Theses and Dissertations--Electrical and Computer Engineering

Electrical and Computer Engineering

2020

FAULT IDENTIFICATION ON ELECTRICAL TRANSMISSION LINES USING ARTIFICIAL NEURAL NETWORKS

Christopher W. Asbery
University of Kentucky, christopher.asbery@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0002-1594-7294
Digital Object Identifier: https://doi.org/10.13023/etd.2020.167

Right click to open a feedback form in a new tab to let us know how this document benefits you.

FAULT IDENTIFICATION ON ELECTRICAL TRANSMISSION LINES USING
ARTIFICIAL NEURAL NETWORKS

_____

DISSERTATION

_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By
Christopher Wayne Asbery
Lexington, Kentucky
Director: Dr. Yuan Liao, Professor of Electrical and Computer Engineering
Lexington, Kentucky
2020

ABSTRACT OF DISSERTATION

FAULT IDENTIFICATION ON ELECTRICAL TRANSMISSION LINES USING
ARTIFICIAL NEURAL NETWORKS

Transmission lines are designed to transport large amounts of electrical power from the point of generation to the point of consumption. Since transmission lines are built to span over long distances, they are frequently exposed to many different situations that can cause abnormal conditions known as electrical faults. Electrical faults, when isolated, can cripple the transmission system as power flows are directed around these faults therefore leading to other numerous potential issues such as thermal and voltage violations, customer interruptions, or cascading events. When faults occur, protection systems installed near the faulted transmission lines will isolate these faults from the transmission system as quickly as possible. Accurate fault location is essential in reducing outage times and enhancing system reliability. Repairing these faulted elements and restoring the transmission lines to service quickly is highly important since outages can create congestion in other parts of the transmission grid, therefore making them more vulnerable to additional outages. Therefore, identifying the classification and location of these faults as quickly and accurately as possible is crucial.

Diverse fault location methods exist and have different strengths and weaknesses. This research aims to investigate the use of an intelligent technique based on artificial neural networks. The neural networks will attempt to determine the fault classification and precise fault location. Different fault cases are analyzed on multiple transmission line configurations using various phasor measurement arrangements from the two substations connecting the transmission line. These phasor measurements will be used as inputs into the artificial neural network.

The transmission system configurations studied in this research are the two-terminal single and parallel transmission lines. Power flows studied in this work are left static, but multiple sets of fault resistances will be tested at many points along the transmission line. Since any fault that occurs on the transmission system may never experience the same fault resistance or fault location, fault data was collected that relates to different scenarios of fault resistances and fault locations. In order to analyze how many different fault resistance and fault location scenarios need to be collected to allow accurate neural network predictions, multiple sets of fault data were collected. The multiple sets of fault data contain phasor measurements with different sets of fault resistance and fault location combinations. Having the multiple sets of fault data help determine how well the neural networks can predict the fault identification based on more training data.

There has been a lack of guidelines on designing the architecture for artificial neural network structures including the number of hidden layers and the number of neurons in each hidden layer. This research will fill this gap by providing insights on choosing effective neural network structures for fault classification and location applications.

<div style="text-align: right;">

Christopher Wayne Asbery
_____

May 11, 2020
_____
Date

</div>

FAULT IDENTIFICATION ON ELECTRICAL TRANSMISSION LINES USING
ARTIFICIAL NEURAL NETWORKS

By

Christopher Wayne Asbery

Dr. Yuan Liao	
Director of Dissertation	
Dr. Aaron Cramer	
Director of Graduate Studies	
May 11, 2020	
Date	

DEDICATION

First, this dissertation and related PhD research has been dedicated to my beautiful wife and son, Kirsten and Peyton Asbery. Working a full-time job while working towards the completion of my PhD degree has been extremely stressful and rewarding for my entire family and myself. I want to thank them both for their continued support for my career and educational goals that help fulfill my aspirations in life. Having my PhD degree in Electrical and Computer Engineering has been a goal of mine since I began my Electrical and Computer Engineering academic career in May of 2004.

Secondly, this dissertation has been dedicated to my loving parents, Jerry and Martha Asbery. My parents have always been extremely supportive of my education and career goals by always encouraging me and providing the support that was needed. Without their continued appreciation of my career development starting at a young age, I'm not sure I would be as successful as I am. Coming from a first-generation college family, my family and I are very proud of my accomplishments and are excited to see how my future unfolds.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

ix

LIST OF FIGURES

Chapter 1   Purpose and Significance of the Research

This dissertation is focused on developing an approach that will identify electrical faults on electrical power systems with specific focus on the transmission system. The context of electric fault identification is meant to recognize the type or classification of an electric fault that has occurred on the transmission system and determine the accurate location of that fault. This dissertation will begin by describing basic background information on the electric power system (which will include the transmission system). This is an important foundation needed to understand the scope of this research. Once the background of the power system has been introduced, the discussion will then adjust its focus to the idea of what electric faults represent and how they might occur on the transmission system. Knowing the consequences that electrical faults present to the transmission network, it becomes critically evident that these faults be identified and restored as quickly as possible. This research will assume that a fault has been detected and the associated fault data is available to analyze the identification of the fault. This research will use a specific intelligent technique based on artificial neural networks (ANN) to assist in providing the identification of these faults. The intelligent technique studied will perform analysis on a two-terminal single transmission line and a two-terminal parallel transmission line.

1.1   Electric Power System Introduction

Electric power systems are expressed in three major components or categories: generation, transmission or sub-transmission, and distribution. Generation, which is also

known as the electrical power sources (machines) for the power system grid, begins the

process by generating bulk amounts of power that will be transported and consumed by

the end users. Generation of electric power is produced in a variety of output levels

between many different types of generation sources. Since this dissertation is focused on

the electric utility power grid, its only appropriate to focus on utility scale generation

sources. Utility scale generation is produced from sources of coal, natural gas, nuclear,

geothermal, wind, and solar photovoltaic. Utility scale generation generates large

amounts of electricity, ranging from a few megawatts (MW) to over a thousand MW

from a single generation site. These generation sources account for approximately 86%

for the total power generation in the United States [1]. Reference [1] focused on the

analysis of baseload and intermediate power plants while ignored the power peaking

plants. Power peaking plants play an important but small role in the total production of

electric power. Figure 1 shows how the distribution of total electric generation fleet is

separated according to [2]. The data is also separated by either electric utility owned or

independent power producer (IPP) owned.

TOTAL U.S. ELECTRIC POWER GENERATION BY GENERATION RESOURCE, OCTOBER 2019

	Coal	Natural Gas	Nuclear	Hydroelectric	Oil Products	Wind	Geothermal	Solar Photovoltaic	All Other Resources
■ Electric Utilities	32.25%	34.97%	18.81%	10.90%	0.61%	1.92%	0.04%	0.30%	0.22%
■ Independent Power Producers	13.95%	40.75%	22.02%	1.27%	0.49%	14.74%	0.87%	4.05%	2.10%

ELECTRIC GENERATION RESOUCES

■ Electric Utilities   ■ Independent Power Producers
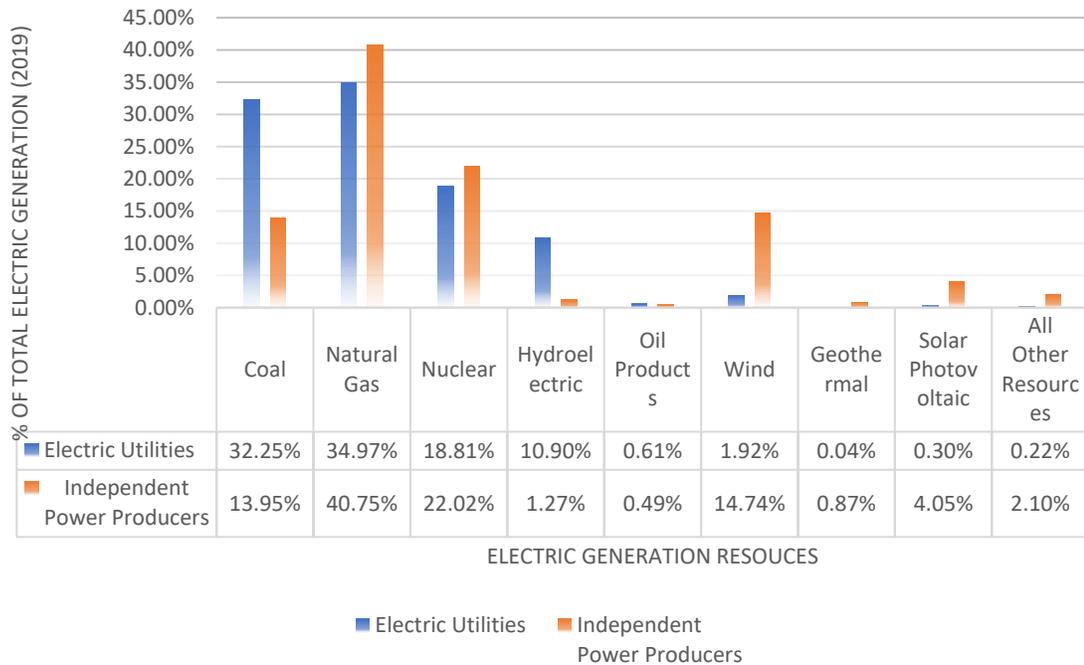
Figure 1 - Total U.S. Electric Power Generation by Generation Resources

Figure 1 proves that most of the total electric generation is produced by coal, natural gas, and nuclear power. The overall goal in the production of electricity is that it can be transported and consumed by the end user (customer) in a reliable and cost-effective manner. The transportation of the generated power is transported via the electric transmission system at higher voltages compared to the generation output or distribution level voltages. Transmission systems should be visualized as a cluster or mesh configuration of electrical connections, known as transmission lines or circuits, in a network arrangement that allows the power to flow from the generation sources to the

distribution system. The power that flows through the transmission system may not always flow in a single direction to the distribution system or other transmission customers. Power may flow in alternate routes to be consumed by the end user since the transmission system is typically designed as a networked or mesh system. Factors that can affect the power flow direction may include distributed generation, transmission contingency (electrical connection disconnected or out of service due to the occurrence of an abnormal condition) situation, schedule transmission element outages, scheduled transfers of power between multiple utilities, or the amount of generation dispatched in a geographical region versus other regions to serve system load requirements.

Transmission systems are designed to transport vast amounts of electrical power from one geographical region to another geographical region at higher voltage and lower current. Transmitting electricity at higher voltage and lower currents reduces the amount of power losses while allowing to send the power over many miles of transmission lines. Equation 1.1**Error! Reference source not found.** relates how the current flowing through a transmission conductor produces power losses.

$$\boldsymbol{Power\ Loss} = \boldsymbol{i^2} * \boldsymbol{R} \tag{1.1}$$

Since the square of the electric current is proportional to the power loss, then a reduction in electric current flowing through the transmission conductor will then produce a reduction of power loss. This allows utilities to maximize the amount of power that they supply to the end users by minimizing the amount of power losses that are lost by

4

transporting the electrical power. In order to send this electrical power over large distances, the voltage drop from the initial point of transmission to the end use of transmission needs to be minimized. Since the current in a transmission conductor is reduced to minimize power losses in the conductor, this process also allows voltage drops across the transmission lines to be reduced.

The electric distribution system, on the other hand, delivers the power from the transmission system to the end-user. The cutoff from the transmission system to the distribution system is mostly decided by equipment in the distribution substations. There is usually a type of substation equipment (distribution transformer, substation bus, distribution feeder breakers, etc.) that will determine this cut off point and it will be vary from utility to utility. These distribution systems are normally designed as radial systems and operate at lower voltages with higher current. But it should be stated that some distribution systems are not always operated in a radial design. It is important to understand that faults on any of the components of the power system are crucial and suspectable to faults. This research will only be focusing on the effects that faults have on the electric transmission system.

1.2   Electric Transmission System Overview

Transmission lines are typically classified by their operational voltage levels and total line length in miles or kilometers (km). In the United States the length of the transmission lines is typically expressed in miles and can be operated in either alternating current (AC)

and direct current (DC) configurations. AC transmission lines are the dominant

configuration within the power system grid and will be the focus of this research. The AC

transmission voltage levels vary throughout the United States but will range from 100 kilovolts

(kV) up to 765 kV. Sub-transmission voltage levels will range from 34.5 kV up to 100

kV. Many sectors of the utility industry are starting to classify 34.5 kV as a distribution

voltage. Table 1 provides an overview of transmission line operation voltage levels with

their associated transmission level classifications [3].

Table 1 - Transmission Voltage Level Based on Transmission Classification

Transmission Line Classification	Voltage Range (kV)	Purpose
Ultra-High Voltage (UHV)	> 765	High Voltage Transmission > 765 kV
Extra-High Voltage (EHV)	345, 500, 765	High Voltage Transmission
High Voltage (HV)	115, 138, 161, 230	High Voltage Transmission
Medium Voltage (MV)	34, 46, 69	Sub-transmission
Low Voltage (LV)	< 34	Distribution for residential or small commercial customers, and utilities

The North American Electric Reliability Corporation (NERC) uses the term bulk electric

system (BES) in their reliability standards to categorize the voltage levels of any

electrical transmission element that is operated at 100 kV and above [4]. BES voltage

levels are divided into two different categories: high voltage (HV) transmission elements

and extra high voltage (EHV) transmission elements. HV transmission elements are

defined to operate on the range of 100 kV to 300 kV where the EHV transmission

elements operate in the range of 300 kV and greater. Transmission lines are typically

supported by steel or wooden structures (also known as towers). These structures are built

in forms of lattice steel structures, wooden, or steel poles. The intent of these transmission structures supports the weight of the transmission lines while withstanding harsh weather conditions. The design specifications of these structures are built to comply with the National Electric Safety Code (NESC) [5]. Most of the time transmission towers, especially in rural areas, support only one transmission line, but there are cases where these towers need to support two or more circuits of conductors. When transmission towers support two or more circuits from one substation to another or located within close proximity to each other, the transmission circuits are known to have the same right of way easement. These transmission configurations are known as parallel transmission line configurations. Of course, circuits that run from one substation to another on the same right of way easement are the most basic representation of parallel line configurations. It is very common for other transmission circuits to only be part of an existing transmission line right of way for a portion of the transmission line distance before diverting into a different direction to different substations. Parallel configurations can consist of lines operating at the same voltage or different voltage levels as well as power flowing in the same or opposite directions.

There are two major identifiable violations or unwanted conditions on the electric power system. These conditions are known as low or high voltage violations and thermal (conductor overload) violations. Low voltage violations are real-time voltage measurements that occur either pre or post contingency where the voltage measurement falls below a specific threshold or value. Likewise, high voltage violations are real-time voltage measurements that occur either pre or post contingency where the voltage measurement is above a specific threshold or value. NERC requires in the TPL-001-4, a

NERC Reliability Standard, that each entity that is registered as a Transmission Planner (TP) or Planning Coordinator (PC) shall have a criteria for acceptable steady state voltage limits [6]. There is no single limit within the TPL-001-4 reliability standard that identifies these low and high voltage violation limits. The second identifiable violation is thermal violations. Each conductor used in transmission line design has specifications that allows a maximum amount of current or power flow to flow through the transmission conductor to ensure that the conductor does not experience the risk of any damage. This power flow can be expressed in terms of either electrical current (measured in unit of amperes (A)) or power-carrying capacity (measured in units of megawatts (MW) or megavolt-ampere (MVA)). Thermal transmission line ratings (or capacity) are generally negatively correlated to the ambient temperature and solar irradiance intensity, but positively correlated with wind speeds [7]. This means that the colder the ambient temperature around the transmission conductors the higher the thermal capacity and the hotter the ambient temperature the lower the thermal capacity for the transmission line.

There are many factors that have been mentioned that can alter power flows through the transmission system. Related to this research, it become important to understand how power flows are altered due to transmission contingency scenarios. When a transmission line experiences a fault or contingency, the power flowing on that transmission line is shifted to another nearby network transmission line(s) connected to the transmission system. If the transmission line is experiencing an contingency situation has the basic task of transmitting power to nearby customer loads and provides only limited amounts of through flow power on that transmission line, then any resultant overload violations will possibly stay local to that geographic area. But if the transmission contingency is a

8

related to a higher-level voltage transmission line that serves the purpose of transmitting electricity to other geographical regions (higher levels of through flow power), then the resultant overload violations that could possibly be created in other transmission elements may be more widespread.

One tool that is used to analyze transmission lines overloads due to another transmission conditions are called "Linear Sensitivity Factors". At a basic level there are two sensitivity factors that are known as power transfer distribution factor (PTDF) and line outage distribution factor (LODF). The PTDF represents the sensitivity of power flow on a transmission line from a shift of power generation from one generator to another. One of the factors that can cause power flows to shift in different directions is a shift in the amount of generation in one area versus another. The LODF sensitivity factor tests for overloads on a transmission circuit when another transmission line has been taken out of service due to a fault on a transmission line or a transmission element malfunction. The LODF will be most relevant to this research and is calculated using equation 1.2 [8].

$$\boldsymbol{LODF_{l,k}} = \frac{\Delta f_l}{f_k^0} \tag{1.2}$$

where:

- $LODF_{l,k}$ is the line outage distribution factor when monitoring line "l" with an outage of line "k".

- Δf_l is the change of MW flow online for line "l".

- $\boldsymbol{f_k^0}$ is the original power flow on line "k" before it was removed from service.

1.3   Electric Transmission System Line Configurations

The transmission system is an important and major component of the electric power

system that is designed to transport electrical power in bulk amounts over large distances

organized within a cluster of networked electrical configurations known as transmission

lines. These networked configurations can consist of radial, single line, and parallel line

configurations, or a variety of different type of configurations that make up the original

networked system as whole. This section will discuss a few important transmission line

configurations in which some of these configurations are used within this research.

1.3.1   Single Terminal Radial Transmission Lines

The first transmission line configuration that will be discussed is the radial transmission

line. There are segments of the transmission system that have end users of electric power

on radial feeds. Radial feeds are transmission lines that are supported by only one

electrical source. The issue with end users that are feed by radial feeds is if an

interruption of power flow from the single source occurs, then no power can flow through

that radial feed to that end user which results in the loss of electricity. Radial feed

configurations are known to have lower reliability since they only have one source

available. These types of electrical transmission lines are very common when feeding

distribution substations in rural areas. Typically, this configuration operates at lower

voltages such as 69 kV transmission level voltages, but they can be used to serve higher

level voltages customers as well. Figure 2 represents a visual representation of a radial

transmission circuit. The AC generator connected to substation A indicates the idea that

radial transmission line has only one source supporting the flow of electricity to the end

users. This transmission topology will not be studied in this dissertation, since networked

transmission lines are the focus.

Figure 2 - One Line Diagram of Transmission Radial Line

1.3.2   Two-Terminal Single Transmission Lines

The second transmission line configuration that is presented is the two-terminal single

transmission line. The two-terminal single transmission line is an example of an electric

transmission line or circuit that travels from one transmission substation to another

without any opportunity for power to divert in a different direction. These transmission

lines normally transmit power between different substations within a networked

configuration. It is extremely common to see transmission breakers in-line with the

transmission line at each connected substation. These breakers provide protection for the

transmission line, which have the task of isolating any fault or abnormal condition that suddenly occurs on the transmission line. This transmission line configuration and the radial transmission line configuration are probably the simplest networked transmission line configurations that protection engineers must provide protection for. In the case of a two-terminal transmission line, power may flow in one or both directions depending on its location and system conditions. The reason for power flow in both directions is because the two-terminal single transmission line is part of a networked configuration that provides power support from both ends of the transmission line. Depending on situational power flows such as load forecast, scheduled or forced outages, scheduled power transfers, and generation profiles power may flow in different directions. Figure 3 shows a one-line representation of the two-terminal single transmission line.

Figure 3 - One Line Diagram of Two-Terminal Single Transmission Line

Two out of the three phases of this research will be utilizing the two-terminal single transmission line to predict the electrical fault identification (fault classification and fault location). Measurement configurations around the transmission lines may occur in a variety of different arrangements. Utilities may have installed potential transformers

(PT's) and/or current transformers (CT's) at both substations that measure and record voltage and current measurements. This research will be using different arrangements of these electrical quantity measurements to predict fault identification. An example of a measurement arrangement would be voltage or current measurements only being available from one substation.

1.3.3   Two-Terminal Parallel Transmission Lines

The third configuration that is presented would be the two-terminal parallel transmission line. This configuration is a topology that extends the idea of the two-terminal single transmission line that parallels multiple circuits. This transmission configuration can be visualized as two or more different transmission lines or transmission circuits sharing a common transmission structure or two or more separate transmission circuits that run beside each other in a single right-of-way easement where mutual coupling is shared between the two circuits. This configuration can cause issues with protection schemes, especially during a faulted condition due to induced currents from magnetic fields caused by mutual coupling. Since these transmission circuits are mutual coupled with each other, a faulted condition on one circuit that contains high fault currents can cause the fault current to be induced into the healthy circuit(s) causing the protection scheme on the healthy circuit(s) to operate pre-maturely. Figure 4 shows a visual representation of the parallel transmission line configuration [9].

