly shows DS-17, DS-18, DS-19 and DS-23 have the lowest average RMSE values when used to predict all of the checking data-sets. Likewise, the data sets which perform the worst exhibit a higher RMSE value. This data differs from the previous results as it is inclusive of all of the checking data sets rather than the 10 most variable data sets. Here the inclusivity of the data has lowered the RMSE value of some of the poor performers. This result can be explained by the ability of a poor performing data set to predict similar data sets or those with low variability with less error than a data set with high variability.

5.2.2 Discussion

When analyzing the results of these tests, a compelling correlation among the characteristics of good and bad training data is observed. The data sets which exhibited the best predictive ability have a mean value that lies within the steep part of the titration curve along the neutral region of the pH scale, as shown in Figure 2.2. Likewise, the data sets which perform poorly in predicting variable data have mean values which lie outside of the steep area in the titration curve. This finding is significant and emphasizes the importance of the nonlinear behavior of the pH treatment process. Furthermore, this finding indicates that to build an accurate Sugeno type ANFIS controller, the balance of the training data should dwell within the region of the titration curve where small changes to the system variables induce a large change in the outgoing pH of the effluent.

(a) Best predicting data sets

(b) Worst predicting data sets

Figure 5.11 – Histogram showing the frequency of pH values for the predicting data set categories.

Figure 5.12 – Boxplot of training error RMSE for all data sets. (DS-05 not shown due to high RMSE value)

Additionally, the quantity and frequency of data must vary across the pH scale to achieve an accurate prediction. As seen in Figure 5.11 the data sets which result in precise forecasts have a wider distribution when compared to the distribution of values seen in unreliable training data. Likewise the quantity of data available to train the ANFIS system is indicative of the data sets performance. While a discrete number for the quantity of data points needed has not been established, it is inferred that the more data available for training the ANFIS architecture the more accurate the prediction will be. This assumes that all other factors (large distribution of data with a mean in the steep region of the titration curve) have been met.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

This work describes the development of a bench-scale system and ANFIS controller to evaluate the feasibility of using this technology at operating mine sites to treat effluent. While pH control has been successfully implemented in a variety of industrial applications in the chemical sector, CAPP mines have yet to adopt advanced control technologies for treating AMD. As a result, the industry has seen increased regulatory scrutiny, the discharge of non-compliant water, and the inefficient use of manpower. These factors increase the cost of maintaining water discharges. This increased cost justifies additional investment in an advanced control system.

An advanced controller designed to control the flow rate of an alkaline material into the bench-scale AMD treatment plant will benefit CAPP coal mine operators. The controls tested here are based on the fuzzy logic control theory proposed by Mamdani (1974). Development of this controller relied on the knowledge and experience of the developer to build the controller. This research verifies the conditions explained in the literature review regarding the conditions required to implement a fuzzy controller.

This controller was evaluated using seven different tests to simulate disturbances that may be encountered by an AMD treatment system used in the CAPP region. Results from this testing array show the controller's response to these disturbances is adequate to maintain a desired set-point around a neutral pH range. These results are encouraging and warrant further research in the development of an automated control system for use at CAPP mine with persistently non-compliant outlets. Despite this optimistic result, the controller does have limitations and cannot adapt to unrealistic swings in the physical states of the system. As an example, it is unrealistic to expect this form of control to maintain a set-point when extreme disturbances are introduced into the treatment system. In §5.1.5 the limits placed on the treatment pump required manipulation to enable the Mamdani controller to compensate for a large change in

the pH of feed water. This difficulty in controlling the disturbance is seen as the controller oscillates outside the acceptable range by approximately 2 pH units.

Additionally, a full-factorial test was conducted to assess the predictive capability of ANFIS models. These tests resulted in the qualitative description of parameter characteristics required to develop a training data set capable of reliably predicting an independent set of data with a high variation. Several key factors were identified through the experiments that are necessary to produce a reliable training data set. First, the mean of the predictor data set should lie in the steep part of the titration curve located around the neutral region of 7.0. This finding is relevant because the steep area of the titration curve yields large variations in pH when small changes to the flow variables are encountered. It seems reasonable that a large array of points in this region are needed to fully discern that relationship. Second, the distribution of pH values used to predict an independent system must be distributed across the pH scale. When data sets with an extreme skew to one side of the pH scale are used to predict variable data sets, a large amount of error results. Third, future predictive capacity of an ANFIS controller is limited to approximately 10 seconds into the future.