Figure 4 - One Line Diagram of Two-Terminal Parallel Transmission Line

To illustrate how the fault current is induced from the faulted transmission circuit to the heathy non-faulted transmission circuit an illustration from the 2019a version of MATLAB and Simulink software is shown in Figure 5 and Figure 6. These figures represent a Simulink simulated three-phase current waveforms recorded from one substation of a parallel transmission line configuration. During this simulation, a 5 Ω, phase A to ground (A-G) fault was applied to one of the transmission lines (circuit #1) at 10 kilometers (km) away from substation A of a 100 km transmission line. The fault was applied to the transmission line at 0.0333 seconds (2 cycles) into the simulation. Figure 5 shows that in the faulted circuit (circuit #1) the fault current in phase A increases from nearly 5 per unit (pu) to nearly 77 pu at 1 cycle after the fault. Once the DC offset settles the phase A waveform amplitude settles to nearly 65 pu. This is the result that is expected, a large increase in the phase A current, since the phase A to ground fault is being simulated. It's the result in the other transmission line (circuit #2) that has the interesting effect (Figure 6). The non-faulted transmission line current in phase A increases from nearly 5 pu to around 13 pu. Also, Figure 6 shows that the phase C current waveform amplitude increases from nearly 5 pu to around 9 per unit. This increase in current amplitude may be a large enough increase to trigger the non-faulted transmission

circuit (circuit #2) breakers to trip based on the designed protection scheme, if the

protection scheme is not designed for mutual coupling effects.
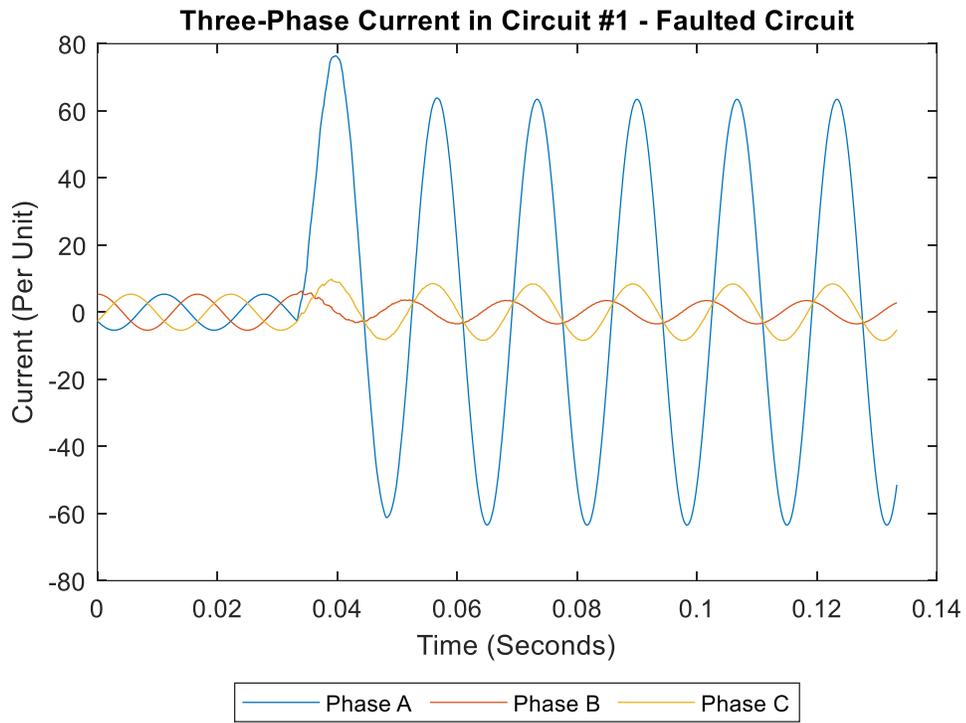
Figure 5 - Three-Phase Fault Current on Faulted Transmission Line with Mutual
Coupling
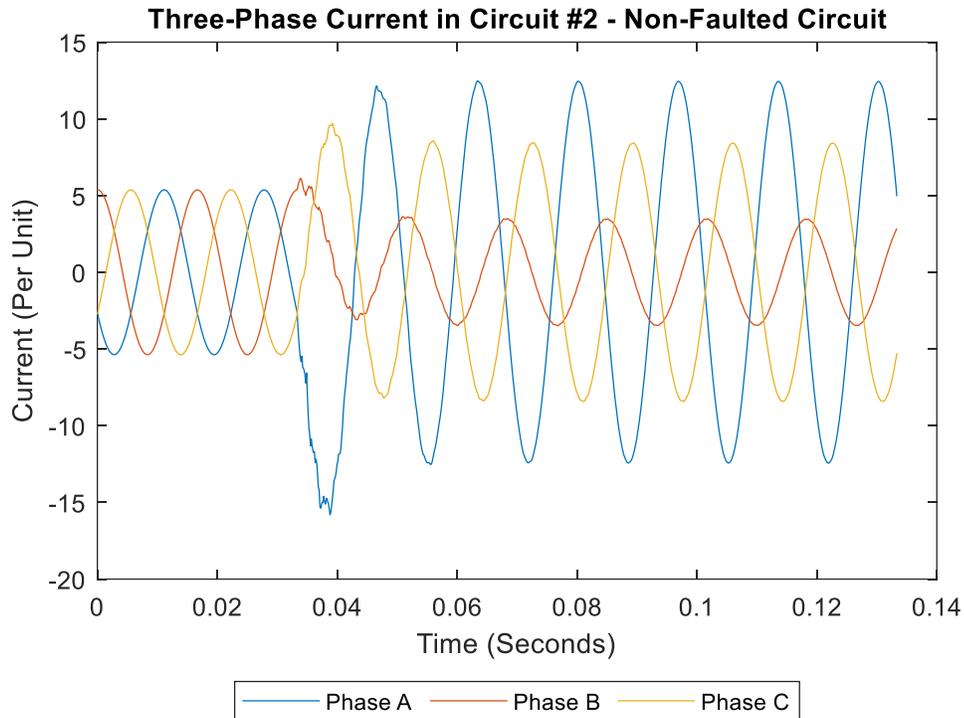
Figure 6 - Three-Phase Fault Current on Non-Faulted Transmission Line with Mutual

Coupling

1.3.4   Multiple Terminal Transmission Lines

The last transmission line configuration that is common within the transmission system is

called the multi-terminal or teed transmission line. This situation originates from the two-

terminal single transmission line which is tapped to provide electrical power to a different

geographical region or to provide power to another substation for new load or reliability

requirements. Most of the protection scheme issues with a multi-terminal transmission

line is determining which transmission line segment the electrical fault is physically

16

located at near the multi-terminal connection point. Figure 7 shows a visual

representation of the multi-terminal transmission line configuration.
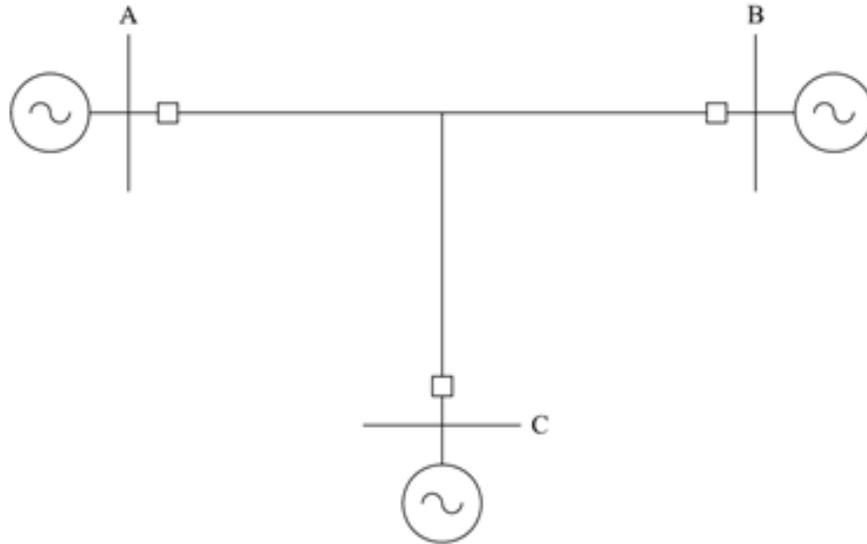
Figure 7 - One Line Diagram of Teed or Three Terminal Transmission Line

As seen in all of the transmission configuration one line diagrams (Figure 2, Figure 3,

Figure 4, and Figure 7) the square boxes adjacent to each bus are representations of

breakers protecting each transmission line. For the multi-terminal transmission line, it

should be noted that there is no protection equipment at the tapped connection point. This

creates the issue of determining the fault identification for the multi-terminal

transmission line. This research does not focus on the multi-terminal transmission line

configuration to identify fault classification and fault location. But this topology most

definitely should be studied in future research.

1.4   Research Purpose Statement

Electric faults on transmission lines are inevitable due to the nature of the system. Detecting faults and restoring the transmission system to its original state can be a time and labor-intensive process where every second counts to prevent further damage. Detecting these faults can become more crucial during system peak conditions. Fault location tools readily available today only exist for the simple two-terminal transmission lines and provide general distance to fault estimates. Performance of these tools is limited and can vary as other transmission line configurations are evaluated. A seamless, automated fault identification and analysis tool is needed to improve the fault location response for complex line topologies such as parallel transmission lines where fault measurement data may be limited. There has been a lack of guidelines on designing the architecture for artificial neural network structures including the number of hidden layers and the number of neurons in each hidden layer. This research will fill this gap by providing insights on choosing effective neural network structures for fault classification and location applications.

1.5   Dissertation Outline

Up to this point, an introduction into the basics of what components make up the power system have been discussed. Most of the attention has been dedicated to the transmission system since this research will be focused on the transmission system. Chapter 2 will continue the discussions by giving a brief introduction to some transmission line

18

characteristics as it relates to transmission lines being vulnerable to faulted conditions.

This research will concentrate on predicting where the faulted condition has occurred,

therefore its best to understand how these faults occur and how often they occur.

Following this introduction of transmission vulnerability to faults, the different types of

fault classifications that can occur on the transmission line will be presented. These fault

classifications will be discussed in detail and describe how these faulted situations may

occur. Chapter 2 will then introduce the intelligent technique of artificial neural network

(ANN) that is used within this research. After providing the ANN overview, some related

work that has occurred as related to the transmission system fault identification problem

using ANNs will be discussed.

This research was completed in three phases. The first phase of the research is presented

within Chapter 3. Chapter 3 begins by describing the two-terminal single transmission

line model that was developed within the 2016a version of MATLAB and Simulink

software. All sections within Chapter 3 describe how the training input and target data

was obtained to begin training the different ANN architectures, and how the ANNs were

tested with the MATLAB and Simulink model testing data. This testing data was

collected so that faulted measurement data was different then the training data that was

used to train the ANNs. Chapter 3 concludes by providing results on how the different

ANN structures predicted transmission fault identification as it relates to using a single

ANN to predict fault classification and fault location together. Chapter 4 is basically a

repeat of Chapter 3 with the exception that multiple ANN were used to predict fault

identification. Chapter 4 proposes an approach that uses one ANN to predict fault

classification and then uses a set of four different ANNs to predict the fault location.

These four different fault location ANNs will correspond to the four basic fault types. Chapter 5 will then finalize the last phase of the research by introducing the parallel transmission line topology. This chapter will use the same approach used in Chapter 4 but will be expanded for the use of the second transmission line. Chapter 6 concludes this dissertation by recapping the conclusions made in the three phases of this research.

Chapter 2 – Background and Related Work

Electric transmission lines transport electrical power for miles throughout the utility scale power system. These transmission lines can range in length from tenths of a mile up to hundreds of miles in length. The United States Department of Energy (DOE) reviews public sources of national information to collect information related to the transmission grid. These public sources are published by the Energy Information Administration (EIA), Edison Electric Institute (EEI), the North American Electric Reliability Corporation (NERC), and the Federal Energy Regulatory Commission (FERC) [10]. Included in reference [10] and presented in Table 2, published in March of 2018, the United States transmission grid consisted of the reported number of transmission line miles for each voltage range at the end of 2016.

Table 2 - Number of Transmission Line Miles in the United States

Miles of Transmission Line in the United States (100 kV and Above)								
Voltage Class	FRCC	MRO	NPCC	RFC	SERC	SPP	TRE	WECC	Total Miles
Total DC	0	1802	26	0	0	0	0	2142	3970
600 kV - 799 kV	0	0	190	2201	0	0	0	0	2391
400 kV - 599 kV	1201	139	0	2431	9093	94	0	13826	26784
300 kV - 399 kV	0	8542	5580	13650	3868	6653	14838	10673	63804
200 kV - 299 kV	6203	7501	1612	6862	22828	3224	0	38167	86397
100 kV - 199 kV	3956	21933	13304	32683	60916	19365	20818	38252	211227
Total Miles by NERC Region	11360	39917	20712	57827	96705	29336	35656	103060	394573
Entity Count	15	25	18	27	30	20	26	61	

The circuit miles as presented provide great insight to the amount of transmission line miles that are used transport power across the United States. Table 2 clearly explain how

these transmission lines are operated geographically throughout the United States if the NERC regions are known geographically. Maps of these NERC regions can be located on any of the public source websites that DOE utilizes to support their Annual U.S. Transmission Data Review. The reported NERC regions are Florida Reliability Coordinating Council (FRCC), Midwest Reliability Organization (MRO), Northeast Power Coordinating Council (NPCC), Reliability First Corporations (RFC), SERC Reliability Corporation (SERC), Southwest Power Pool (SPP), Texas Reliability Entity (TRE), and Western Electricity Coordinating Council (WECC). Figure 8 presents the existing transmission line miles located within the United States separated by NERC Regions. Figure 8, is a graphical representation of the data presented in Table 2 to make it easier to define how the different NERC regions operate transmission lines located within the geographical areas.
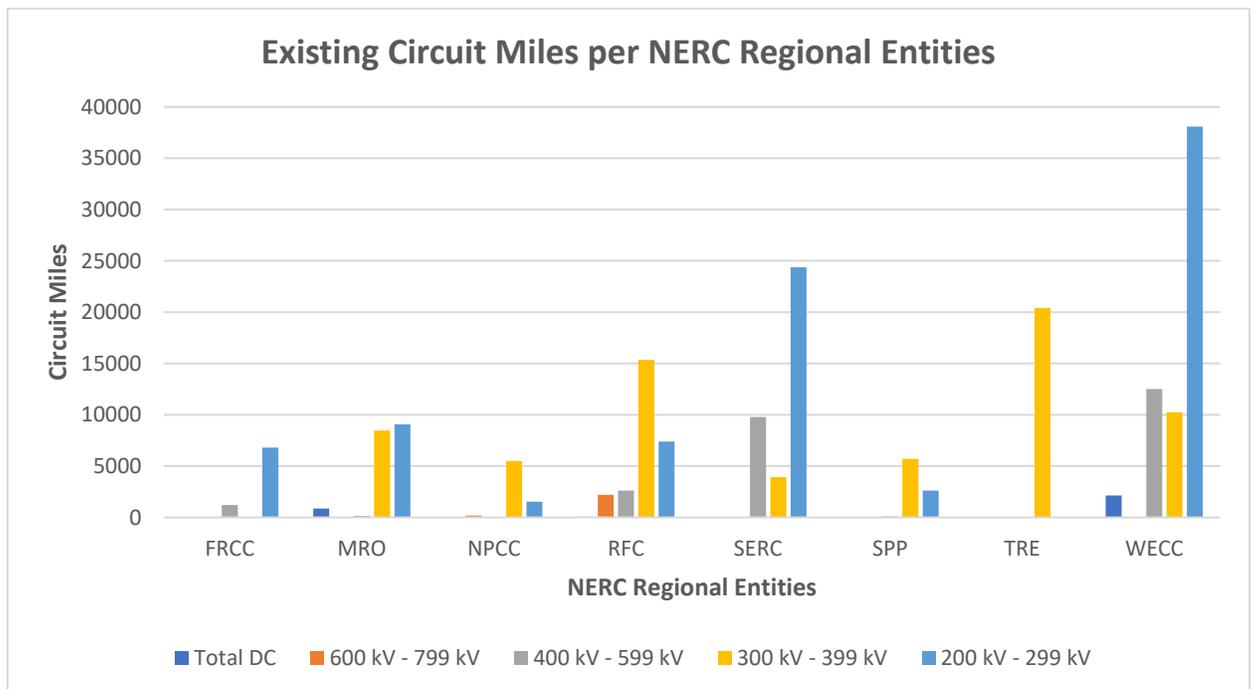
Figure 8 - Existing Transmission Circuit Miles in the United States

The information in Table 2 and Figure 8, was extracted from the NERC Transmission

Availability Data System (TADS) database. The TADS database contain data that is

collected quarterly on existing transmission equipment inventory and outage frequency

experienced by the different transmission equipment. This data is voluntarily provided by

transmission owners (TO) and is reviewed by the eight NERC regional entities. The

collected data is categorized by voltage class and only contains information related to the

transmission infrastructure that is operated at 100 kV and above [10]. It should be noted

that there are many transmission facilities that operate at voltage levels less than 100 kV

which are not reported in the TADS database. Table 2 and Figure 8 demonstrates that

there are over 394,000 miles of overhead transmission lines that support the

transportation of electric power in the United States alone. This does include both high

voltage AC and high voltage DC operated transmission facilities. As electrical load

continues to grow throughout the United States, the design and installation of the United

States transmission system will continue to grow to keep up with the demand. The NERC

Electricity Supply & Demand (ES&D) database, is a database that contains information

on existing and planned transmission facilities that will operate at voltages of 100 kV and

above. For the planned portion of the data, the ES&D database provides transmission

assets that are under construction, planned, or under conceptual development. This

information is provided in Figure 9 and Figure 10 to provide additional insights on the

amount of new transmission line miles that will be operated in the United States. Just

evaluating the total amount of circuit miles that are planned to be built by the year of

2020 to 2025, will add up to an additional 14,117 circuit miles (2,852 miles: Under

Construction and 11,265 miles: Planned). Most of the planned construction that include

transmission lines are to be built by the end of 2020. Transmission assets that are only

planned and no construction has taken place have to option to withdraw the planned

project. This would reduce the number of miles for future planned transmission lines.
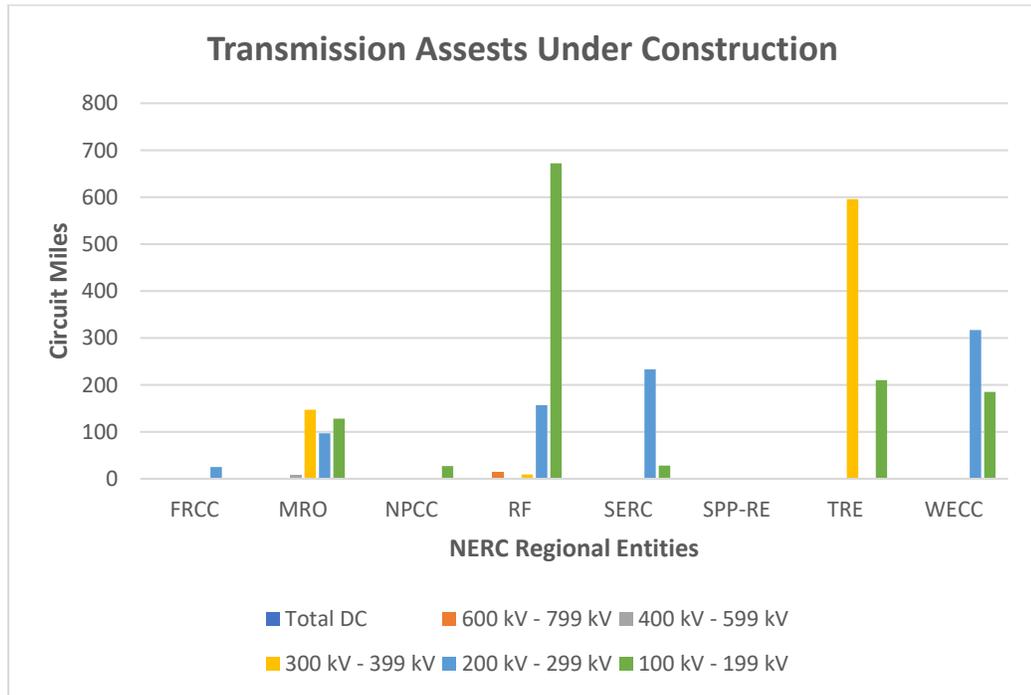
Figure 9 - Transmission Assets Under Construction
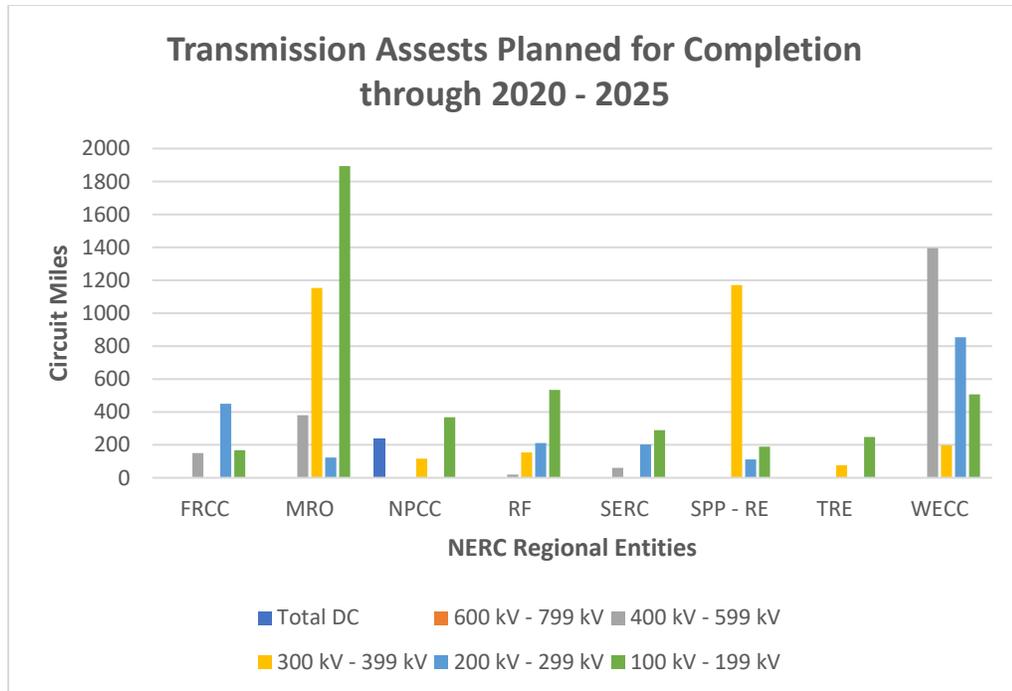
Figure 10 - Transmission Assets Planned for Completion through 2020 - 2025

Looking at a future perspective (through 2025) for transmission lines operated within the
United States, the data shows that 408,690 miles of transmission lines will be in service
operating at 100 kV and above.

2.1   Electric Transmission Power System Faults

With substantial miles of overhead transmission lines being operated throughout the
United States, transmission lines are deliberately exposed to a variety of potential
external events. These events can create abnormal or faulted condition on these active
transmission lines. With such large distances of exposure to transmission lines, it is

inevitable that electrical transmission line faults are going to occur, and it is just a matter of when these faults are going to occur. These faults can originate from many sources including weather, natural disasters events, animals, or from human interaction to name a few. Table 3, published by NERC, defines the categories of different causes of transmission line faults and how frequent these electrical faults have occurred between 2012 and 2016 [10].

Table 3 - Transmission Line Fault Cause Codes and Outage Frequency

TADS Transmission Line Fault Cause Code and Outage Frequency						
Initiating Cause Code	2012	2013	2014	2015	2016	2012 - 2016
Lightning	852	813	709	783	733	3890
Unknown	710	712	779	830	773	3804
Weather Excluding Lightning	446	433	441	498	638	2456
Failed AC Circuit Equipment	261	248	224	255	362	1350
Miss Operation	321	281	314	165	249	1330
Failed AC Substation Equipment	248	191	223	221	214	1097
Foreign Interference	170	181	226	274	258	1109
Contamination	160	151	149	154	289	903
Human Error	212	191	149	132	153	837
Power System Condition	77	109	83	96	81	446
Fire	106	130	44	65	72	417
Other	104	64	77	77	78	400
Combined Smaller ICC Groups Study 1-3	57	53	49	37	47	243
Vegetation	43	36	39	32	34	184
Vandalism, Terrorism, or Malicious Acts	10	9	8	1	7	35
Environmental	4	8	2	4	6	24
All with ICC Assigned	3724	3557	3467	3587	3947	18282
All TADS Events	3753	3557	3477	3587	3947	18321

Most electrical faults come from weather related events with the majority of those being due to lightning or an unknown cause. An electric transmission fault is defined as an abnormal condition that has the opportunity to occur on the electrical power system that interferes with the normal flow of electrical current [11]. Faults can be classified as either temporary or permanent. Temporary faults that occur on the transmission system are only sustained for a short period of time. This fault category is known to clear the fault itself. An example of a temporary fault is a tree limb falling on a transmission line that causes the faulted condition and then after the contact between the current carrying or grounded conductor(s) and the tree limb occur the tree limb falls off the transmission line. This results in the fault clearing itself from the transmission system. Permanent faults are abnormal conditions that occur on the transmission system where the condition cannot be cleared or removed on its own. An example of a permanent fault would be a current carrying conductor breaking in mid-span between two transmission towers and the conductor contacting a transmission structure that is grounded. This would cause a sustained line to ground fault that would require physical assistance to remove the conductor contact from the transmission tower. Abnormal flows of electrical current can flow between conductors to ground, between multiple conductors, or between multiple conductors and the ground. How electrical current is flowing during these faulted conditions defines the fault classifications (also known as the fault types). These fault classifications that the electrical power system can experience define the faults that are studied in this research. To the electrical utility industry, the different fault classifications are known as single line-to-ground faults, double line (line-to-line) faults, and double line (line-to-line) to ground faults, and three-phase faults. The three-phase fault is the only

fault that is known to be a symmetrical fault. Where on the other hand, the single line-to-ground, the double line, and the double line to ground fault are classified as asymmetrical faults. According to reference [11], most faults on transmission systems at voltages of 115 kV or higher are caused by lightning which results in flashover of the insulators. Experience has shown that 70% to 80% of transmission line faults result into single line to ground faults. Where roughly only 5% of all transmission faults involve all three phases [11].

2.1.1 Line to Ground Faults

The line to ground fault is the most common electrical fault that occurs on the transmission system. Each transmission line is composed of three current carrying conductors and a static or ground conductor wire that is grounded at nearly every transmission structure. This type of grounding system is known as the multi-grounded system. The three current carrying conductors are mostly classified as phases and contain the labels of phase A, phase B, and phase C. Which phases that are classified as phase A, phase B, or phase C are arbitrary if the phase designation is keep consistent. A line to ground fault is considered an abnormal condition that contacts one of the three current carrying conductors to a physical element of the transmission system that operates at a zero-voltage potential. Figure 11 provides a visual representation of a hypothetical point on a transmission line in which a phase A to ground fault has occurred. The Z_f fault impedance represents the fault impedance through the current carrying conductor to

grounded equipment. This Z_f impedance value can vary depending on the physical

condition that is causing the fault.
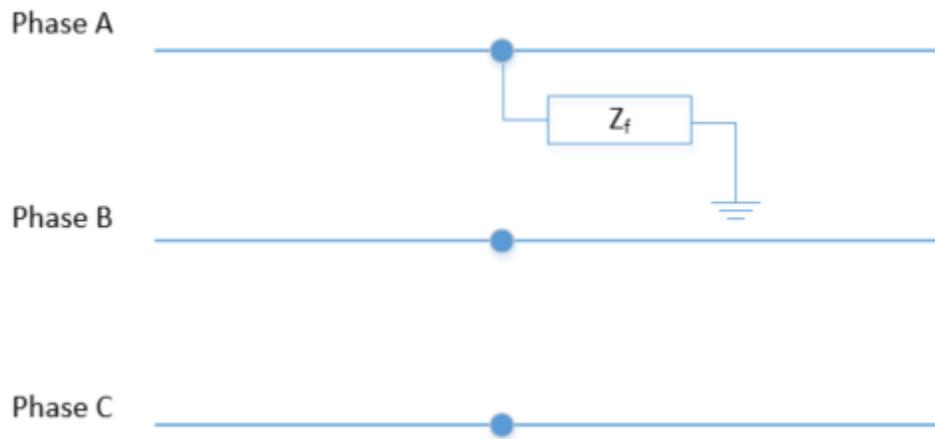
Figure 11 - Line to Ground Fault with Z_f Fault Impedance

For a complete and detail derivation of line to ground faulted conditions, it is encouraged

that the reader of this dissertation should review reference [11]. In order to follow this

derivation or any unsymmetrical fault, an understanding of symmetrical components is

needed.

2.1.2 Line to Line Faults

Line to line faults are faulted conditions that encompass connections between two of the

current carrying conductors of the transmission line. The possible faulted classifications

for these types of faults would consist of abnormal conditions between any two of the

three current carrying conductors: phase A to phase B, phase A to phase C, or phase B to

phase C. These line fault classifications are considered and analyzed within this research

dissertation. Figure 12 provides a visual representation of a hypothetical point on a

transmission line in which a phase B to phase C line to line fault has occurred. The Z_f

fault impedance represents the impedance of the line to line contact. This Z_f impedance

value can vary depending on the physical condition causing the fault.
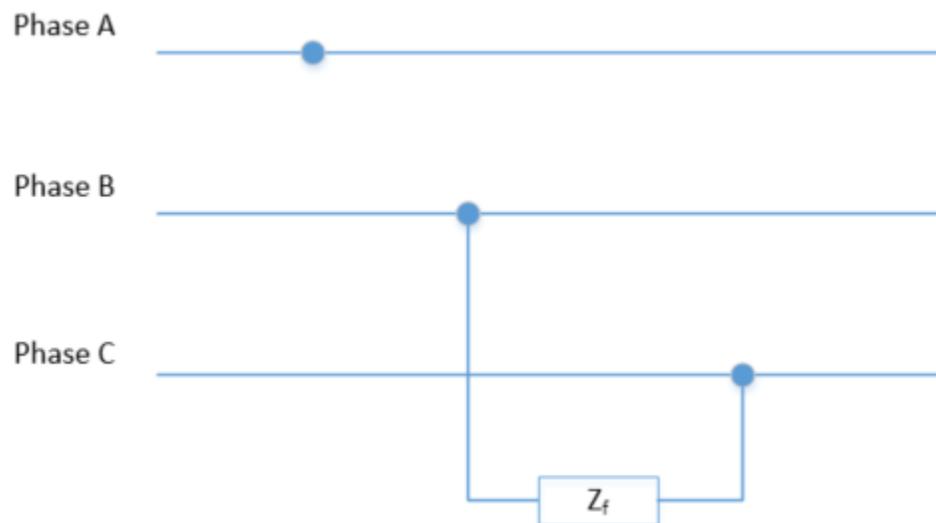
Figure 12 - Line to Line Fault with Z_f Fault Impedance

For a complete and detail derivation of line to line faulted conditions, it is encouraged that the reader of this dissertation should review reference [11].

2.1.3 Double Line to Ground (Earth) Faults

Double line to ground faults are faulted conditions that encompass connections between two of the current carrying conductors and a grounding connection of the transmission system. The possible faulted classifications for these types of faults would consist of one of the following three arrangements:

- Phase A – Phase B – Ground

- Phase A – Phase C – Ground

- Phase B – Phase C – Ground

These line fault classifications are considered and analyzed within this research dissertation. Figure 13 provides a visual representation of a hypothetical point on a transmission line in which a phase B to phase C to ground fault, double line to ground fault, has occurred. The Z_f fault impedance represents the impedance of the double line to ground contact. Since the fault current for this faulted condition would flow through the current carrying conductors and then through ground, the total faulted impedance is shown on the grounding connection. This Z_f impedance value can vary depending on the physical condition causing the fault.
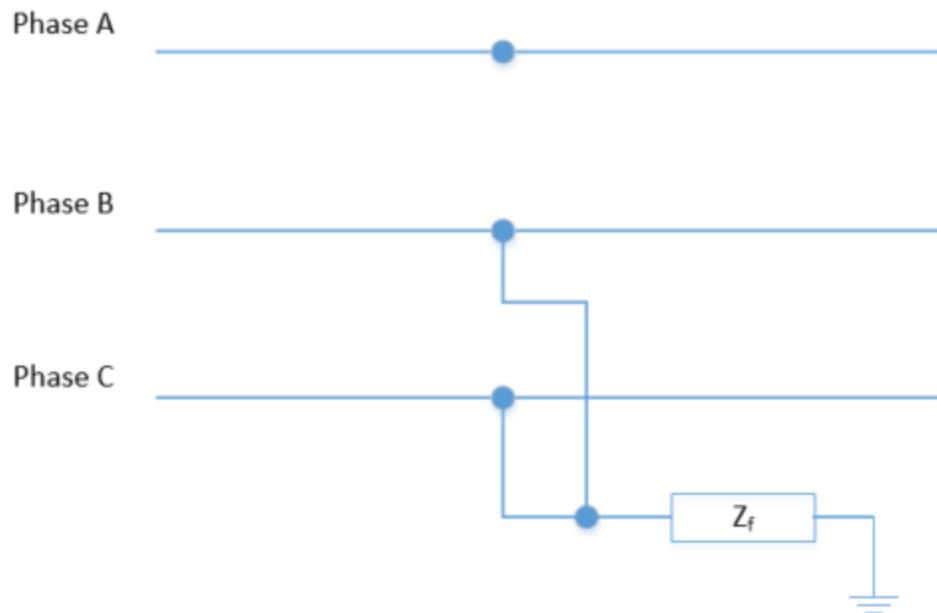
Figure 13 - Line to Line to Ground Fault with Z_f Fault Impedance

For a complete and detailed derivation of line to line to ground faulted conditions, it is encouraged that the reader of this dissertation should review reference [11].

2.1.4 Symmetrical Three-Phase Faults

The three-phase fault is a last fault type that will be studied within this dissertation. This fault type is also the rarest of all faulted conditions to occur. Three-phase fault conditions occur when all three current carrying conductors have become in contact with each other. The only possible faulted combination that can happen on a three-phase system is when phases A, B, and C comes in contact with each other. These faults, as with the other three

fault classifications we have discussed previously, will contain some amount of fault impedance. The fault impedance, Z_f, within the three-phase fault condition is modeled such that the fault current in each phase must flow through the fault impedance within each phase. Figure 14 provides a visual representation of the three-phase fault classification.
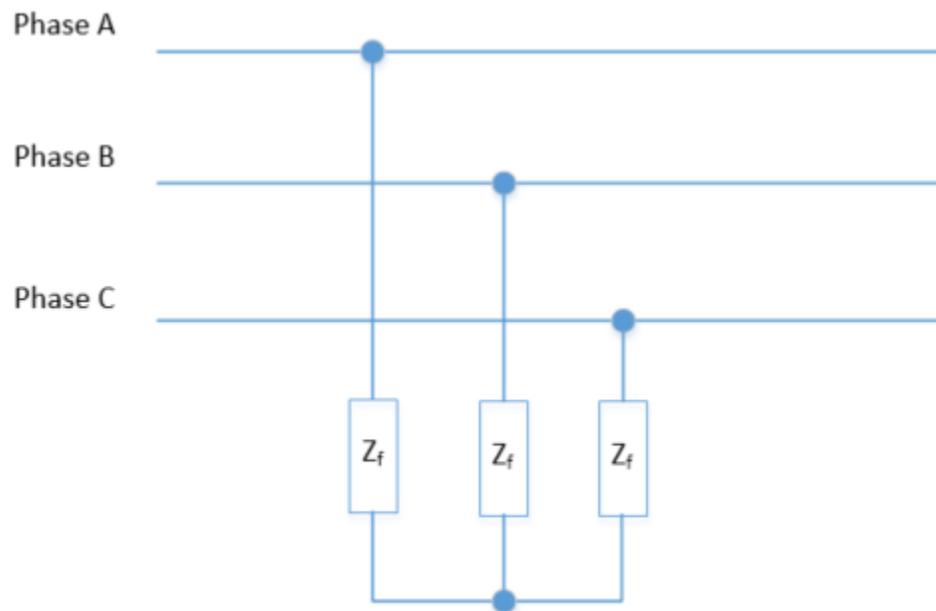
Figure 14 - Three-Phase Fault with Z_f Fault Impedance

For a complete and detail derivation of three-phase faulted condition, it is encouraged that the reader of this dissertation should review reference [11].

2.1.5 Open Conductor Faults

There is one other type of fault that can occur which is the open conductor fault. The past four fault classifications that have been discussed are shunt fault types. The open conductor fault is a series fault type. During this faulted condition, an open circuit occurs in one or more phases of the transmission circuit. This type of fault classification was not studied within this research. Figure 15 provides a diagram of an open conductor fault example. This example shows that phase A has become an open circuit while phases B and C remain intact.

Figure 15 - Open Conductor Fault

For a complete and detail derivation of open conductor fault condition, it is encouraged that the reader of this dissertation should review reference [11].

2.2   Artificial Neural Network Overview

This research uses an intelligent technique based on artificial neural networks (ANN). This section will provide an overview of a type of ANN, feed forward neural networks, and describe how feed forward networks will be used to solve the fault identification problem. This dissertation so far has provided a lot of discussion on the types of vulnerable issues that the transmission system is exposed to and some statistics that describe how the exposure of the transmission system can create unwanted power flow conditions. This research assumes that a transmission fault has occurred. This assumption continues by assuming that all protection devices that are designed to protect that specific transmission line has operated to isolate the fault from the rest of the transmission system. This may mean that the transmission operators have segmented the system even further than normal system protection to isolate the fault from the transmission system while restoring the maximum amount of customer loads as possible.

Transmission protection engineers have many different types of line configurations that they must protect transmission elements from when it comes to electrical faults. One of the sources that protection engineers use to protect the transmission system is the use of relays and breakers. There are multiple relay types and protection schemes that can be used to help protect transmission systems against transmission faults. Not every relay type is acceptable to be used on any transmission line configuration. Table 4 provides a list of some examples of protective relay functions that could be implemented to provide protection to transmission lines [12]. Keep in mind that this is only a subset of the full list of protective relay functions that protection engineers have at their fingertips. Reference

[12] has a full list of the protective relays that protection engineers can use for protecting transmission lines.

Table 4 - Standard Protection Relay Functions (IEEE/ANSI C37.2 Standard)

Standard Protection Relay Functions (IEEE/ANSI C37.2)	
Relay Function	Device/Function Number
Distance Relay - A device that functions when the circuit admittance, impedance, or reactance increases or decreases beyond a predetermined value	21
Directional Power Relay - A device that operates on a predetermined value of power flow in a given direction	32
Instantaneous Overcurrent Relay - A device that operates with no intentional time delay when the current exceeds a preset value	50
Instantaneous Overcurrent Relay with Time Delay	50TD
AC Directional Overcurrent Relay - A device that functions at a desired value of AC overcurrent flowing in a predetermined direction	67
Differential Protective Relay - A device that operates on a percentage, phase angle, or other quantitative difference of two or more currents or other electrical quantities	87

With all these relay types and protection schemes, it is not always easy to point to exactly where the fault is located along the transmission line. Some of the industry may have developed good practices over the years to get a general area of where the actual fault has occurred. But this devotes time and resources by reviewing data and breaker operations to

determine the location of the transmission line fault before sending field personnel out to fix any repairs.

The artificial neural network is a concept that is related to the idea behind the operation of biological neural networks or how the human brain functions. As discussed in references [13] and [14], the human brain consists of large numbers of interconnected elements known as neurons. Through life experiences and lessons learned these biological neurons will adjust and allows human recognition to occur. In simplified terms, neurons consist of three primary components: the dendrites, the cell body, and the axon. The dendrites will carry electrical signals into the cell body. Where the cell body will then sum the electrical signals and threshold the incoming electrical signals from the dendrites. Finally, these modified electrical signals will flow out of the cell body and into the axon so that the signals can be transported to other networked neurons. Another important function of the biological neurons is the point where the axon of one neuron meets another neuron. This connection point is known as the synapse. As it will be seen shortly that the artificial neuron network model will contain weighted inputs. The weighted inputs hold a similar function as the synapse [13].

2.2.1   Artificial Neural Network – Multi-Input Single-Neuron Models

Artificial neural network architectures are developed in different categories of structures. This research will be utilizing the multi-layer feed-forward ANN architecture. Before discussing the full design parameters of the feed-forward neural networks used within

this research some introduction into feed-forward architectures should be presented. It is

best to begin the introduction into neural network design by discussing the simplest ANN

architecture, the single-neuron model. Single neuron models can be introduced with

either a single input or multiple input characteristic. Most available references that

discuss feed forward neural networks will present both single-input and multiple-input

networks. Figure 16 presents an example of a single neuron model with multiple data

inputs [13].
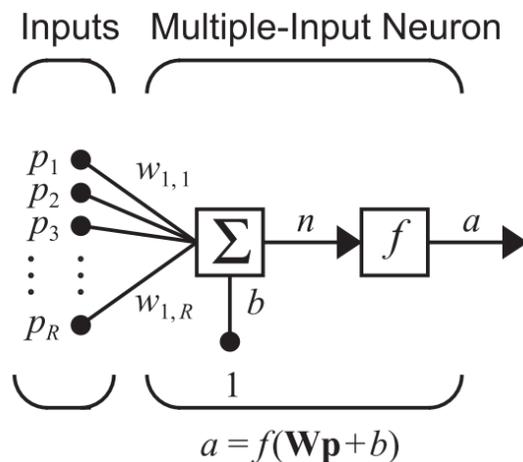
Figure 16 - Multi-Input Single Neuron Model

The multiple-input or single-input single-neuron model contain the following contents:

data input(s), weighted links between the data inputs and the neuron model, a weighted

bias value, a net input function block, and the activation or transfer function block. The

data inputs, if more than one input is provided to the model, will be in the format of a

column vector as the data is presented to the network. Each data entry in the input

column vector is considered an input "P_n" to the feed-forward network, where n = 1, 2, 3, …, R. Each of the inputs have weighted links between the input value and the neuron net input function block. This weighted link is an adjustable scalar parameter of the neuron model that is adjusted during the ANN training process. The data input is multiplied by the associated weighted value which is known as the weight function. The output value of the weight function is then sent to the net input function. There also a weighted bias value that is introduced to the neuron net input function block as well. This weighted bias input has a constant input value of one. As with the other inputs into the neural network the bias input is adjusted by the weighted value of the bias link before entering the net input function. The neural network designer does have the option to omit this bias value and bias weight from the neuron model if desired. But the bias does add some flexibility when attempting to use neural networks to perform a desired behavior. The net input function will use a summing function for this research. This summing function will sum all weighted inputs along with the weighted bias if applicable. The output of the summing function, known as the net input, of the single-neuron model is shown in equation 2.1 [13].

$$net\ input\ (n) = \sum(Wp) + b \qquad\qquad (2.1)$$

where:

W = weighted link value of the associated input

p = neural network input value

b = weighted bias value of the neuron

This net input value is then presented to the transfer function. Some resources will call this transfer function an activation function. The transfer function may be linear or non-linear depending on the application of the ANN which is set by the ANN designer. The ANN transfer function is also known to be part of the ANN single neuron architecture. A common list of transfer functions and the associated MATLAB programming function is provided below.

- Hard Limit (hardlim)
- Symmetrical Hard Limit (hardlims)
- Linear (purelin)
- Saturating Linear (satlin)
- Symmetric Saturating Linear (satlins)
- Log-Sigmoid (logsig)
- Hyperbolic Tangent Sigmoid (tansig)
- Positive Linear (poslin)
- Competitive (compet)

Many of these functions were tested during this research before it was decided that the hyperbolic tangent sigmoid function (MATLAB function: tansig) and the linear (MATLAB function: purelin) transfer functions would be used in this research. The hyperbolic tangent sigmoid function and the log-sigmoid function are known as a squashing function. These functions take any value between -∞ to +∞ as an input and provides an output that is within the range of -1 to +1 and 0 to +1 respectively.

The hyperbolic tangent sigmoid transfer function is known to be highly used in multi-layer networks that are trained with the backpropagation algorithm. Figure 17 and equation 2.2 represent the hyperbolic tangent sigmoid function by providing the input and output relationship [13], [14].
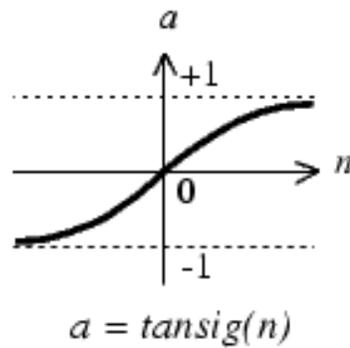
Figure 17 - Hyperbolic Tangent Sigmoid Transfer Function

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}} \qquad (2.2)$$

Where n is any value between -∞ to +∞. It will be shown shortly that the hidden layers

and output layer of the neural network architecture design will both contain transfer

functions. This research will use the hyperbolic tangent sigmoid function within all

hidden layers of this research. As for the output layer, the linear transfer function was

used. This transfer function is very basic as signified by its name and simply provides the

same value to the output as the input to the function. Figure 18 represent the linear

transfer function by providing the input and output relationship [13], [14].
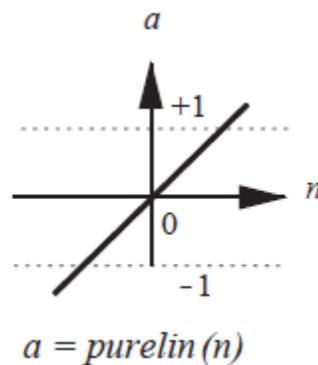
Figure 18 - Linear Transfer Function

Where n is any value between -∞ to +∞. Finally, the output of the transfer function is

then called the neuron output. This output value is the result of the neuron model that is

presented to the user. The function equation for the entire single-neuron model is

presented in Figure 16.