Finally, this research has demonstrated the ability to produce a reliable controller is a non-trivial process. While many hours were devoted to developing and validating the controller for the bench test, the resulting product is not yet suitable for operation in the field. Given the unique characteristics seen at AMD treatment sites, any form of control must be developed on a site specific basis. Additionally, this control algorithm must be adapted for alternate hardware and software platforms before a comprehensive control system can be utilized by CAPP coal operators. This adaptation is required because small processing units are used in the remote locations where NPDES outlets are located. These units typically rely on solar power to operate, limiting the computational ability of the processor.

6.2 Recommendations / Future Work

While this research has demonstrated an ANFIS type of controller is feasible for use with AMD treatment systems, further work is required to develop a controller suitable for industrial use. The author of this thesis recommends the following areas where additional research may be continued:

1. Further validation of the control scheme is required to confirm the suitability of this controller with alternate treatment chemicals and feed stocks. The substitution of AMD collected from various CAPP AMD sites will confirm the suitability of this controller in more realistic applications. This type of testing is necessary as chemical interactions that occur when metal oxides are released

from the treated water may change the membership function in relation to the control variables. Additionally, sodium carbonate was used exclusively in this research and additional alkaline chemicals at different concentrations will impact the parameters used in this testing.

2. Further work is recommended to test the ability of the controller for use with different water quality parameters. For example, a turbidity sensor may be incorporated into this system to autonomously administer flocculant during heavy rain events to reduce the chance of exceeding aluminum or TSS parameters. Additionally, a dissolved oxygen sensor could also be beneficial to the overall treatment system to control oxidizing agents to aid in the precipitation of certain metals.
3. The Mamdani style controller should be converted to a Sugeno FIS to increase operational efficiency of the system. This operational efficiency will be necessary for implementation in an industrial application where utilities are non-existent. Furthermore, this control scheme will require modification to allow implementation on the micro-computing platforms (e.g. Arduino, Raspberry Pi) capable of operation in remote locations using minimal power. Once possible solution, which was not tested, is decreasing the frequency which measurements are recorded. By using fewer data points, a lower power consumption by the controller may be achieved.
4. The results from the predictive capability study are incomplete and a more detailed investigation is warranted. This objective can be accomplished by increasing the different variables used in developing the ANFIS model. For example, the effect of adding more membership functions or changing the defining shape may significantly decrease the resulting error in predictions. Furthermore, additional research is required to determine the optimum number of epochs which are used to train the ANFIS system. After this work is complete, a controller based on the ANFIS model should be developed.

BIBLIOGRAPHY

- Afram, A. and F. Janabi-Sharifi (2014). Theory and applications of HVAC control systems - A review of model predictive control (MPC). *Building and Environment* 72, 343–355.
- Ahmad, J., J. Purbolaksono, and L. C. Beng (2010). Tube failures due to cooling process problem and foreign materials in power plants. *Nuclear Engineering and Design* 240(7), 1882–1886.
- Akcil, A. and S. Koldas (2006). Acid Mine Drainage (AMD): causes, treatment and case studies. *Journal of Cleaner Production* 14(12-13 SPEC. ISS.), 1139–1145.
- Alpha Natural Resources (2014, mar). Alpha Natural Resources Concludes Comprehensive Environmental Agreement.
- Andone, D., R. Dobrescu, A. Hossu, and M. Dobrescu (2006). Application of fuzzy model predictive control to a drum-boiler. *Integrated Computer-Aided Engineering* 13(4), 347–359.
- Astrom, K. J. and T. Hagglund (2001). The future of PID control. *Control Engineering Practice* 9(11), 1163–1175.
- Berenji, H. R. (1992). Fuzzy logic controllers. In *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, pp. 69–96. Springer.
- Bhatt, M. S. (2006). Effect of water side deposits on the energy performance of coal fired thermal power plants. *Energy Conversion and Management* 47(9-10), 1247–1263.
- Booth, R. B. and L. C. Mead (1969). Prevention of and removal of scale formation in water systems.
- Bridle, M. (2006). Method and apparatus for treating water to reduce boiler scale formation.