2.2.2 Feed-Forward Multi-Layer Artificial Neural Networks

Single-neuron models are not very powerful neural networks when attempting to solve complex problems by themselves. Most useful and developed ANNs will consist of different series and parallel combinations of the single neuron models to allow more complex problems to be solved. A combination of two or more single-neuron models that are in a parallel configuration will begin to form a layer of neuron models within a network. Figure 19 show the basic construction when the single-neuron model is expanded with multiple "S" parallel neuron models [13].
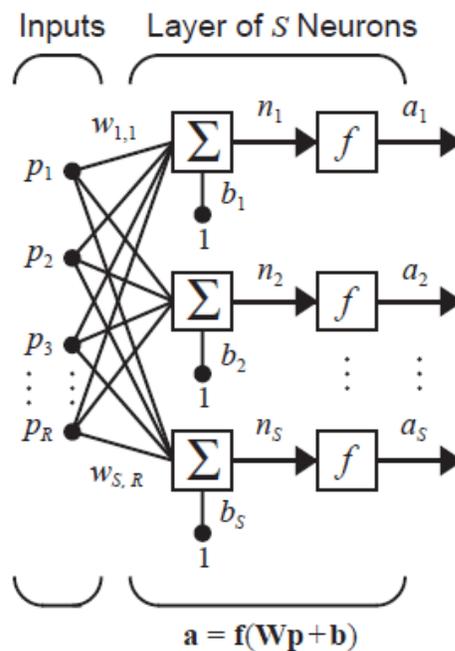
Figure 19 - Multiple Input Multiple Neuron Neural Network in a Single Layer

In Figure 19 it is shown that the neural network consists of a single layer of "S" neurons. This would mean that within the single layer there would be "S" net input functions and transfer functions. The "S" number of neurons and transfer functions used within the layer is determined by how many outputs are needed out of the neural network. A layer as defined in reference [13], is identified by the incoming weighted inputs, the weighted biases, the net input functions (a summing function for this research), the transfer function, and the output column vector for a set of parallel single-neuron models. When a neural network only contains the weighted input values and one layer of neurons, the network is ideally consisting of a set of inputs and an output layer. Even though, these networks are more developed then the single-neuron model they still are very limited on the type of complex problems they can solve. To broaden the type of complex problems that feed forward neural networks can solve, neural networks can be further expanded to contain series combinations of various numbers of differently designed layers. When neural network architectures begin containing multiple layers of neurons, the neural network architecture become known as multi-layer neural networks. Figure 20 provides a basic representation that defines an example of a multi-layer neural network [13]. This research will be using multi-layer ANNs and the next section within this chapter will discuss how they will be used to solve the fault identification problem.

<div align="center">
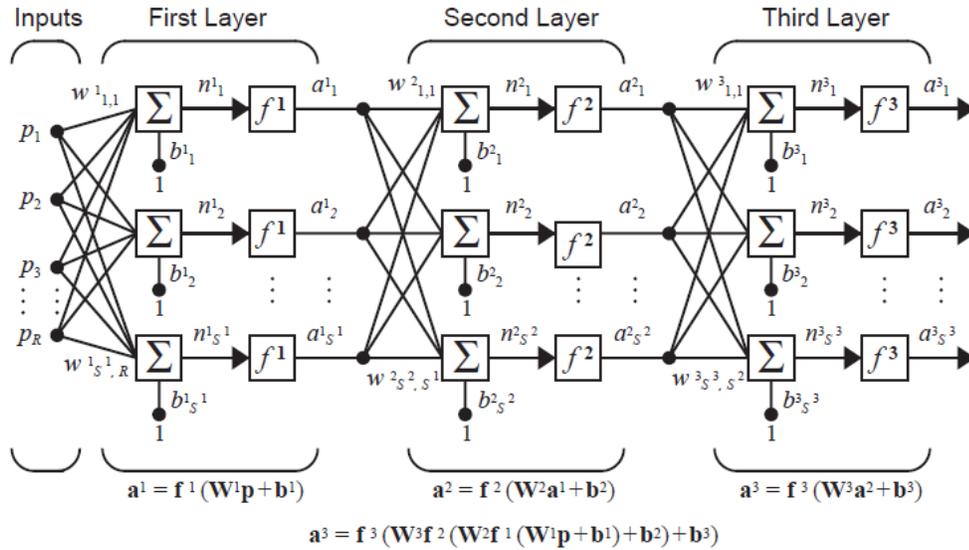
Figure 20 - Multiple Input Multiple Neurons with Multiple Layers

</div>

Multi-layer ANNs will consist of one or more hidden layers and one output layer. Each

defined hidden layer within these networks do not necessarily contain the same number

of neurons in each layer. Layers can be adjusted with different number of neurons

between all hidden and output layers of the network which is a characteristic of neural

networks that make them flexible to solve complex problems. Along with having a

different number of neurons throughout the different layers of the neural network, the

transfer function used within the different layers can also be different. The only exception

with the transfer function is that the same transfer function must be used within each

layer for all neurons. The output values for each neuron in the hidden layers will contain

weighted links to all neurons of the next hidden layer or the output layer. These weighted

links operate in the same way as with the weighted links between the inputs and the first

layer of neurons that was previously discussed in the single neuron model.

Each hidden layer contained in the neural network architecture design will lie between the provided network inputs and the output layer. It is stressed that the network inputs are not identified as a layer. Unlike the number of neurons designed within each hidden layer of the network, the number of parallel neurons used in the output layer will be defined by the target data sets used to train the network. This is known as configuring the neural network which takes place during the training process.

Training the ANNs is an important introductory concept that should be well understood. Training the neural network is a procedure that modifies and adjusts all network and bias weights based on the data it is provided. This process of adjusting the network and bias weighted links is known as the learning rule or referred to as a training algorithm [13]. Learning rules can be categorized into three broad categories: supervised learning, unsupervised learning, and reinforcement or graded learning. Supervised learning is a method of updating network and bias weights based on input and target mapping combination that are provided to the network during training. While in the training process, the inputs are applied to the network to allow the ANN to produce some output values. These predicted output values are then compared to the actual target values. Depending on the error between the ANN predicted output and the actual target value, the learning rule keeps adjust network and bias weights for the ANN predicted output to produce less error. Each iteration of this process is identified as an epoch. This research will be using the supervised learning rule process. The second learning rule category that is available to ANN designers is the unsupervised learning rule. The unsupervised learning rule is a method of updating the network and bias weights based only on the inputs that are applied to the network. Unsupervised learning is a great approach to use if

46

there are no output target values available. The last categorized learning rule is the reinforcement or graded learning rule. Reinforcement or graded learning is very similar to supervised learning. The learning method instead of providing the network the correct target value associated with the network input, the learning rule is given a grade or score. This grade or score is a measure of the network performance over some sequence of inputs. Again, unsupervised and reinforcement learning rules will not be studied to solve fault identification problems at this time.

The full data set of the input and target training data is not used to train these neural networks. The default in the MATLAB neural network toolbox sets 70 percent of the collected training data to be used for training the neural network. Then 15 percent of the data is used to test the network during the training process while the last 15 percent is used to validate the network. These percentage breakdowns can be adjusted from the default values at the ANN designer's discretion. The designer also can select how the training, testing, and validation data will be selected for the training process. MATLAB uses the default approach to select the breakdown of the training, testing, and validation data points in a random fashion. The random approach that was used to select the breakdown of the training values was used within this research.

2.2.3 Fault Identification Approach using Artificial Neural Networks

The entire section 2.2 of this dissertation has been devoted to providing the reader with a basic introduction on how ANNs are designed and how some of the parameters need to

be considered when designing these types of networks. Nearly every resource available that has studied ANNs, will state that there is no specific ANN that can be developed that will solve any and every complex problem. Developing these networks need special attention in a case by case basis. Related to this issue, this research attempts to provide some basic rule of thumb concepts that were observed and to provide the user with some techniques to design ANNs when attempting to solve fault identification problems related to power systems.

This research will be using the versions 2016a and 2019a of MATLAB and Simulink software to perform all model building and neural network tasks. Within the MATLAB and Simulink software versions, the artificial neural network toolbox will be used to design, train, and the test all ANN architectures in order to analyze the accuracy that the ANN approaches can predict fault identification.

2.3 Research Related Work

There are various fault classification and location approaches in existing literature, which can be classified into impedance based [15] [16] [17], traveling waved based [18], and artificial intelligence based methods. This dissertation and thus this section will focus on intelligence-based methods.

There have been and is still ongoing research in many forms that are devoted to solving a variety of complex problems using the artificial intelligence, in particular artificial neural networks. This even holds true using ANNs to solve the problem of identifying electrical

transmission faults. As reviewed in [19], between the years of 2000 to 2005 ANNs had attracted most of the research attention, related to power system, in load forecasting, fault diagnosis, economic dispatch, security assessment, and transient stability. Out of these five top research categories most of the research was using ANNs for load forecasting at 25 percent followed by fault diagnosis at 18 percent. For the research that is being devoted to fault diagnosis, in particular fault identification, there seems to be a leading majority studying the use of ANNs to perform fault classification and fault location on two-terminal single transmission lines [20], [21], [22], [23], [24], and [25]. But this is not the only transmission topology that has been gaining popularity. The two-terminal parallel line topology has gained some attention and is presented in [20], [26], [27], [28], [29], [30], [31], and [32]. Both configurations will be discussed in this section to see how previous works have handled fault classification and fault location problems. There has been other dedicated research to fault identification using other transmission topologies, such as three-terminal or teed transmission lines, but these works are far less common [33]. This dissertation will not cover fault identification techniques into teed networks, but the prior research is worth mentioning and should be a high priority research effort since these lines can cause many challenges for protection engineers and current fault identification techniques.

It was observed that many of the authors that have worked on related research to fault identification on transmission lines, have provided prior research into all transmission network topologies, with some work related to transmission network fault identification. As previously stated, single transmission lines are the most common network topology as seen in a power system. But each of these transmission line topologies share an equal

level of importance and priority. Mainly single and parallel transmission lines can practically operate at any transmission level voltage and can travel a variety of different distances. Teed transmission lines usually operate at lower voltages and are not usually very long in distance. It was not surprising that many of the previous work identified in this dissertation used a wide variety of different type of transmission line parameters. Reference [33], used a single ANN to determine the classification of the fault and which line segment of the multi-terminal line the fault was located on. The model used in this research was a 220 kV multi-terminal (three-terminal) line. The author decided to model a transmission network where all three of the line segments were modeled at different lengths (200 km, 120 km, and 110 km). The input signals derived from the modeled transmission system was normalized between a range of -1 to +1. The transmission model was only tested on the double line to ground faults at different fault locations between 0% to 90% of the line total length, fault inception angles of 0° and 90°, and fault resistances of 0Ω, 50Ω, and 100Ω. The total number of faulted scenarios simulated was 774. From these 774 faulted simulations, the inputs for training the ANN used the fundamental frequency magnitude values for the voltage and current measurements recorded for all three buses. This results in the ANN being trained with 18 inputs of both voltage and current magnitude signals. The target values consisted of a 7-entry column matrix, where the first 4 entries of the column matrix determined the classification of the fault and the last 3 entries determined the faulted line segments. The author shows accurate fault location using ANNs to detect fault identification on multi-terminal lines.

The different system parameters that the identified prior work has used include transmission line operational voltage levels, total length of the modeled transmission line,

the voltage and/or current configuration measurements taken from the power system

models, transmission line parameters, and generation (source) parameters. These

transmission line models contain operational voltages levels between 100 kV to 400 kV.

Most of the transmission line lengths have been modeled in the range of 100 km (62.1371

miles) to 150 km (93.2057 miles), with most of the models using 100 km. This research

uses a 100 km line length to perform all simulations. The main differences within the

previous works have been the approach of using voltage and/or currents as inputs into the

ANN and how the ANNs have been used for fault classification and fault location.

References [22], [23], [24], and [34] use voltage and current measurements as the inputs

into the ANNs. These voltage and current measurements have been obtained at one

terminal of the transmission line through substation equipment of current transformers

(CT) and voltage transformers or potential transformers (VT or PT).

Within reference [22], the author has decided to determine if the transmission system is

experiencing any electrical fault by using a single ANN. The output of this ANN is either

a value of zero or one. If the output value is zero, then the transmission system is not

experiencing any electrical fault. Whereas, if the ANN output value is one, then the

transmission system is experiencing some type of electrical fault. If the fault detection

ANN determines an electrical fault is in existence, then another ANN is used to

determine the classification of the fault. In parallel with the classification of the fault, a

different ANN is used to determine where the fault is located by using pre-defined zones

or protection. These ANNs all share the same input vector to determine their outputs.

This input vector is the fundamental frequency phase voltage and phase current

magnitudes. The input magnitude values have been normalized between values of zero and one before presenting them to the ANNs.

The work presented in [23] was only looking at transmission fault detection and classification. The approach to derive at the fault detection and fault classification was the same as in [22]. The only difference with this approach is the definition of the voltages and current inputs. The faulted measurements were normalized to the pre-fault values of voltage and current. Also, the zero-sequence voltage and current values were inputs into the ANNs to help define when faulted conditions contained ground connections.

A slightly different approach was taken in [34], where the author decided to use two ANNs to identify the classification of the fault. The idea here is that only one of the ANNs will be activated at a time. The fault connection that contains a connection with ground will be the deciding factor for which ANN will be activated to determine faults. The author uses a level detector that takes in the zero-sequence current and outputs a logical zero or one signal that assigns that value to the ground connection. One of these two ANNs will output faults that do not contain ground connections (three-phase faults and line to line faults). Whereas the second ANN will output faults that do have connections with ground (single line to ground faults and double line to ground faults). For a visual representation of this ANN approach using multiple ANNs with a level detector see Figure 21 [34].
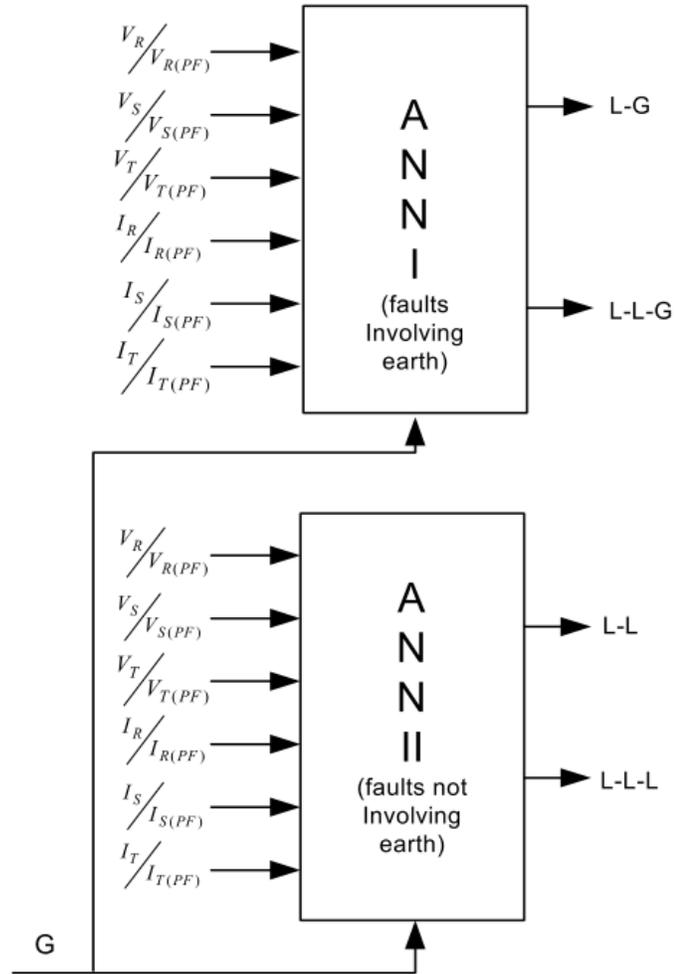
Figure 21 - Two ANNs Fault Classification Approach

Once the fault classification has been determined, another ANN will be triggered to determine the electrical fault location. During the fault location step, each fault classification output will be tied to separate distinct ANN. Therefore, there are four different ANNs that are used to identify the fault location. Fundamental phase voltages and current magnitudes were selected as the inputs into the ANNs. The voltage and current measurements were normalized with the pre-fault measurements before submitting the measurements as inputs into the ANNs.

The author in reference [24] took a very similar approach to determine if a transmission fault exist on the transmission system as mentioned in previous works. But when the author looked at fault location there was three different sets of inputs analyzed. The author used three independent ANNs to analyze the fault location. These different ANNs analyzed inputs for only current magnitude measurements, only voltage magnitude measurements, and voltage and current magnitude measurements.

Each of these ANNs were examined extensively to determine the networks architecture. For instance, there were many iterations of trial and error to determine the optimal number of hidden layers and hidden layer neurons for each tested scenario. This was a large effort in this research as well, which was used to determine the best architecture to use of each measurement configuration studied. The previous works were also trained with data from many different fault locations, fault resistances, fault inceptions angles, and fault types. References [24] and [34], the results of the ANN predictions for fault location used the formula for percent error as given in equation 2.4.

$$Error\ (\%) = \frac{|ANN\ Output - Fault\ Location|}{Length\ of\ Line} * 100 \tag{2.4}$$

This research will present the results in a similar but different approach. Since the ANN fault location results will be passed along to the field personnel, it was decided to present the results in absolute error. This absolute error will present the amount of error that exist between the actual fault location to the predicted ANN fault location in units km. The

results were decided to be present this way since the fault identification will be relayed to field personnel. It is the goal of this research to provide fault identification data that contains low kilometers of error to keep the field personnel from having to look for the faulted condition for a long period of time. If field personnel are searching for the fault conditions for long periods of time, this would defeat the purpose of providing fault identification values to field personnel to locate, isolate, and correct the faulted condition as quickly as possible.

While reviewing previous work with two-terminal parallel transmission lines it was identified that mutual coupling between the transmission circuits on the same structure or transmission lines running near each other can cause pre-mature breaker operations in a healthy non-faulted transmission circuit. This pre-mature breaker tripping is a common point that all prior art has mentioned and focused on. References [30], [31], and [32] focused on the two-terminal parallel transmission line model, which it was determined that the exact same model with the same parameters was used. This was not surprising since the same authors were related to the resources. The model that was studied was a 220 kV transmission system with both circuits at 100 km in length. These models all considered that mutual coupling between the two circuits did exist. Reference [31], the author looked at two different approaches to solve the transmission fault location problem. The first approach looked at a single ANN, where the voltage and current measurements were inputs to the ANN and the output of the ANN was the fault location. There was no mention of the fault classification in this approach. This approach used the fundamental magnitude phase currents and bus voltage at one end of the transmission lines as the inputs into the ANN (9 inputs into the network). These input values have

55

been normalized to an input level of -1 to +1. As with the prior literature used for the

two-terminal single transmission line ANN architecture, mainly the hidden layer

architecture, was determined by a trial and error approach. The other approach in [31],

was classified as a modular approach which uses multiple ANNs to determine fault

classification and fault location. In this approach, the fundamental magnitude bus

voltages and phase currents were used as inputs. The fault detector/classifier ANN will

identify the type of fault as either single phase to ground, phase to phase, double phase to

ground, or three-phase. Based on the output of the fault detector/classifier ANN, another

ANN will be activated to estimate the fault location on the transmission system. The fault

locator is made up of four independent ANNs which are activated from the fault

detector/classifier output. As with the single transmission line, the author uses percent

error to determine the performance of the ANNs. References [30] and [32], the authors

used the modular approach just described, but they only use phase to phase faults and

single line to ground faults to train their networks respectively.

Within reference [20], there were mentions that there is no single neural network that can

detect faults on any transmission system. But there are neural network structures that can

be used in many architecture forms to solve all fault location problems on transmission

systems. This is what has been seen in all prior works. Each one of these sources use

different numbers of inputs, different number of hidden layers, different number of

neurons in the hidden layers, and all have normalized the inputs and outputs differently.

As a collection, all prior literature has been successful in using ANNs to solve fault

location to a low percentage of error.

Chapter 3 – Fault Identification with Single Transmission Lines using a Single ANN Approach

This chapter begins the discussion on the approach for fault identification for the first transmission line configuration that will be evaluated. The transmission line configuration that is used in this chapter is the single transmission line connected between two distinct substations. As stated within the introduction, fault identification within the context of this dissertation is identifying the fault classification or fault type and the location of that fault which has occurred on the unique transmission line. This research assumes that the faulted scenario has occurred on the transmission line and all transmission protection devices protecting the transmission line have operated to isolate the fault. The fault classification ANN output will relate the faulted scenario to one of the ten different fault types that could have possibly occurred. These ten different fault types where discussed in chapter 2 in detail and to recap they are known as the line to ground fault (LG), line to line fault (LL), double line to ground fault (LLG), or three-phase fault (LLL). All phases of this research will be using MATLAB and Simulink software to develop the transmission line topology model. The transmission line model used in chapters three and four will both use the single transmission line configuration that is connected between two substations. The single transmission line model used within this phase of the research utilizes the 2016a version of the MATLAB and Simulink software.

The approach proposed in this chapter attempts to use a single artificial neural network to identify both the fault classification and fault location. The ANN architecture will be discussed in detail later in this chapter. The following section in this chapter will provide

a breakdown to the Simulink modeling details that make up the single transmission line

model. Each type of modeling block/data will be discussed by describing the role that

each modeling block takes to provide input and target values as an output to the

transmission line model. Once the model has been discussed in detail the process of

designing and training the ANNs will be described. This will encompass building and

training the ANNs, gathering testing data, and providing the testing data to the trained

ANNs. Fault identification results and conclusion on the ANN architecture results will be

presented to describe how well the ANNs can predict fault classification and fault

location.

3.1 Two-Terminal Single Transmission Line Model for Phase 1

This section provides the modeling details as an overview of the two-terminal single

transmission line model. The single transmission line model was developed using

MATLAB and Simulink software version 2016a. The objective for the development of

the transmission line model was to create a model that would provide voltage and current

measurements at both substations connected to the transmission line. These measurement

values will have magnitude levels which could be experienced by a real-world utility.

The transmission line models used within all phases of this research will be simulating a

60 hertz (Hz), 500 kV transmission line that is modeled at 100 km (62.13712 miles) in

length. The conversion between kilometers to miles can be calculated using equation 3.1.

$$Number\ of\ Miles = \frac{Number\ of\ Kilometers}{1.60934\frac{Kilometers}{Miles}} \qquad (3.1)$$

The single transmission line model consists of two generator modeling blocks, two

equivalized mutual impedance blocks, two voltage and current V-I measurement blocks,

and the transmission line topology. The transmission line is modeled as two distributed

parameter line blocks. The transmission line is modeled with two distributed parameter

line blocks to allow for any type of fault to be applied at any point along the transmission

line by changing the distance parameters of the distributed parameter line blocks. There

will be more discussion on applying faults to transmission line when discussing the

distributed parameter line blocks in more detail. Figure 22 provides an illustration of the

single transmission line model that was developed in Simulink.
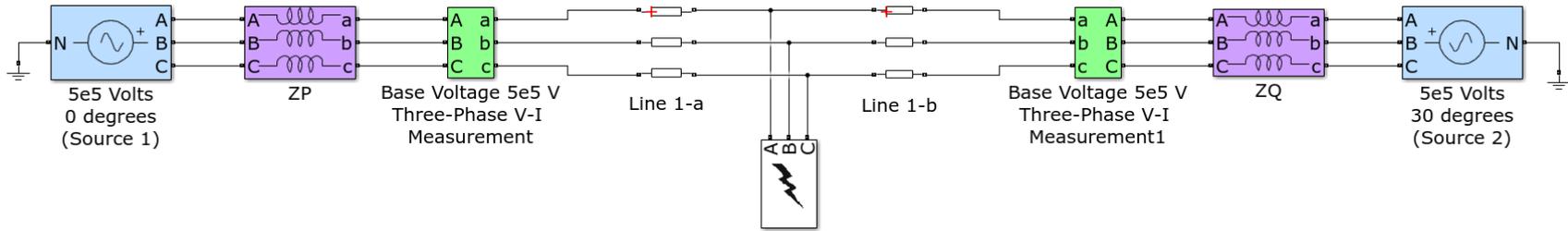
Figure 22 – Single Transmission Line Simulink Model

3.1.1 Generation Modeling – Single Transmission Line Model

The construction of this Simulink model started with the development of two generator sources. In a traditional utility scale power system, there are many different generators that are connected to the same power grid that help support the generation of electrical power which is then transported through the transmission lines to be consumed by the electrical loads. For transmission line modeling purposes, power only needs to be transported across the transmission line that is being studied so there is no need to provide detail for numerous generators. To provide power flow across the modeled transmission line only two modeled generators are needed. Therefore, the two modeled generators shown in Figure 22, should be viewed as an equivalization to all the generators seen by each end of the transmission line that would be connected to the power system. The two generation sources will be located at the endpoints of the model and again are intended to simulate the electrical power to flow across the modeled transmission line. For power to flow across the transmission line there needs to be a potential difference between the two generation sources. Since both generators will be generating at a magnitude of 500 kV, the potential difference was created by changing the phase angles between the two generators. It should be noted that both generators are set to generate electricity at a 60 Hz frequency. Table 5, displays the generator parameters used within the single transmission line model.

Table 5 - Single Transmission Line Model Generator Parameters

	Generation Modeling Parameters	
	Generator Connected to Substation A	Generator Connected to Substation B
Amplitude (Vrms Ph-Ph)	500 kV	500 kV
Phase Angle (Degrees)	0˚	30˚
Frequency (Hz)	60	60

3.1.2 Transmission System Impedance Modeling – Single Transmission Line Model

The next modeling elements that were added were the equivalized mutual impedance

blocks. Transmission systems encompass many transmission elements (i.e. transformers

and transmission lines) that contain impedance values that contribute and limit the

amount of power that will flow throughout each transmission line. Without modeling the

equivalized transmission system impedance that would be seen by the modeled

transmission line, the total generation from the modeled generators would flow across the

transmission line and would not provide a realistic (real-world) modeled scenario.

Therefore, the mutual impedance blocks contain Thevenin system impedance values as

seen at each end of the transmission line to simulate closer to realistic transmission line

flows. Table 6 and Table 7 provide the positive and zero sequence (resistance and

inductance) modeled equivalized mutual impedance values as seen by both substations.

Table 6 - Equivalized Mutual Impedance at Substation A

	Equivalized System Impedances at Substation A	
	Positive Sequence	Zero Sequence
Resistance (Ω)	17.177	2.5904
Inductance (H)	0.1208	0.0391

Table 7 - Equivalized Mutual Impedance at Substation B

	Equivalized System Impedances at Substation B	
	Positive Sequence	Zero Sequence
Resistance (Ω)	15.31	0.7229
Inductance (H)	0.1218	0.0401

The inductance values that are shown in Table 6 and Table 7 have been converted into the units of Henrys (H). Typical inductance values for a power system will be given in reactance and should be converted using equation 3.2 where the system frequency for this research is 60 Hz.

$$Inductance\ (H) = \frac{Inductive\ Reactance\ (X_L)}{(2*\pi*System\ Frequency)} \qquad (3.2)$$

3.1.3 Current and Voltage Measurement Modeling - Single Transmission Line Model

The next set of details that are placed in the model are the three-phase voltage and current (V-I) measurement blocks. The purpose of the measurement blocks is to output instantaneous voltage and/or current measurements that would be collected at either substation. Since measurements are to be taken at both ends of transmission line inside the substations, these measurement blocks are placed in the model at the end points of the transmission line and should be assumed that the measurement devices live inside the substation fences. Within the model in Figure 22, its assumed that the point between the equivalized mutual impedance blocks and the three-phase V-I measurement blocks should signify the location of each substation or bus that the transmission line is connected. The output of the voltage and current measurements from the measurement block are in per unit quantities at a sampling rate of 128 samples per cycle. The per unit measurements are based on voltage and power base values specified by the development of the Simulink model. All models within this research use a power base of 100 MVA and voltage base of 500 kV. Voltage measurements are recorded based on a phase to ground orientation. The modeled three-phase V-I measurement blocks use voltage and current tags to allow the model to access the voltage and current measurement outputs.

This research uses ANNs to predicts the fault classification and fault location using voltage and/or current phasor measurements at a time stamp after the fault has been applied to the transmission line. Since the transmission line model is recording instantaneous per unit voltage and current measurement from the output of the three-

phase V-I measurement blocks, these instantaneous voltage and current measurements need to be converted into phasor values for each sample during the entire simulation. This ensures that at any point during the simulation can be used to analyze the prediction for fault identification if needed. Figure 23 and Figure 24 show the high level Simulink diagrams that are used to convert the recorded instantaneous V-I measurements into phasor (magnitude and angle) values. These figures only show the instantaneous values being convert at one substation. The Simulink model contains another conversion process for the second substation.
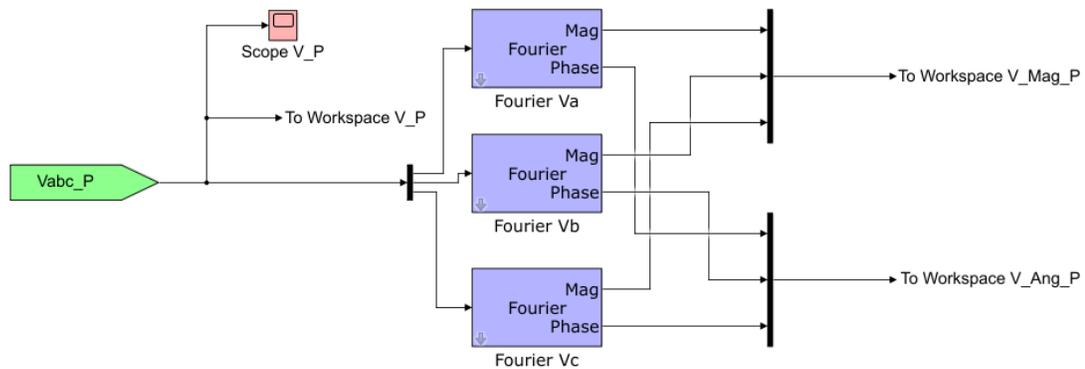
Figure 23 - Single Transmission Line Voltage Phasor Conversion
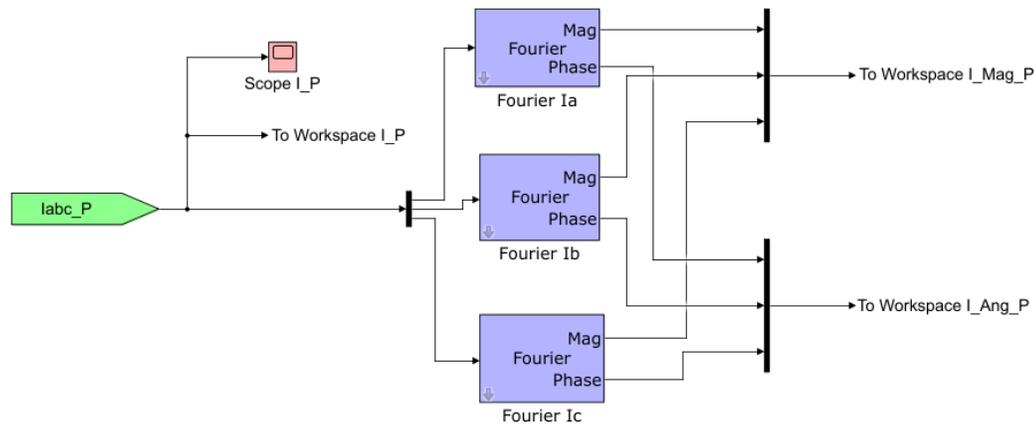
Figure 24 - Single Transmission Line Current Phasor Conversion

The measurement tags are labeled Iabc_P, Iabc_Q, Vabc_P, and Vabc_Q and are used to

associate the instantaneous voltage and current measurements from the V-I measurement

blocks. The designation of "P" and "Q" are used to represent the two distinct substations,

substation A and substation B respectively, that connects the transmission line. To

provide a mechanism for troubleshooting, the instantaneous measurements of voltage and

current at each substation are recorded to the MATLAB workspace. This model also

captures the voltage and current waveforms using the scope block. The collected data and

the scope provide the user a visual troubleshooting tool to see if there may exist any

issues with the voltage and current waveforms after the simulation completes. The

voltage and current instantaneous measurements are then separated by each phase using

the de-mux block. Now since the waveforms are separated by phases, the instantaneous

phase measurements can be converted into phasor values by performing the Fourier

analysis of the voltage or current signals. Phasor values provide the measurement

66

quantities in magnitude values measured in per unit and phase angles measured in degrees.

Once the Fourier analysis on the signals are complete, the phasor quantities are joined back together by the grouping of phase magnitudes and phase angles. These join conditions are completed by using the mux block in Simulink.

3.1.4 Transmission Line Modeling – Single Transmission Line Model

The final detail of the single transmission line model defines how the transmission line is modeled. The transmission line is modeled by using the distributed parameter line block in Simulink. Since this research will be simulating a fault moving down the transmission line, there will need to be two distributed parameter line blocks used to define the specifications of the entire transmission line. It should be viewed that placing the data from the two distributed parameter lines together will form the complete data representation for the entire transmission line. One of these distributed parameter line blocks will represent the portion of the transmission line from substation A to the faulted point along the transmission line. While the second distributed parameter line block will model the portion of the transmission line from the faulted point to substation B. During any faulted simulation of this research, if the two distributed parameter line blocks are viewed as one, their combined line distance parameter should sum up to equal the total length of the transmission line. Some of the other parameters of the distributed parameter line block allows the model to define how many phases are contained within that specific

transmission line. This parameter will be used when discussing the difference between single and parallel transmission line configurations. If there are only one transmission line being modeled, then the transmission line will contain only three phases. This will be the case for the single transmission line covered in chapters three and four. Within the distributed parameter line block the resistance, inductance, and capacitance of the line should be specified based on per unit length. The resistance values should be provided in ohms per km (Ω/km), inductance in henrys per km (H/km), and capacitance in farads per km (F/km). These transmission line characteristics should also be provided in positive sequence, zero sequence, and mutual zero sequence components if applicable. The positive sequence and zero sequence are known to be self-impedance quantities of the transmission line. When there is more than one transmission line near each other, either contained in the same right of way easement or on the same transmission structures, there can be impedance added to the transmission line by mutual inductance. The data represented in Table 8 expresses the impedance sequence data for the two distributed parameter line blocks for the single transmission line model. It should also be noted that the impedance sequence data that is presented are in units of ohms, henrys, and farads. MATLAB expects the values to be in per unit length therefore, the actual values should be divided by 1.61 to express the values in per kilometer.

Table 8 - Single Transmission Line Sequence Impedance Values

	Distributed Parameter Line Characteristics	
	Positive Sequence	Negative Sequence
Line Resistance (Ω)	0.249168	0.60241
Line Inductance (H)	0.00156277	0.0048303
Line Capacitance (F)	1.9469E-08	1.206678E-08

It is not a true detail of the transmission line model, but a three-phase fault block was used to apply faults to the transmission line. The three-phase fault block allows the model builder to select the type of fault that should be applied to the model and the location of that fault. These fault types could be any of the ten fault types that have been discussed throughout this dissertation. The three-phase fault block also sets the fault resistance of the fault. The fault resistance values are set within the block parameters and the values are set differently depending on the type of fault that is being applied to the model. Within the parameters of the fault block there are four check boxes that are used to select the type of fault that will be applied to the model. These check boxes correspond to the three current carrying conductors (phases) of the transmission line (phase A, phase B, and phase C) and a grounding (static or earth) connection. There are also two text boxes that allow for fault impedance values. These text boxes correspond to the labels of R_{ON} and R_G. R_{ON} is the fault impedance located in the phase conductor where R_G is the fault impedance in the ground connection. Figure 25 provides a representation of how faults can be applied to the transmission line [35].
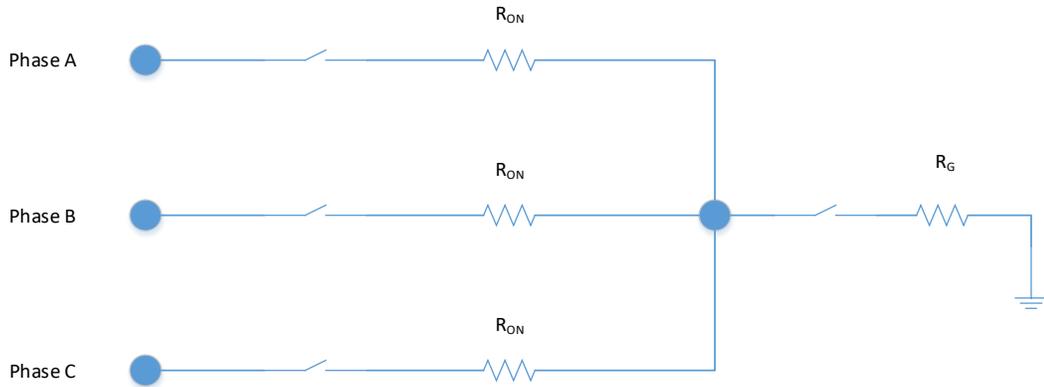
Figure 25 - Three-Phase Fault Block Impedance Diagram

Once the fault block is selected to perform line to ground faults, the algorithm of the fault

block closes the switch on the faulted phase and the ground connections. For the line to

ground fault all the fault current will flow through the phase and ground impedance. The

sum of the phase and ground impedance would determine the total fault impedance. It

was selected that phase impedance would be set to $0.01 \, \Omega$ (since no impedance values

can be a bolted fault impedance at 0 ohms in MATLAB) and the ground impedance

would be set to the total fault resistance value. This approach is very similar for the

double line to ground fault (LLG). For LLG faults applied to the transmission line model,

the switches for the two faulted phases and the switch for the faulted ground connection

are closed. Again, since all the entire fault current will flow through the faulted ground

impedance, the R_{ON} fault impedance is set to $0.01 \, \Omega$ and the R_G fault impedance is set to

the total fault resistance value. Faults that have no ground connection are applied to the

transmission system in a little different way. Phases that are contained in the fault, either

the fault type be line to line (LL) or three-phase (LLL) faults, only the check boxes of the

faulted phases are checked. If the check box for the ground connection is not selected the

R_G text box will be grayed out and no value can be entered. During a three-phase fault the R_{ON} field is set equal to the total fault resistance value. But for a line to line faults, the fault current will through one phase and then flow back through the other phase. Therefore, the R_{ON} fault resistance value should be set equal to half of the fault resistance value ($R_F/2$).

3.2 Development of Input and Target Training Data for Single Transmission Line using Single ANN Approach

Artificial neural networks relate to a technique that is used to solve complex problems by teaching or presenting a set of actual data with the expectation that the neural network can recreate that scenario by predicting network output(s) assuming the network is provided the same or similar input values. As for this research, simulated electrical transmission fault data will be used as the neural network input and target training data from the developed MATLAB Simulink model. A problem starts to surface when trying to determine where the fault data is derived or obtained. Since no electric utility is going to have actual fault data for any line on their system and for nearly every point on any transmission line, the transmission line topology needs to modeled and fault simulation data should be obtained by simulating a comparable transmission line model to obtain this input and target fault data.

3.2.1 Development of Input Training Data for Single Transmission Line

The idea behind this research is to determine the how accurate ANNs can predict fault identification (fault classification and accurate location of the fault) when a variety of different phasor measurement, voltage and/or current phasors, may or may not be available. This correlates to nine possible measurement arrangements that are identified as the voltage and/or current phasors being available. These nine measurement arrangements are:

- Voltage phasor from substation A

- Voltage phasor from substation B

- Current phasor from substation A

- Current phasor from substation B

- Voltage and current phasors from substation A

- Voltage and current phasors from substation B

- Voltage phasors from substation A and substation B

- Current phasors from substation A and substation B

- Voltage and current phasors from substation A and substation B

Since there are numerous combinations that could be selected to evaluate the fault identification performance with the use of ANNs and the time allocation that it takes to perform the fault identification analysis on each of these combinations, not all identified combinations were studied in this research. This research will focus on using the

following measurement combinations to analyze the predictability for fault identification

using ANNs:

- Voltage phasors from Substation A

- Current phasors from Substation A

- Voltage and current phasors from Substation A

- Voltage and current phasors from Substation A and Substation B

When ANNs are used to predict fault identification, the networks need to be trained with

sufficient fault training data so that the identification of the fault can be accurately

identified for any faulted situation (fault resistance, fault type, and fault location). It was

shown in Chapter 2, that electrical faults can occur in ten different classifications. Any of

these fault classifications can occur within any point along the transmission line and the

ANN that is selected and developed should have the ability to identify the location of any

fault. Therefore, one of the important characteristics of the fault data used for training

purposes need to contain simulated data that was collect from the numerous fault

locations on the transmission line model. The data collected from each simulated fault

location should contain data for every fault type as well. It was important to decide where

to initially begin the first fault location simulation and what would be the final fault

location simulation. For the initial phase of this research, the initial fault location was set

to 1 km from substation A while the final fault location was be set no closer to substation

B then 1 km. Within the distributed parameter line blocks, in the Simulink transmission

line model, the fault location was set based on a reference location using substation A of

the 100 km transmission line as the reference. Since this research is crucial in being able

to provide fault identification analysis using ANNs, the initiated fault on the transmission line must be shifted down the transmission line starting at the initial fault location and moving the fault to the final fault location. This shift in the fault location was set based on a fault step size that was set by a pre-defined step distance to move the fault down the transmission line toward substation B. After each fault location scenario was simulated, this faulted scenario was then shifted down the entire transmission line until the faulted scenario reached a minimum of 1 km in distance from substation B. Since the transmission line model uses two distributed parameter line blocks and if a fault were placed right on the substation bus, this would result in the distance parameter of one of the distributed parameter line blocks to be set to 0 km. When the distance parameter is set to 0 km, MATLAB will flag an error in the simulated model stating that the distrusted parameter line block cannot contain a zero value for the distance parameter. Therefore, the process used in this phase of the research stops the fault at 1 km from substation A and B. Figure 26 provides a visual aid when applying the fault on the transmission line starting at an initial fault location and then moving that same faulted scenario down the transmission line to the other fault locations on the transmission line.
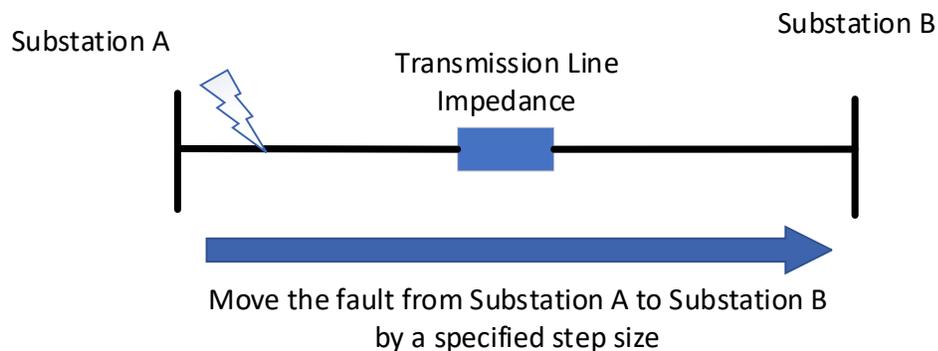
Figure 26 - Moving the Faulted Condition Down the Transmission Line for Simulation

This approach was performed for the ten different fault classification. Two sets of fault training data were created by decreasing the fault step distance during the model faulted simulations to provide more training data. By varying the step size of moving the transmission fault down the transmission line, the input and target training data was able to contain more data for training the ANN. The moving step size was analyzed for 0.1 km, and 0.05 km. This allows for the results to be evaluated to determine if more training data provided more accurate fault identification results.

When faults occur on the power system, they create an abnormal path for fault current to flow through some amount of fault impedance that is associated with the fault. This fault impedance can vary and will vary for every fault. Different sets of fault impedances were studied in this research. This process of simulating faults with different fault resistances adds more input and target training data and provides the ANN the ability to predict fault identification with a more robust set of fault data containing different fault impedances. Since modifying the step size used to move the fault down the transmission line and adding more fault resistance to the simulation data, the input and target training sets can potentially increase or decrease based on the parameters used during the simulation of the different faulted scenarios.

This phase of the research uses the single transmission line configuration and attempts to use one ANN to predict both the fault classification and fault location with fault step distances of 0.1 km and 0.05 km were tested. A fault resistance range of 1 ohm (Ω) to 50 Ω was used for simulating the faulted conditions. Between the fault resistance range faulted condition were also simulated with step sizes of fault resistance multiples of 10 Ω. This results in fault resistances of 1 Ω, 10 Ω, 20 Ω, 30 Ω, 40 Ω, and 50 Ω being used.