- Carvajal, J., G. Chen, and H. Ogmen (2000). Fuzzy PID controller: Design, performance evaluation, and stability analysis. *Information sciences* 123(3), 249–270.
- CH2M HILL (2010). Review of Available Technologies for the Removal of Selenium from Water. Technical report, North American Metals Council.
- Chang, F.-J., Y.-H. Tsai, P.-A. Chen, A. Coynel, and G. Vachaud (2015). Modeling water quality in an urban river using hydrological factors–data driven approaches. *Journal of environmental management* 151, 87–96.
- Chen, Q., R. Long, S. Quan, and L. Zhang (2014). Nonlinear Recurrent Neural Network Predictive Control for Energy Distribution of a Fuel Cell Powered Robot. 2014.
- Dutta, S. (1993). Fuzzy logic applications: Technological and strategic issues. *Engineering Management, IEEE Transactions on* 40(3), 237–254.
- Elbisy, M. S., H. M. Ali, M. A. Abd-Elall, and T. M. Alaboud (2014). The use of feed-forward back propagation and cascade correlation for the neural network prediction of surface water quality parameters. *Water Resources* 41(6), 709–718.
- Fahmy, R., H. Zaher, and A. Kandil (2015). A Comparison between Fuzzy Inference Systems for Prediction (with Application to Prices of Fund in Egypt). *International Journal of Computer Applications* 109(13), 6–11.
- Fogler, H. S. (2005). *Distributions of Residence Times for Chemical Reactors* 13 (Fourth Edi ed.). Upper Saddle River: Pearson Education, Inc.
- Gazea, B., K. Adam, and a. Kontopoulos (1996). A review of passive systems for the treatment of acid mine drainage. *Minerals Engineering* 9(1), 23–42.
- Ghee, N. T., S. Kumaresan, and L. C. Fan (2002). Fuzzy PID controller to control the pH neutralization process. *Student Conference on Research and Development*, 330–334.
- Gustafsson, T. K. and K. V. Waller (1992). Nonlinear and Adaptive Control of pH. *Ind. Eng. Chem. Res.* 31, 2681–2693.
- Hayward, G. and V. Davidson (2003). Fuzzy Logic and Applications. *Analyst* 128(I), 1304–1306.
- Hedin, R. S., R. W. Nairn, and R. L. Kleinmann (1994). Information Circular 9389 : Passive treatment of coal mine drainage. Technical report.

- Heishman, A. S. and R. G. Mclusky (2012). Quality Standards Through Senate Bill 562. *W. Va. Law Review 2012 - 2013 115*, 1031–1050.
- Henson, M. A. and D. E. Seborg (1994). Adaptive nonlinear control of a pH neutralization process. *IEEE Transactions on Control Systems Technology 2*(3), 169–182.
- Hill, D. W. (1969). Neutralization of Acid Mine Drainage. *Journal (Water Pollution Control Federation) 41*(10), 1702–1715.
- Hoehn, R. C. and D. R. Sizemore (1977). Acid mine drainage (AMD) and its impact on a Virginia stream. *Water Resources Bulletin 13*(1), 153–160.
- Huntsman, B. E., J. G. Solch, and M. D. Porter (1978). Utilization of Sphagnum species dominated bog for coal acid mine drainage abatement. In *Proc Geol Soc America, 91st Annual Meeting (Abstracts), Toronto, Ontario, Canada*, Volume 322.
- Ibrahim, R. (2010). Practical modelling and control implementation studies on a pH neutralization process pilot plant. (March).
- Jain, L. C., M. Seera, C. P. Lim, and P. Balasubramaniam (2014). A review of online learning in supervised neural networks. *Neural Computing and Applications 25*(3-4), 491–509.
- Jang, J., N. Gulley, and I. MathWorks (1998). Fuzzy Logic Toolbox: For Use with MATLAB: User’s Guide. *Order A Journal On The Theory Of Ordered Sets And Its Applications*, 289.
- Jang, J.-S. (1996). Input selection for ANFIS learning. *Proceedings of IEEE 5th International Fuzzy Systems 2*, 1493–1499.
- Jang, J. S. R. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man and Cybernetics 23*(3), 665–685.
- Jiménez, A., J. Aroba, M. L. de la Torre, J. M. Andujar, and J. A. Grande (2009). Model of behaviour of conductivity versus pH in acid mine drainage water, based on fuzzy logic and data mining techniques. *Journal of Hydroinformatics 11*(2002), 147.
- Johnson, D. B. and K. B. Hallberg (2003). The microbiology of acidic mine waters. *Research in Microbiology 154*, 466–473.
- Johnson, D. B. and K. B. Hallberg (2005). Acid mine drainage remediation options: A review. *Science of the Total Environment 338*(1-2 SPEC. ISS.), 3–14.