With all the different combinations of faulted scenarios that are being tested, equation 3.3 was used to determine the number of training sets or columns of data that resided in the input and target training data.

$$Number\ of\ Training\ Data\ Sets = (\#\ of\ Parallel\ Lines) *$$
$$(\#\ of\ Fault\ Classifications) * (\#\ of\ Fault\ Resistances) *$$
$$(\#\ of\ Fault\ Applied\ to\ Transmission\ Line) \tag{3.3}$$

Within equation 3.3 the number of fault steps may not seem clear. This portion of the equation is related to how many steps the fault was moved down the transmission line for the fault to be simulated on the transmission line. This portion of equation 3.3 is calculated using a two-step process shown in equations 3.4 and 3.5.

$$Number\ of\ Fault\ Steps = floor(\frac{Total\ Transmission\ Line\ Length}{Fault\ Step\ Size}) \tag{3.4}$$

Based on the value of the step size, equation 3.4 has the potential of providing a whole number or decimal number. If a decimal number is the outcome of equation 3.4 before the floor function is applied, then the last fault location will be just short of substation B. The floor function is a function in most programming languages that takes the argument of a value and reduces that value to the next lowest whole number. But if equation 3.4 outputs a whole number before the floor function is applied then the value will not

change. This could result in the last fault location to applied at the bus of substation B.

Since the distance parameter of the distributed parameter line block in MATLAB cannot

contain a zero value, the number of fault steps is reduced by 1 km, such that the last fault

occurs 1 km before substation B. This conditional equation is shown in equation 3.5.

$$If\ Number\ of\ Fault\ Steps = whole\ number:$$

$$Number\ of\ Fault\ Steps = \ Number\ of\ Fault\ Steps - 1 \qquad\qquad (3.5)$$

While reviewing the results of the ANN predictability for fault identification this research

will be attempting to correlate how well the fault identification ANNs perform related to

how many training data sets were used to train the neural networks. Using equation 3.3

the number of training sets that reside in each input and target training data are provided

in Table 9.

Table 9 - Number of Training Data Sets Used for ANN Training - Phase 1

Number of Training Data Sets Used for ANN Training - Phase 1		
Fault Resistances	Fault Location Step Size	Number of Training Data Sets
1Ω, 10Ω, 20Ω, 30Ω, 40Ω, 50Ω	0.1 km	56,940
1Ω, 10Ω, 20Ω, 30Ω, 40Ω, 50Ω	0.05 km	119,940

These faulted scenarios were simulated for eight cycles. This simulation of eight cycles

consisted of normal power flow across the transmission line for the first two cycles of

simulation time. The faulted conditions were applied to the model at the beginning of the

second cycle. This fault was then never removed from the transmission line for the

remainder of the simulation time frame. This results in the fault being applied to the

transmission line for a total of six cycles. Faulted phasor measurements of voltage and

current were recorded at each substation. The data was sampled and recorded at a

sampling rate of 128 samples per cycle. This creates a total of 1,025 data samples for

each measurement type for the entire eight cycles of simulation time. This research uses a

single sample data point within the simulation time frame for each faulted scenario. The

input phasor measurements that were collected to possibly be used for training the ANNs

collected the data sample at the beginning of the third, fourth, fifth, sixth, seventh, and

eighth cycles. This provided the opportunity to use any of these cycles as data points to

train and evaluate the ANNs. During the first phase of this research, it was decided that

all phases of this research would utilize the fifth cycle data point and only use the other

cycle data points if needed. Equation 3.6 provides an example of a of the input training

data layout in the input training matrix for voltage and current fault data being collected

at substation A.

$$
Input\ Data: Voltage\ and\ Current\ from\ One\ Substation = \begin{bmatrix} Magnitude\ V_a \\ Magnitude\ V_b \\ Magnitude\ Vc \\ Angle\ V_a \\ Angle\ V_b \\ Angle\ V_c \\ Magnitude\ I_a \\ Magnitude\ I_b \\ Magnitude\ I_c \\ Angle\ I_a \\ Angle\ I_b \\ Angle\ I_c \end{bmatrix} \quad (3.6)
$$

3.2.2 Development of Training Target Data for Single Transmission Line

This research will be using the supervised artificial neural network learning rule when

training the neural networks for all phases. Supervised learning as discussed in chapter 2

is a technique that performs a mapping algorithm for inputs presented to the neural

network to the associated targets that are also presented to the neural network during the

training process. The training target data for the first phase will contain five data points

orientated in columns for every set of data collected or for the number of simulated

faulted scenarios that were performed. The five data points will correspond to a physical

connection with phase A, phase B, phase C, and ground, and then the actual location of

the fault. Equation 3.7 provides a visual representation of the orientation for each column

of data collected for each simulated faulted condition.

$$ANN\ Training\ Target\ Matrix = \begin{bmatrix} Phase\ A \\ Phase\ B \\ Phase\ C \\ Ground \\ Fault\ Location \end{bmatrix} \quad\quad (3.7)$$

The first four entries in each column of the ANN training target data will only contain

discrete values of either 0 or 1. These top four entries are going to describe the fault

classification that has occurred on the transmission line. Each fault classification entry

will be assigned a value based on fault connection or no-connection algorithm. If the fault

that has occurred on the transmission line creates an abnormal path for current to flow

between that a phase conductor or the ground conductor, then that entry for the conductor

representation in the training target matrix will be assigned a value of 1. Any phase or

ground conductors that has not experienced the faulted condition, the corresponding data

entry will be assigned a 0 value. The fifth entry of the target matrix is the actual fault

location from substation A (using substation A as the reference substation). The fault

location entry will be a floating-point value between the range of zero and the total

transmission line length.

3.3 Training the Single ANNs for the Single Transmission Line Model

After the voltage and current phasor input training data and the associated target data has

been collected, the focus of developing and training the ANN architectures comes to the

forefront. This phase will be using single and multi-hidden layer feed forward neural

networks. The developed multi-hidden layer neural networks will only contain two

hidden layers. Figure 27, shows the high-level layout of the multi-hidden layer perceptron

neural network that this research utilizes.
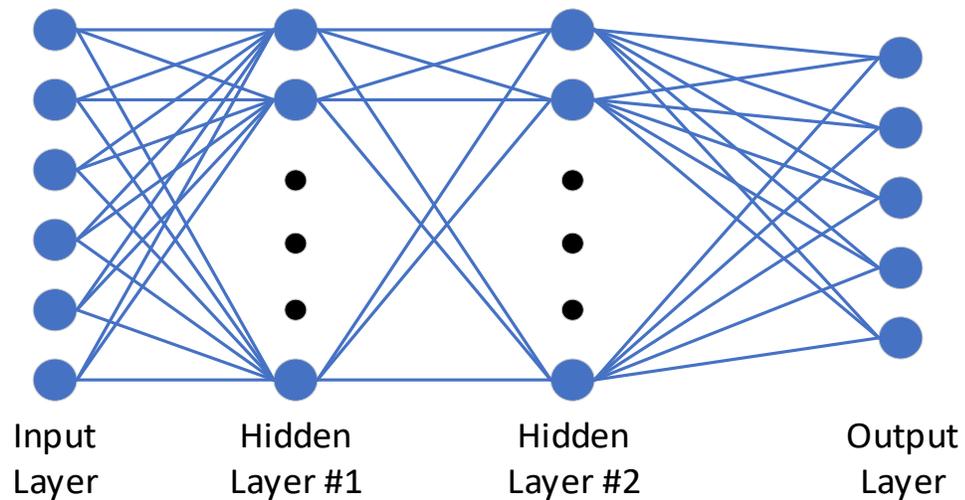
Figure 27 - Multi-Layer Perceptron Neural Network (Phase 1)

The left most layer of the network is the input layer. The diagram shown in Figure 27 might look as if the input layer contains neurons as with all other layers of the network. But this is a false assumption and should be understood that the input layer does not contain any neurons. The input layer should not be thought of as a layer of neurons rather a data entry port that accepts inputs into the networks. The number of inputs that each ANN will accept in the input layer will be configured by the input training data which will be presented to the network during the training process. Each input in the input layer will correlate to one entry of the input training data. Throughout this research the input data will contain combinations of voltage and current phasors collected at the transmission line buses located within the substations. Table 10, links the number of inputs of the input training data to the type of phasor measurements being used to train the ANNs. Keep in mind that this table is only acceptable for the single transmission line topology that is being used in phases 1 and 2 of this research.

Table 10 - Number of ANN Inputs per Type of Phasor Measurement

Number of ANN Inputs per Type of Phasor Measurement	
Phasor Measurements with Orientation	Number of ANN Inputs
Current (I) @ Substation A	6
Voltage (V) @ Substation A	6
Voltage (V) and Current (I) @ Substation A	12
Voltage (V) and Current (I) @ Substations A and B	24

The layer to the far right of the neural network will be known as the output layer. The output layer shares some of the same attributes as the input layer when compared to configuring the output layer. The output layer once it has been configured will contain the same number of network outputs as present in the target training data. For the first phase of this research the trained ANN output layer will have five network outputs as shown in equation 3.7. The output layer begins to diverge from the input layer since the output layer will contain neurons. The output layer neurons will utilize the pure linear transfer function (MATLAB function: purelin) within all phases of this research. The pure linear transfer function was reviewed in chapter 2.

All other layers associated with the neural networks between the input layer and output layer are known as the hidden layers. This research concentrates on analyzing the benefits of using either one or two hidden layers and varying the number of neurons used in each layer of the network. While using a single hidden layer ANN, the hidden layer neurons were ranged between 6 to 36 neurons to analyze the ANNs ability to detect fault identification. But not every integer value of neuron between the range of 6 to 36 neurons were evaluated. The neurons that were evaluated began at 6 neurons and then were varied by steps of three neurons until 36 neurons were applied to the ANN structure. Each neuron associated with the hidden layer used the hyperbolic tangent sigmoid transfer

function (MATLAB function: tansig). The hyperbolic tangent sigmoid transfer function was reviewed in chapter 2. Within the multi-hidden layer (two hidden layers) architecture, the hidden layer neurons in both layers were ranged between 12 to 21 neurons. Again, neurons in each layer were varied by steps of 3 neurons. The neurons in the multi-layer ANNs also used the tangent sigmoid transfer function.

Before any of the training data (input or target data) was presented to the network for training, the training data was normalized in some fashion. The input data is going to contain a range of numeric data. This input data contains different combinations of voltage magnitudes, current magnitudes, and voltage and current phase angles. For input data collected at the initial sample of the 5th cycle of simulation, Table 11 shows the maximum and minimum values for the voltage and current phasor measurements.

Table 11 - Voltage and Current Maximum and Minimum Input Data Values

Voltage and Current Maximum and Minimum Input Data Values		
	Max Value	Min Value
Voltage Magnitude (per unit)	1.022	0.0014
Voltage Angle (degrees)	180	-179.98
Current Magnitude (per unit)	82.21	2.3305
Current Angle (degrees)	180	-180

As seen in Table 11, voltage magnitudes will range between values of 0 and just above 1 per unit. Current magnitudes can have a high range of positive values that will depend on the transmission line parameters, but for this phase of the research it was observed that the values ranged between values of 2 to approximately 82 per unit. The voltage and current phase angles will both range from -180 to +180 degrees. Since the input data can

have a diverse range of values, each type of input data was normalized differently.

Voltage and current magnitudes were normalized to the maximum voltage magnitude or

current magnitude recorded within all simulated faulted scenarios. This will keep all

voltage and current magnitudes between the values of 0 to 1 before applying the

magnitudes to the neural network. The voltage and current phase angles were normalized

to the maximum positive phase angle recorded for all simulated faulted scenarios. This

will keep all phase angles between -1 to +1. MATLAB documentation recommends that

all input and target values be normalized before introducing the data to the neural

network for training [14]. Most of the training data already had the values in the target

training matrix between zero to one. This data corresponds to the fault classification. But

this is not true for the fault location value. Since the fault location value will be between

0 km to the total length of the transmission line using substation A as a reference, the

fault location will be normalized to the total length of the transmission line. All phases of

this research will normalize the fault location within the target training matrix to the total

length of the transmission line. For this research, the total length of the transmission line

will be 100 km. This will normalize all values in the target training matrix between 0 and

1.

The last development before training any of the ANNs that will be tested for fault

identification is deciding on the training function to use. Multiple training function were

attempted within this phase of the research, but it was decided that the best training

function to use for fault identification using phasor measurements was the Levenberg-

Marquardt algorithm. According to [13], the Levenberg-Marquardt training function is

very well suited for neural network training where the performance index is using the

mean square error (mse). The MATLAB training function for the Levenberg-Marquardt algorithm is: "trainlm". This will complete all pre-training developments for the input and target training data and the ANNs can be developed and trained so that the ANNs can be tested against the developed testing data.

3.4 Development of Testing Data for Single Transmission Line Using Single ANN Approach

While training the ANNs which are used to predict fault identification, there is a performance index calculating the error between the trained ANN output predictions and the actual fault identification values. Since this research is using the Levenberg – Marquardt training function the mean square error (mse) index is used. In most cases the predictability of the fault classification and location from trained ANNs can be estimated to be good or poor performance by reviewing the value of final mse displayed during the training steps. For all phases of this research it was observed that if the mse values was less than or equal to 1e-7, then the fault classification and fault location predictions became close to the actual fault identification. But the predictability of the fault identification should not be based on the mse values alone. Using his approach does not provide any validation on the actual fault identification performance of the ANNs. In order to validate how well the trained ANNs can predict fault identification, the ANNs need to be tested on fault data that is different from the input and target training data that was used to train the neural networks. For the first phase of this research there was only

one testing data set developed. This testing data set varied the fault locations using the

MATLAB random generator (MATLAB function: rand). The fault resistances that were

used when applying the faulted conditions to the transmission line while developing the

testing data were not varied from the fault resistance values that were used to develop the

training data sets. This results in the same six fault resistance values being simulated

while applying the faults to the transmission line (1 Ω, 10 Ω, 20 Ω, 30 Ω, 40 Ω, and 50

Ω). However, the fault locations where varied from the original training data sets. There

were 20 random fault locations used to develop the testing input and target data. Since

there was only one transmission line studied within phase 1 of this research, applying the

ten different faults with the same six fault resistance values, and then using the 20

random fault locations, this created 1200 sets of faulted measurement data within the

testing input and target data set. Since the step sizes used within phase 1 of this research

was so small (0.1 km and 0.05 km), the fault locations that were used for testing

incorporated an accuracy of four decimal points. This would provide confidence that it

would be very unlikely that the same fault location would be used to develop the testing

data set as used in the development of the training data sets.

3.5 Results for Single ANN Approach using Single Transmission Line Model

This section presents the ANN fault identification prediction results for phase 1 of this

research. Just as a recap this phase uses a single ANN approach to predict fault

classification and fault location within the same ANN structure. There were multiple sets

of the single ANN structures developed with the modifications centered around the different number of hidden layers and the number of neurons in the hidden layers. The ANNs were then trained with data that contained different phasor measurement arrangements and different step sizes that moved the faulted conditions down the transmission line. These trained ANNs were evaluated to test how well different ANN architectures with different ranges of input data can predict the fault identification. Retrieving the ANN prediction results began once all the ANN architectures were trained and the complete testing data was gathered. This testing data was then supplied to the trained ANN structures such that each ANN would provide the fault identification predictions. Once the ANN output predictions were obtained, each data set in the actual testing target fault data was compared with the ANN output predictions for fault classification and fault location. During this comparison step, the error difference was collected based on absolute error as shown in equation 3.8.

$$Fault\ Identification\ Error = |Actual\ Fault\ Identification -$$

$$Predicted\ Fault\ Identification| \hspace{4cm} (3.8)$$

Once the absolute errors were calculated for the fault classification and fault location predictions, performance metrics were then determined for each ANN architecture. For the fault classification metrics, the maximum absolute error, minimum absolute error, and average absolute error was determined for all ANN output predictions for each ANN structure developed utilizing the different measurement arrangements being available. As

for fault location the maximum absolute error, minimum absolute error, and average absolute error metrics were also recorded. But there was an additional metric that was calculated for the number of instances of absolute errors that exceeded a threshold value. For phase 1 of this research this threshold value was set to 5 km. This metric was used to determine how many tested scenarios out of the total number of tested faulted conditions where the ANN fault location predictions exceeded 5 km of absolute error. At 5 km, which converts into 3.10686 miles, it was believed that the error becomes too large to provide field personnel with any related fault identification information with high levels of confidence. The selected ANN should have the ability to provide tighter tolerances of absolute error then 5 km when it comes to fault location.

Fault classification absolute error is the difference between the actual discrete value of 0 or 1 for a faulted connection or no-connection versus the ANN output that corresponds to that related phase or ground connection. This use of absolute error would be unitless for fault classification. When calculating the absolute error for fault location, the equation is comparing the difference between the actual location of the fault from a reference substation and the ANNs fault location prediction. Since both inputs into the equation are describing the fault location which is expressed in units of km then the absolute error calculations will be expressed in units of km.

The rest of the information related to this section will discuss the actual results of phase 1. The results will be discussed in sub-sections of the different fault measurement arrangements. The sub-sections will discuss the performance metric of maximum absolute errors. In order to determine how well these ANNs are performing to predict the fault identification, the maximum error needs to be as low as possible. This ensures high

confidence in the ANNs ability to identify the fault type and its location on the transmission line.

There were four sets of ANNs trained which attempt to predict fault identification using the different measurement arrangements. The ANN training parameters that were used to train the network are identified using the following:

- 1 hidden layer of neurons with training data containing fault steps of 0.1 km

- 2 hidden layers of neurons with training data containing fault steps of 0.1 km

- 1 hidden layer of neurons with training data containing fault steps of 0.05 km

- 2 hidden layers of neurons with training data containing fault steps of 0.05 km

3.5.1 Fault Identification Results using Current from Substation A (Phase 1)

When using current phasors from one substation on a single transmission line, there will be six inputs provided to the ANNs for the training process. When the one hidden layer ANN was used to predict the fault identification using current phasor measurements from one substation, the fault classification portion of the results are very positive. Figure 28 provides the fault classification results with using one hidden layer ANN structure trained with current phasors from one substation. This ANN structure was trained with data containing fault steps at 0.1 km. These results show that the fault classification contains high errors at low number of neurons in the hidden layer. But as the number of neurons increase in the hidden layer, the maximum absolute error begins to decrease below 10 percent. From 12 to 36 neurons, if the ANN output is rounded to the nearest whole

number using traditional rounding techniques, the error comparison for every tested

scenario of the fault classification had perfect fault classification prediction. This can be

seen in Figure 28, knowing that the error provided is the difference between the actual to

predicted fault output and knowing the actual fault value is either a value of 0 or 1, the

absolute error has to be greater than 0.5 for traditional rounding to create a miss

prediction for fault classification. At 36 neurons the error gets less than 0.02 or 2 percent

for all phases and ground connection predictions.
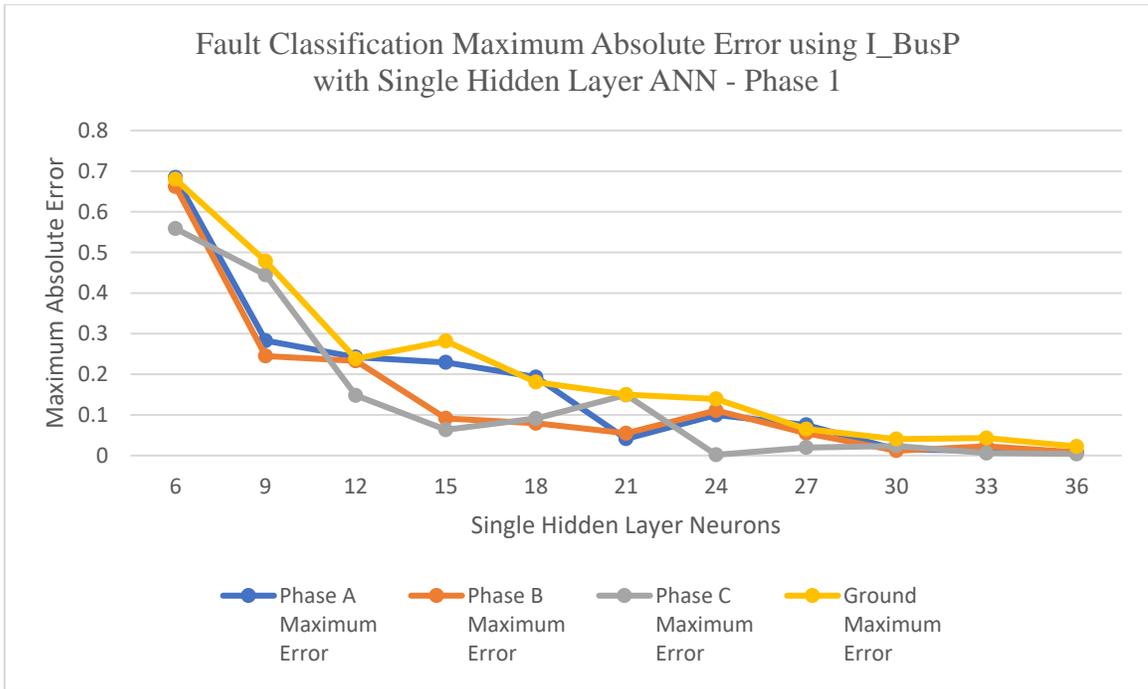
Figure 28 - Fault Classification Maximum Absolute Error using I_BusP with Single
Hidden Layer ANN – Phase 1

When the fault steps are decreased to 0.05 km, the errors are very similar as shown in

Figure 28. If current measurements for one substation are used with a single ANN to

predict fault classification any neuron greater than 24 neurons provides a maximum of 10 percent error or less.

When evaluating the fault location maximum errors, the absolute error gets as low as 9 km. Figure 29 shows the maximum absolute error for fault location as the number of neurons are increased in the hidden layer. The data shows that the best ANN structure for fault location predictions would be for any neurons after 27 neurons. At 27 neurons we begin getting maximum errors near 10 km. Using 27 to 36 neurons in the hidden layer, the ANN predictions for fault location produces a maximum of 34 instances of absolute error over 5 km. This means that only 34 instances of the 1200 tested scenarios provided absolute errors of 5 km or greater.
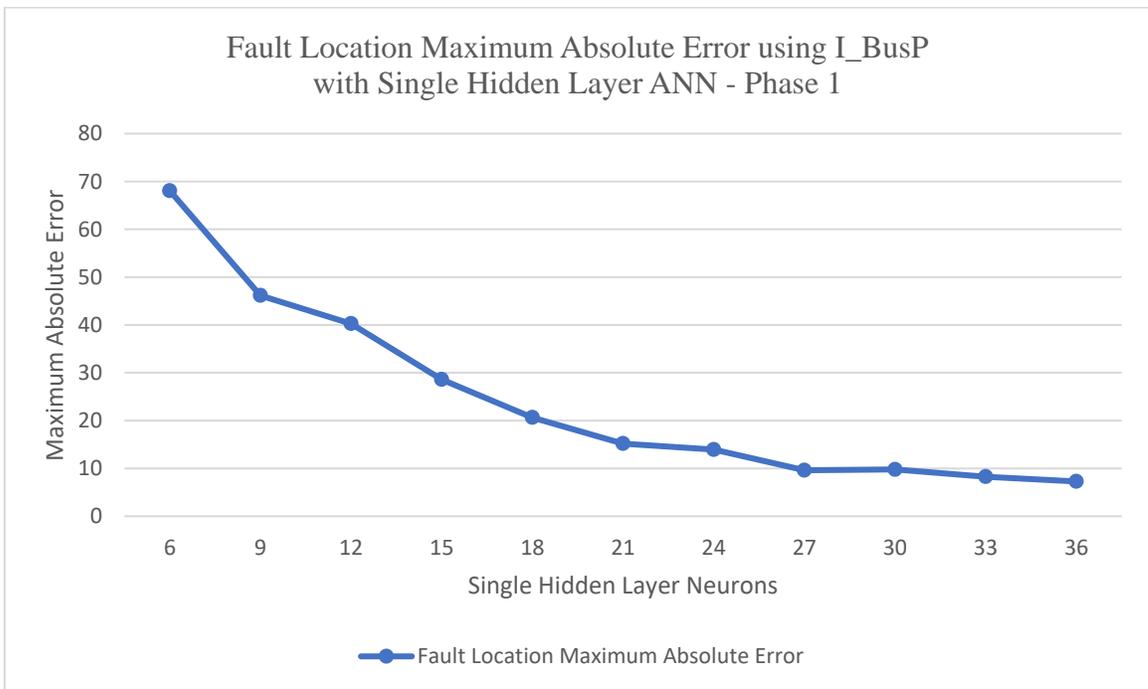
Figure 29 - Fault Location Maximum Absolute Error using I_BusP with Single Hidden Layer ANN - Phase 1

This output data is using ANNs that have been trained with fault steps of 0.1 km. When

ANNs are trained with 0.05 km training data and the results do not improve drastically.

For 24 to 36 neurons in the hidden layer produce a maximum number of 35 testing

instances of absolute error that exceeded the 5 km error threshold.

Looking at the two hidden (multi-layer) neural networks all hidden layer structures

produced absolute errors less than 10 percent for fault classification. It is considered that

any multi-hidden layer structure that was trained in this research is an acceptable choice

for fault classification. If the ANN output is rounded to the nearest whole number using

traditional rounding this will produce no errors in fault classification predictions. Figure

30 provides the fault classification results showing the trends of the maximum absolute

error for each phase and ground connection. As with the single hidden layer ANNs for

fault classification when 0.05 km data was used, the results for the multi-layer ANN did

improve. But the results do not improve drastically to conclude that one set should be

used over the other.
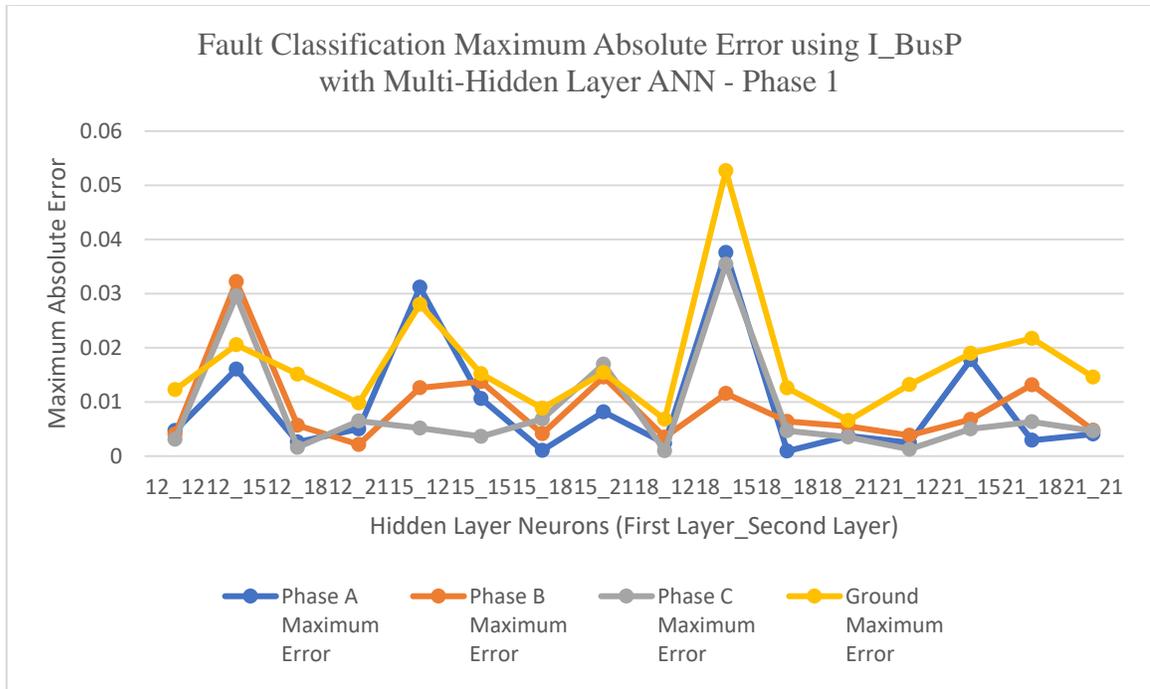
Figure 30 - Fault Classification Maximum Absolute Error using I_BusP with Multi-Hidden Layer ANN – Phase 1

When the multi-hidden layer ANNs are used, the number of maximum absolute errors for fault location is well improved. The highest number of maximum absolute error for any structure was 9.206 km for 18 neurons in the first layer and 15 neurons in the second layer. The maximum error got as low as 3.46 km for 18 neurons in the first layer and 21 neurons in the second layer.
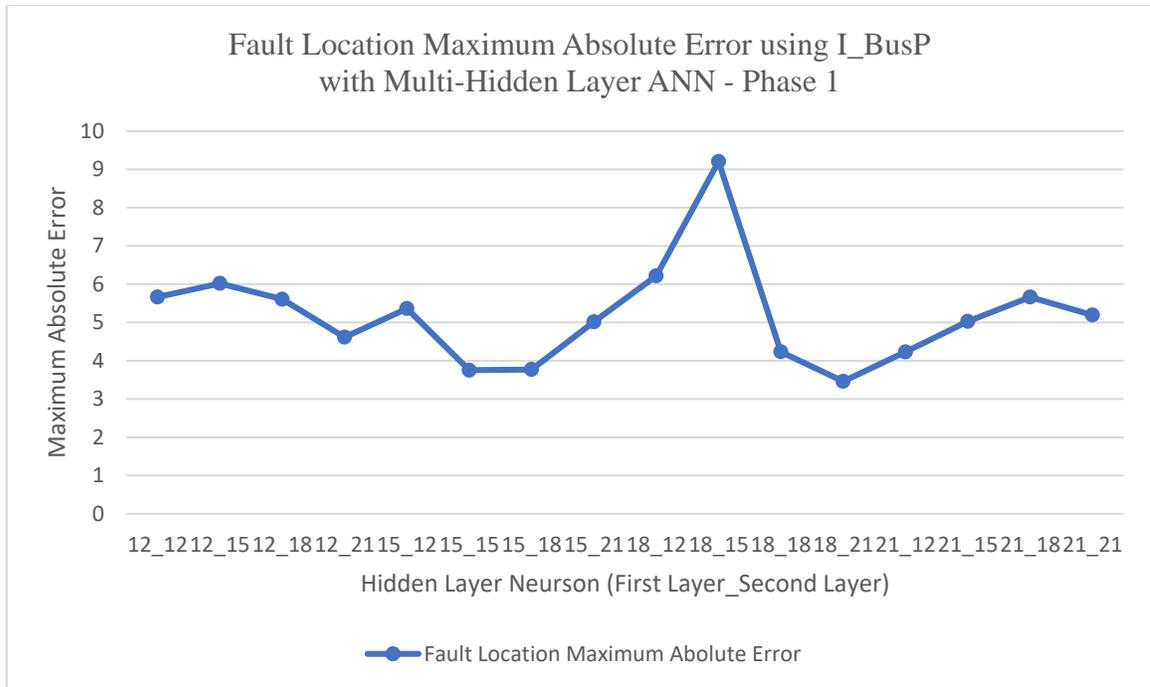
Figure 31 - Fault Location Maximum Absolute Error using I_BusP with Multi-Hidden Layer ANN – Phase 1

There were many ANN structures that were evaluated that lead to the number of maximum absolute errors for fault location being less than 5 km. The multi-hidden layer structure that outperformed any of the other structures tested for fault location would be 18 neurons in the first layer and 21 neurons in the second layer.

3.5.2   Fault Identification Results using Voltage from Substation A (Phase 1)

The next measurement configuration that was evaluated used voltage phasor measurements recorded from only one substation that is connected to the single transmission line. These voltage measurements will contain six values that correlate to

six inputs into the trained ANNs. When single hidden layer ANNs were evaluated for using voltage measurements from one substation it was identified that trends observed with current from one substation also held true. The lower the number of neurons used in the hidden layer the higher the maximum absolute error. But as the neurons increase in the hidden layer the maximum absolute error decreased. As for the fault classification portion of the ANN, the maximum absolute error recorded its lowest values near 10 percent at 30 neurons in the hidden layer. Figure 32 provides the trends for fault classification using voltage from one substation in the single hidden layer ANN. The ANNs used in these results were trained with fault steps of 0.1 km. Using fault steps of 0.05 km did not improve any of the fault classification errors, in fact the results were very similar to ANNs trained with 0.1 km fault steps. Using ANN structures with 18 to 36 neurons in the hidden layer, if traditional rounding is used with the ANN output for the fault classification portion, the ANN would have perfect predictions for all tested scenarios.
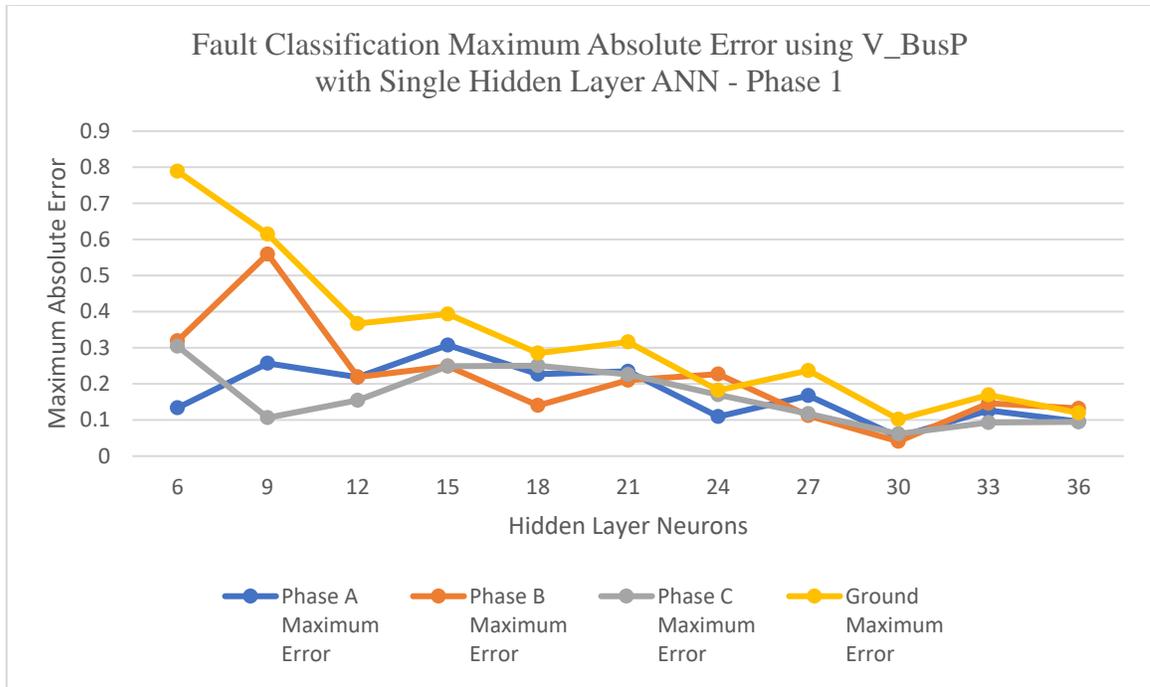
Figure 32 - Fault Classification Maximum Absolute Error using V_BusP with Single Hidden Layer ANN - Phase 1

As for the fault location portion predictions from the tested ANNs, the maximum absolute errors were very high in lower number of neurons used in the hidden layer. These absolute errors improved drastically by increasing the number of neurons in the hidden layer. The maximum absolute errors hit the lowest value of 15.276 km at 30 neurons in the hidden layer. During that ANN structure it was identified that the ANN predictions resulted in 117 instances out of 1200 tested scenarios having absolute errors over 5 km. Even though this seems like a high value of instances, it is only around 10 percent of the tested scenarios and considered a low probability of occurrence. Figure 33 shows the fault location maximum absolute errors that were recorded using voltage

phasors from one substation in a single hidden layer ANN structure. This data was

recorded on ANNs that were trained with 0.1 km fault steps.
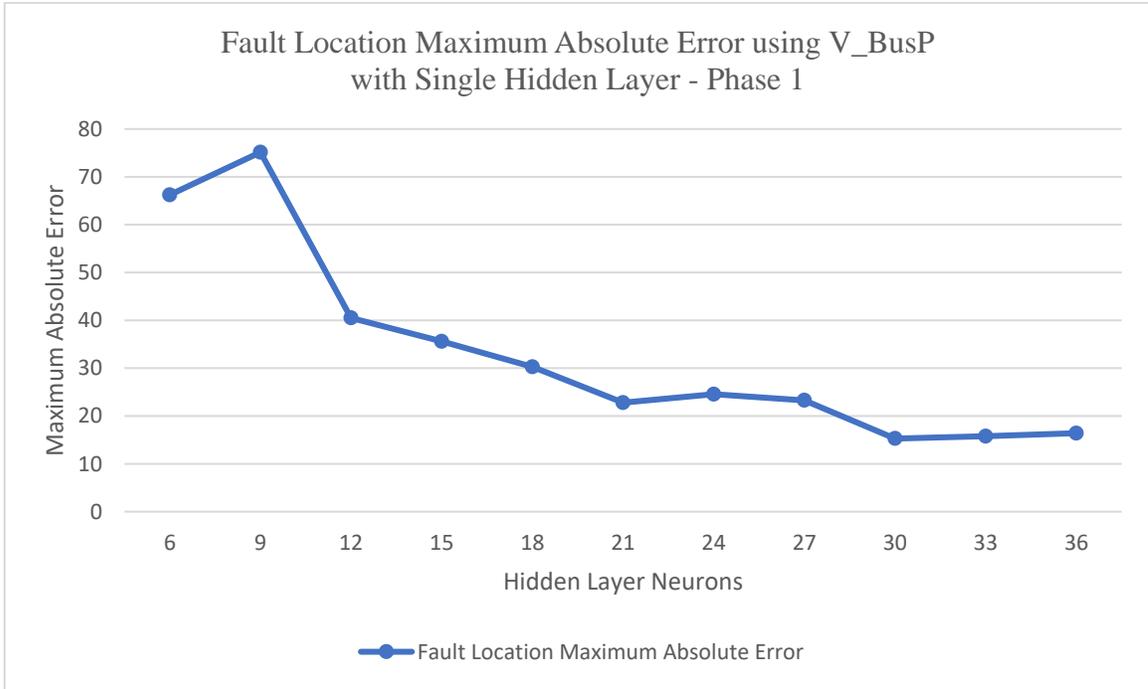
Figure 33 - Fault Location Maximum Absolute Error using V_BusP with Single Hidden
Layer ANN - Phase 1

When multi-layer neural networks were introduced with voltage measurements from one

substation, the ANN predictions seemed to improve. All ANN structures that were tested

produced maximum absolute errors less than 10 percent for fault classification. This is

deemed acceptable for all tested hidden layer scenarios that were used to predict fault

classification. The lowest maximum error results were obtained with the hidden layer

containing 18 neurons in the first layer and 15 neurons in the second layer.
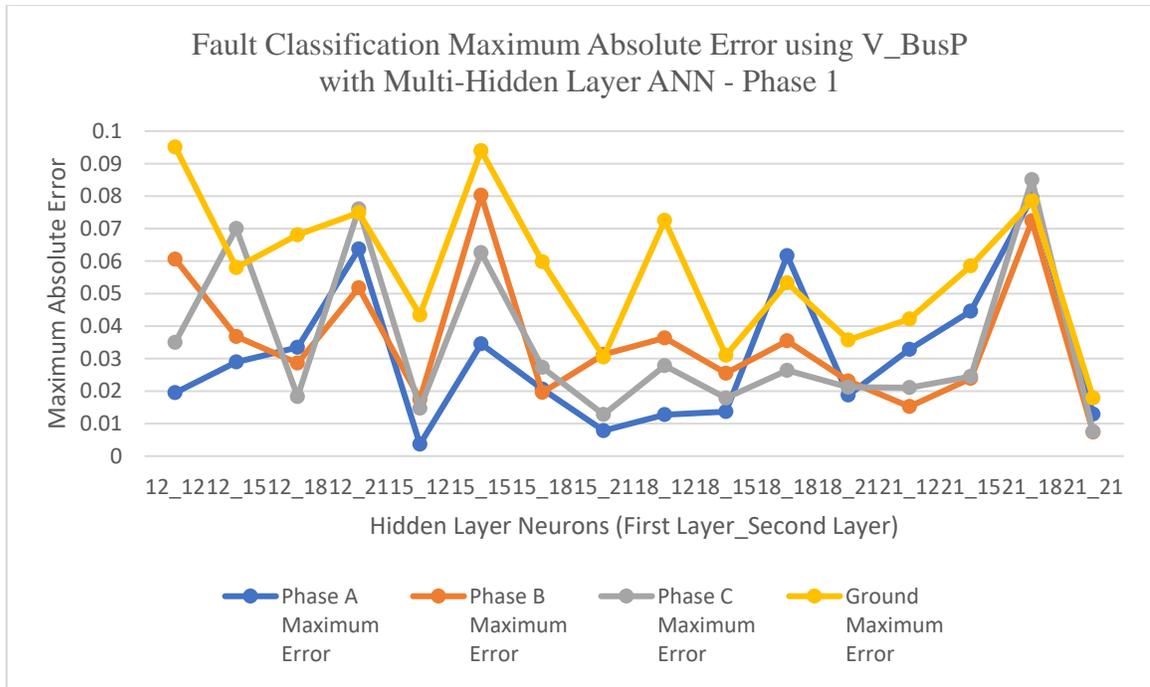
Figure 34 - Fault Classification Maximum Absolute Error using V_BusP with Multi-Hidden Layer ANN – Phase 1

The maximum error recorded for fault location reduced when using multi-hidden layer ANNs. The lowest maximum error recorded for fault location was 6.682 km within an ANN structure that contained 21 neurons in both the first and second layers. This ANN structure produced only 3 instances of 1200 faulted test scenarios over the 5 km absolute error threshold that was set. This was a huge improvement over the single hidden layer structure. There were in fact multiple ANN multi-hidden layer structures that produced less than 10 instances over the 5 km threshold. Figure 35 provides the maximum absolute errors for fault location using a multi-hidden layer ANN trained with voltage phasors from one substation.
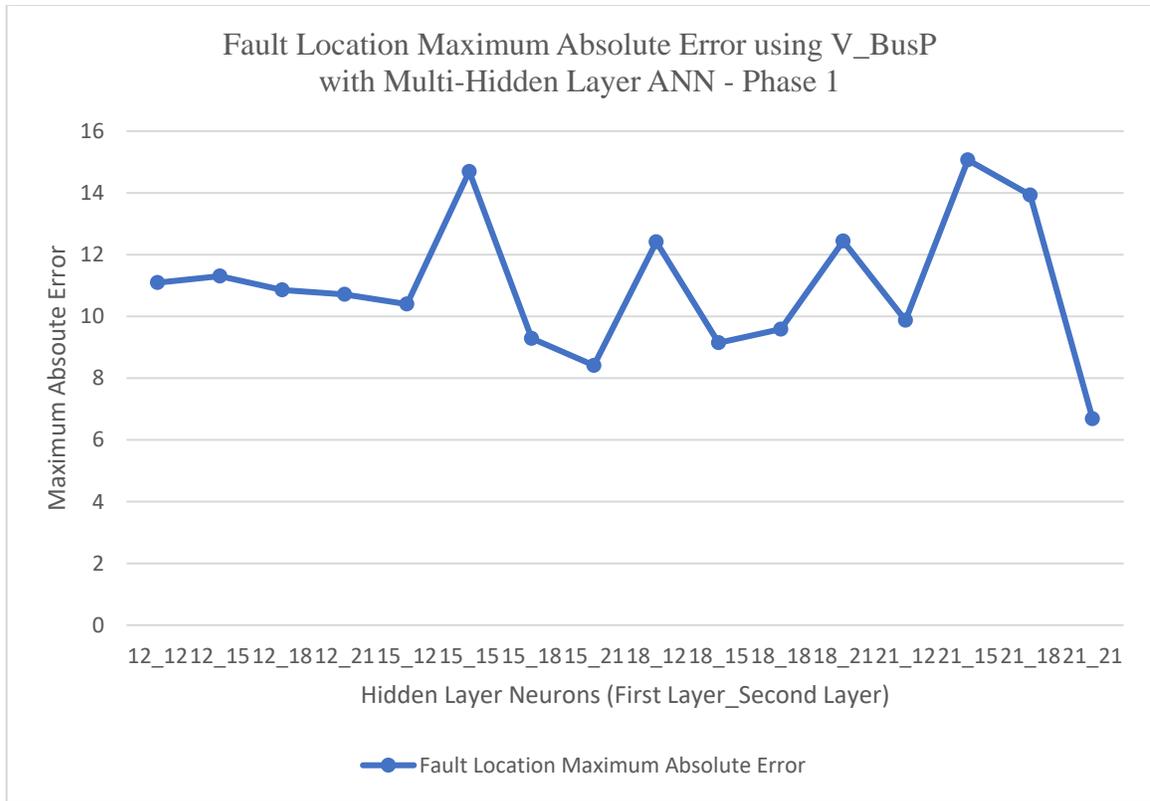
Figure 35 - Fault Location Maximum Absolute Error using V_BusP with Multi-Hidden Layer ANN - Phase 1

3.5.3 Fault Identification Results using Voltage and Current from Substation A (Phase 1)

The single ANN approach continues by evaluating how each ANN structure developed improves when there is more input data available to train the neural network. This section will evaluate the ability to perform ANN predications with voltage and current phasors being available at one substation that is connected to a transmission line that is faulted. Having both voltage and current phasors available increases the number of inputs into the ANN from six to twelve.