- Kalin, M., A. Fyson, and W. N. Wheeler (2006). The chemistry of conventional and alternative treatment systems for the neutralization of acid mine drainage. *Science of the Total Environment* 366(2-3), 395–408.
- Kaur, A. and A. Kaur (2012). Comparison of Mamdani-Type and Sugeno-Type Fuzzy Inference Systems for Air Conditioning System. *International Journal of Soft Computing & Engineering* 2(2), 323–325.
- Kickert, W. J. M. and H. R. Van Nauta Lemke (1976). Application of a fuzzy controller in a warm water plant. *Automatica* 12(4), 301–308.
- Kim, B. H. (2014). An adaptive neural network learning-based solution for the inverse kinematics of humanoid fingers. *International Journal of Advanced Robotic Systems* 11(1), 1–9.
- King, S. L. (1999). Neural Networks vs. Multiple Linear Regression for Estimating Previous Diameter. In *12th Central Hardwood Forest Conference*, Lexington, pp. 159 – 166.
- Kleinmann, R. L. P. (2001). Prediction of Water Quality at Surface Coal Mines. *The National Mine Land Reclamation Center Located at West Virginia University in Morgantown, West Virginia* 239, 247.
- Laudani, A., G. M. Lozito, F. Riganti Fulginei, and A. Salvini (2015). On Training Efficiency and Computational Costs of a Feed Forward Neural Network: A Review. *Computational Intelligence and Neuroscience* 2015, 1–13.
- Lei, Y., Z. He, Y. Zi, and Q. Hu (2007). Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mechanical Systems and Signal Processing* 21(5), 2280–2294.
- Li, H., M.-K. Hsieh, S.-H. Chien, J. D. Monnell, D. A. Dzombak, and R. D. Vidic (2010). Control of mineral scale deposition in cooling systems using secondary-treated municipal wastewater. *Water Research* 45(2), 748–760.
- Mamdani, E. (1974). Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers* 121(12), 1585.
- Mamdani, E. (1976). Advances in the linguistic synthesis of fuzzy controllers. *Int. J. Man-Machine Studies* 8(6), 669 – 678.
- McAvoy, T. J., E. Hsu, and S. Lowenthal (1972). Dynamics of pH in controlled stirred tank reactor. *Industrial Engineering Chemical ProcessDes. Developmen* 11(1), 68–70.

- McMurry, J. and R. C. Fay (1998). *Chemistry*. Prentice Hall.
- McNeill, D. and P. Freiburger (1994). *Fuzzy logic: The revolutionary computer technology that is changing our world*. Simon and Schuster.
- Mehta, R. and S. K. Jain (2009). Optimal operation of a multi-purpose reservoir using neuro-fuzzy technique. *Water Resources Management* 23(3), 509–529.
- Misir, D. and H. A. Malki (1996). Design and analysis of a fuzzy proportional-integral-derivative controller. *Fuzzy Sets and Systems* 79(3), 297–314.
- Najah, A., A. El-Shafie, O. A. Karim, and A. H. El-Shafie (2014). Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring. *Environmental Science and Pollution Research* 21(3), 1658–1670.
- Narendra, K. S. and K. Parthasarathy (1990). Identification and control of dynamical systems using neural networks. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* 1(1), 4–27.
- Navghare, S. S., G. G. Bodhe, and Shruti (2011). Design of Adaptive pH Controller using ANFIS. *International Journal of Computer Applications(0975-8887)* 33(6), 41–48.
- Nilashi, M., K. Bagherifard, O. Ibrahim, N. Janahmadi, and M. Barisami (2011). An Application Expert System for Evaluating Effective Factors on Trust in B2C Websites Trust, Security, ANFIS, Fuzzy Logic, Rule Based Systems, Electronic Commerce. *Engineering* 03(11), 1063–1071.
- Nise, N. S. (2007). *Control Systems Engineering* (4th ed.). John Wiley & Sons.
- Ozger, M. (2009). Comparison of fuzzy inference systems for streamflow prediction. *Hydrological Sciences Journal* 54(2), 261–273.
- Petchinathan, G., K. Valarmathi, D. Devaraj, and T. K. Radhakrishnan (2014). Local Linear Model Tree and Neuro-Fuzzy System for modeling and Control of an Experimental pH Neutralization Process. 31(02), 483–495.
- Rutherford, D. A. and G. C. Bloore (1976). The Implementation of. pp. 572–573.
- Schulze, E. and R. Eberhard (1992). Device for eliminating boiler scale using arrangement of magnetic coils and baffle plates.
- Shen, H.-F., X.-J. Zhu, M. Shao, and H.-f. Cao (2013). Neural Network Predictive Control for Vanadium Redox Flow Battery. *Journal of Applied Mathematics* 2013, 1–7.

- Singh, K. P., A. Basant, A. Malik, and G. Jain (2009). Artificial neural network modeling of the river water quality-A case study. *Ecological Modelling* 220(6), 888–895.
- Singh, P., H. K. Mohanta, and A. Kansal (2011). Fuzzy logic control of a pH neutralization process. In *8th Control Instrumentation and System Conference (CISCON-11)*, Manipal.
- Skousen, J. (2003). Total Maximum Daily Loads and Anti-Degradation. pp. 1 – 6.
- Skousen, J., T. Hilton, and B. Faulkner (1997). Overview of acid mine drainage treatment with chemicals. *Agriculture & Natural Resources*, 21.
- Skousen, J. G., A. Sexstone, and P. F. Ziemkiewicz (2000). Acid mine drainage control and treatment. *Reclamation of Drastically Disturbed Lands* (41), 1–42.
- Stewart, B. T., A. N. Venkat, J. B. Rawlings, S. J. Wright, and G. Pannocchia (2010). Cooperative distributed model predictive control. *Systems and Control Letters* 59(8), 460–469.
- Takagi, T. and M. Sugeno (1985). Fuzzy Identification of Systems and Its Applications to Modeling and Control. *Elsevier Inc*, 116 – 132.
- USEPA (1994). Acid Mine Drainage Prediction. Technical Report December, Washington, D.C.
- USEPA (2000). Abandoned mine site characterization and cleanup handbook (EPA 910-B-00-001). (August), 408.
- USEPA (2004). 40CFR § 230.3.
- USEPA (2008). United States of America v. Massey Energy Company.
- USEPA (2009). United States of America v. Patriot Coal Corp.
- USEPA (2011a). United States of America v. Arch Coal, Inc.
- USEPA (2011b). United States of America v. CONSOL Energy, Inc.
- USEPA (2014). United States of America v. Alpha Natural Resources, Inc.
- USEPA (2015). United States of America v. Arch Coal, Inc.
- Vasickaninova, A., M. Bakosova, A. Meszaros, and J. J. Klemes (2011). Neural network predictive control of a heat exchanger. *Applied Thermal Engineering* 31(13), 2094–2100.

- Visioli, A. (2001). Tuning of PID controllers with fuzzy logic. *IEE Proceedings - Control Theory and Applications* 148(I), 1.
- Viswanathan, R., J. Sarver, and J. M. Tanzosh (2006). Boiler Materials for Ultra-Supercritical Coal Power Plants - Steamside Oxidation. *Journal of Materials Engineering and Performance* 15(June), 255–274.
- Wang, D., S. Li, and X. Zhou (2013). Control-oriented modeling and real-time control for the ozone dosing process of drinking water treatment. *Environmental Science and Technology* 47, 2197–2203.
- Wieder, R. K. and G. E. Lang (1982). Modification of acid mine drainage in a freshwater wetland. In *Symposium on wetlands of the unglaciated Appalachian region*, pp. 43–53. West Virginia University Press Morgantown, WV.
- Yu, D. L. and J. B. Gomm (2003). Implementation of neural network predictive control to a multivariable chemical reactor. *Control Engineering Practice* 11(11), 1315–1323.
- Zadeh, L. A. (1965). Fuzzy Sets.
- Zadeh, L. A. (1996). Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems* 4(2), 103–111.
- Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences* 178(13), 2751–2779.
- Zamarreño, J. M. and P. Vega (1998). State space neural network. Properties and application. *Neural networks : the official journal of the International Neural Network Society* 11, 1099–1112.
- Zhang, J. and a. J. Morris (1999). Recurrent neuro-fuzzy networks for nonlinear process modeling. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council* 10(2), 313–326.
- Zhang, J. and J. Morris (1997). Neuro-Fuzzy Networks for Process Modelling and Model-Based Control. In *Neural nad Fuzzy Systems: Design, Hardware and Applications*, Volume 133, pp. 6/1 – 6/4.
- Ziemkiewicz, P. (1998). Steel Slag : Applications for Amd Control. (304), 44–62.
- Ziemkiewicz, P. F.; Skousen, J. G.; Simmons, J. (2003). Long-term Performance of Passive Acid Mine Drainage Treatment Systems. *Mine Water and the Environment* 22, 118 – 129.