This analysis begins by reviewing the single hidden layer ANN prediction results. The

fault classification portion of the single ANN approach results are very similar to the

multi-hidden layer ANNs with only voltage or current phasors from one substation being

presented to the ANN. It was observed that 12 neurons and beyond in the hidden layer

will produce maximum absolute error near 10 percent for fault classification. There were

a couple of ANN architectures that performed extremely well with relatively low

maximum absolute errors, but when the hidden layer contains 33 neurons the ANN fault

classification prediction produces the best results. Figure 36 presents the fault

classification data using voltage and current from one substation using the single hidden
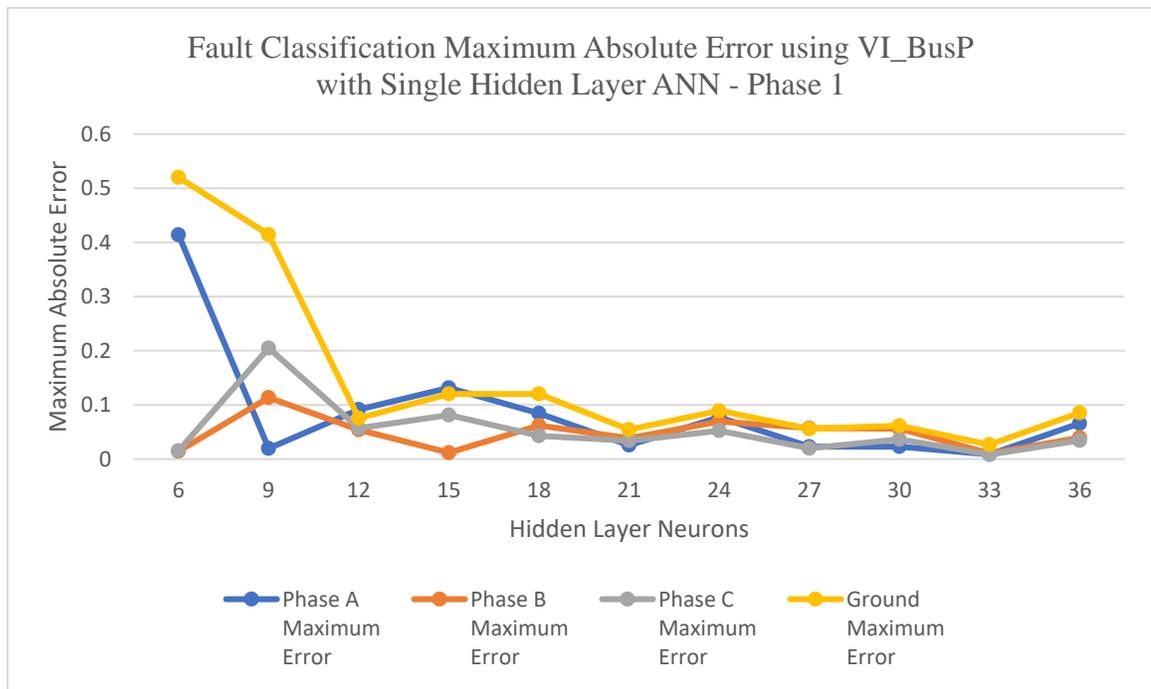
layer ANNs.

Figure 36 - Fault Classification Maximum Absolute Error using VI_BusP with Single
Hidden Layer ANN – Phase 1

Reviewing the fault location errors when using single hidden layer ANNs, the maximum absolute errors are still decreasing while the number of neurons increase in the hidden layer. The only hidden layer structure that did not obey this trend was 36 neurons in the hidden layer. The lowest maximum error occurred with 33 neurons in the hidden layer with 6.1 km being the value. At 33 neurons in the hidden layer the ANN produced only 4 instances out of 1200 faulted test scenarios being over the 5 km threshold. Having more measurement data available seems to produce better results with simpler networks. Figure 37 provides the maximum absolute error results for fault location using voltage and current from one substation using single hidden layer networks.
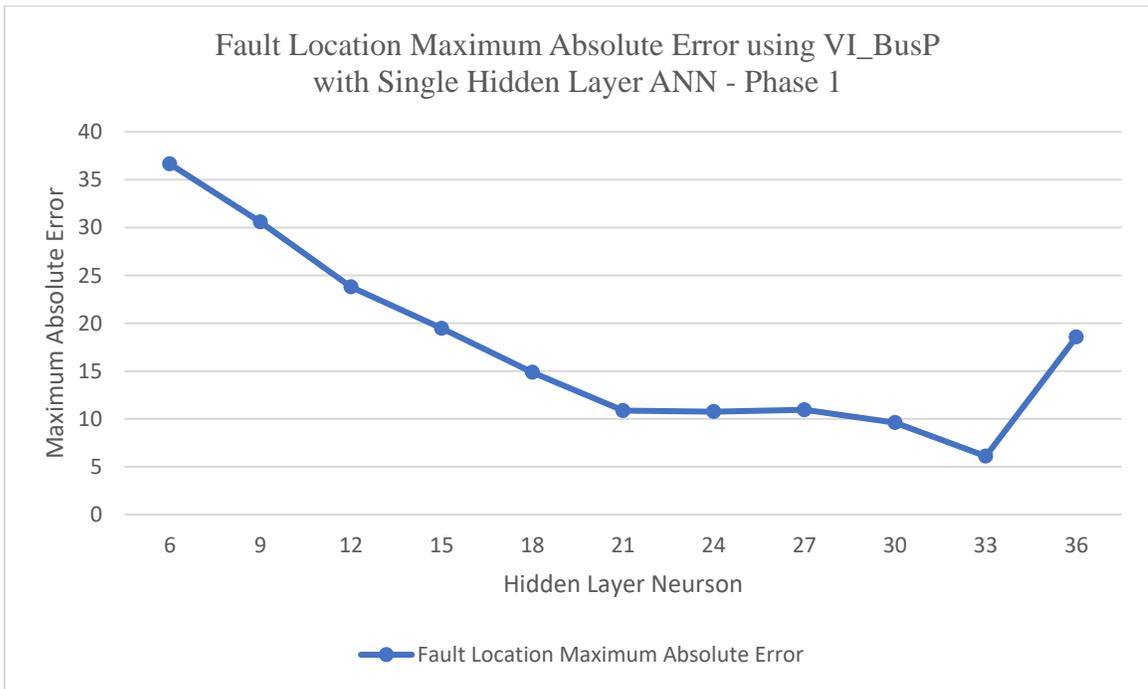
Figure 37 - Fault Location Maximum Absolute Error using VI_BusP with Single Hidden Layer ANN – Phase 1

When multi-layer networks are used with voltage and current phasors from one

substation, most ANN structures produce of the maximum absolute errors less than 4

percent on all phase and ground connections for fault classification. The strongest

performing ANN for the fault classification portion was 21 neurons in both the first and

second hidden layers. This structure produced maximum errors of 0.8 percent for all

phases and ground connections. Figure 38 provides the maximum absolute errors

recorded for fault classification for each phase and ground connection for each hidden

layer structure that was tested.
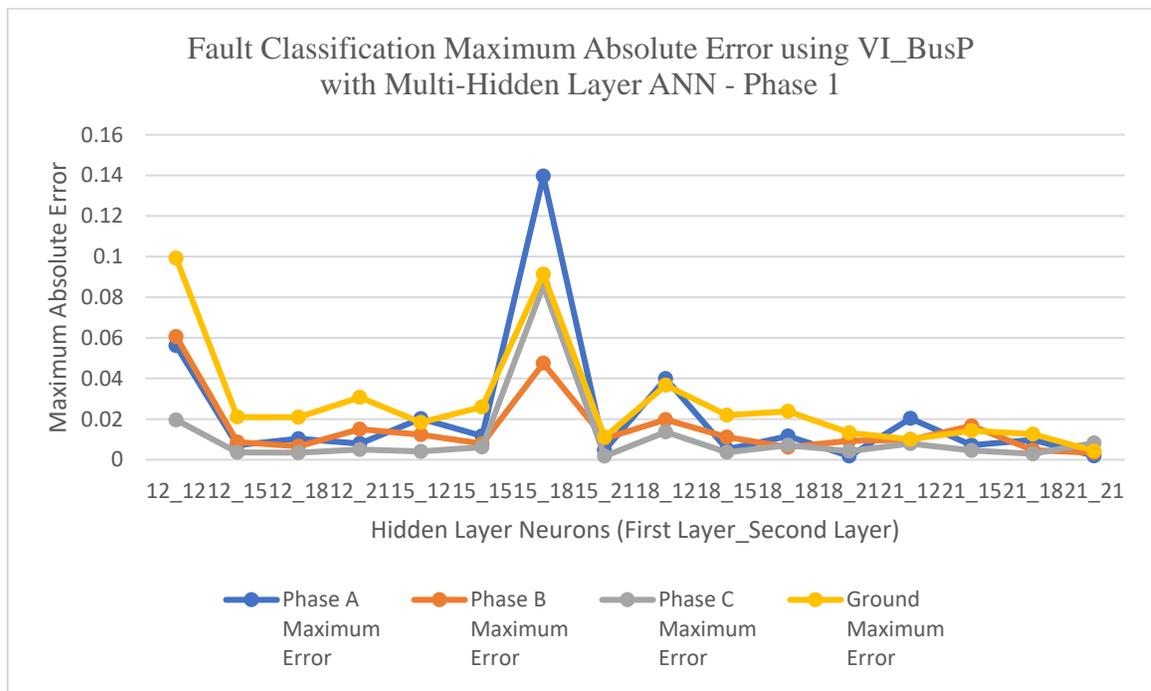
Figure 38- Fault Classification Maximum Absolute Error using VI_BusP with Multi-Hidden Layer ANN – Phase 1

As with the fault classification, the same trend continues to hold true with the fault location maximum absolute errors. The lowest maximum error occurred on the hidden layer neuron structure was 21 neurons in both hidden layers at a maximum error of 2.326 km. Since this maximum absolute error value is less than the 5 km threshold, there were no tested instances over the threshold value of 5 km. Figure 39 provides the fault location maximum absolute errors recorded using the multi-hidden layer ANNs as discussed.
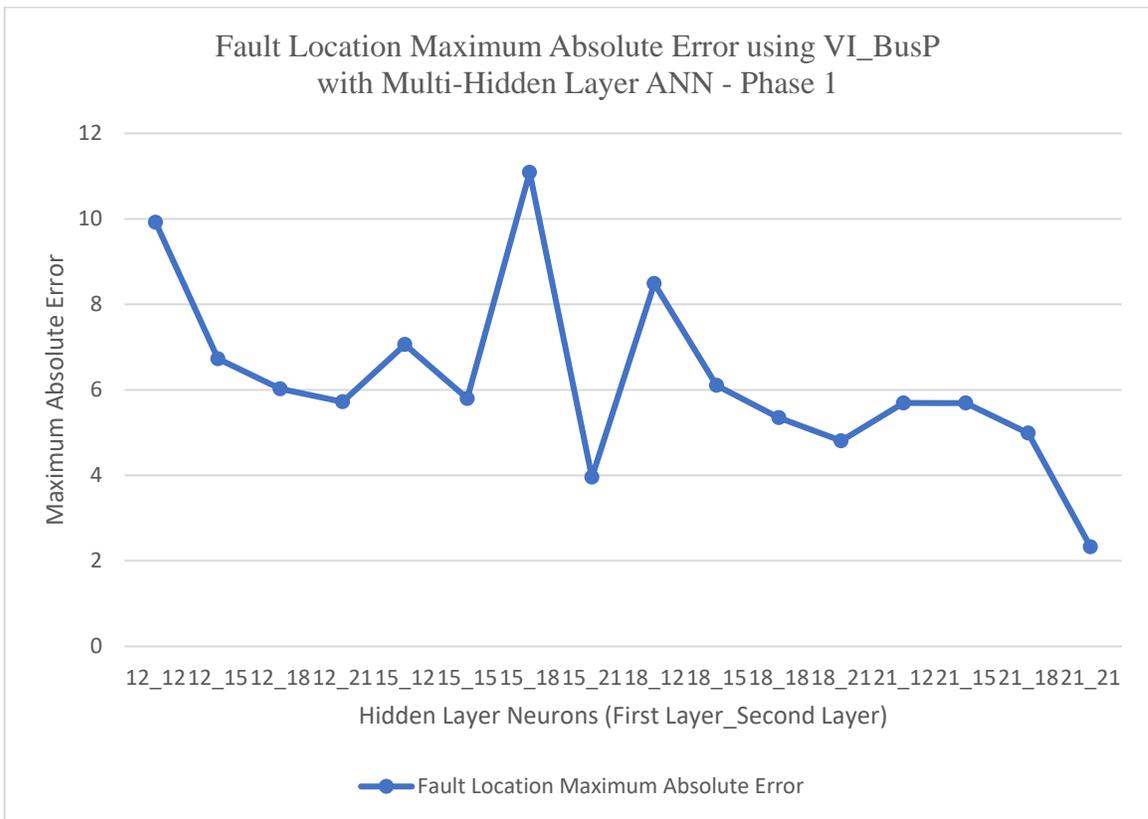
Figure 39 - Fault Location Maximum Absolute Error using VI_BusP with Multi-Hidden Layer ANN – Phase 1

3.5.4 Fault Identification Results using Voltage and Current from Substation A and Substation B (Phase 1)

The last measurement configuration that was evaluated for this phase was using voltage and current phasors from both substations connecting the transmission line. Using measurements from both substations will produce 24 inputs into the ANNs. It is also expected before any analysis was evaluated that this measurement configuration would produce the closest ANN prediction results. When the single hidden layer ANN was used to produce fault classification ANN predictions, it was observed that the maximum absolute error for most of the ANN structure contained errors at less than 5 percent. Only looking at hidden layer structures of 24 to 36 neurons, the maximum absolute error was less than 3 percent. The largest error occurs at 6 neurons with the error less than 20 percent. Therefore, if traditional rounding is used there would be perfect fault identification predictions for all tested scenarios. Using the ANN structure of 27 neurons in the hidden layer would produce the lowest maximum error. For the fault classification portion of the single ANN approach using voltage and current phasors from both substations, using ANNs that have been trained with fault data of 0.05 km fault steps seems to produce lower errors at lower hidden layer neurons. The lowest maximum errors recorded using ANNs trained with 0.05 km fault steps is near 1.5 percent error at 27 neurons in the hidden layer. Figure 40 provides the ANN fault classification results using voltage and current measurements from both substations in a single hidden layer ANN trained with 0.1 km fault step training data.
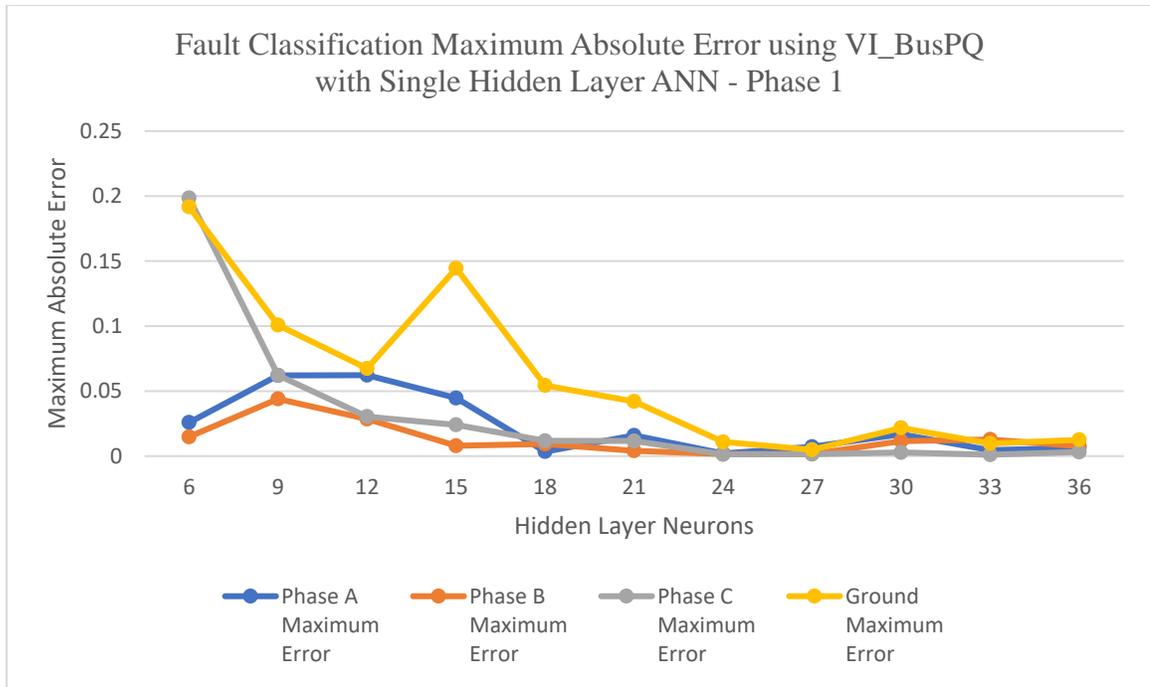
Figure 40 - Fault Classification Maximum Absolute Error using VI_BusPQ with Single Hidden Layer ANN – Phase 1

The maximum absolute error in the fault location portion of the trained ANN predictions are considerable low compared to the other measurement configuration results. The lowest maximum absolute error recorded was at 27 neurons in the hidden layer with a value of 0.56 km. This produces no instances of the 1200 faulted test scenarios over the threshold value of 5 km. Figure 41 provides the ANN fault location results using voltage and current measurements from both substations in a single hidden layer ANN.
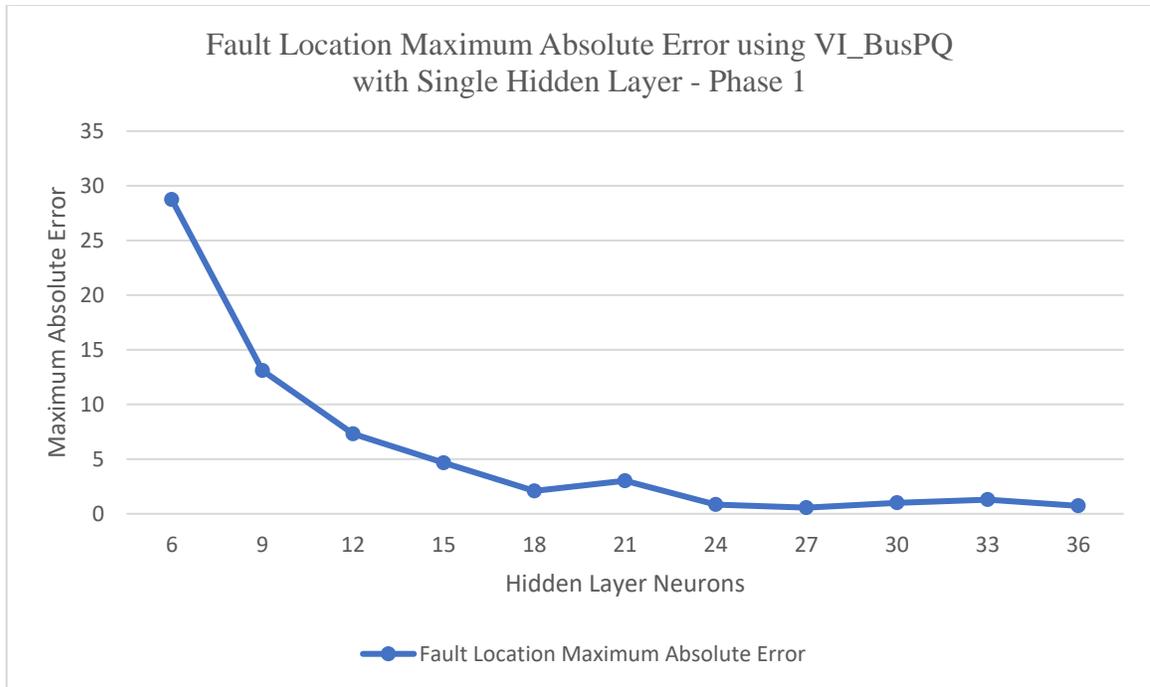
Figure 41 - Fault Location Maximum Absolute Error using VI_BusPQ with Single
Hidden Layer ANN – Phase 1

Using multi-layer hidden ANNs to perform fault identification with voltage and current

phasors from both substations only reduces the absolute error even more. The fault

classification portion of the ANN produces maximum absolute errors at less than 1

percent for all phases and ground connections. The best performing multi-layer ANN

structure for fault classification is 21 neurons in the first hidden layer and 12 neurons in

the second hidden layer. This hidden layer structure contains maximum errors near 0.02

percent for all phases and ground connections. The fault location maximum absolute

error contains values as low as 0.621 km. This absolute error happens with an ANN

structure of 21 neurons in the first hidden layer and 12 neurons in the second hidden

layer.

Chapter 4 – Fault Identification with Single Transmission Lines using Multiple ANN Approach

This chapter discusses the second phase of this research. During the second phase, it was determined to analyze the effects of how the ANN fault identification predictions would differ if multiple ANNs were used versus the single ANN approach used within phase 1. This approach will use a single ANN to identify the fault classification. Then there would be a set of four distinctly different ANNs used to predict the fault location. These four different ANNs are developed and trained based on the four type of faults (LG, LL, LLG, LLL). The drive behind this phase of the research was that the input and target training data are trying train both aspects of fault identification, which are two complex problems that may provide better ANN predictions if the two problems are separated. This could possibly result in the trained ANNs used within phase 1 to lean toward the ANN output predictions having poor accuracy. The idea of splitting up the algorithm by using multiple ANNs to predict fault identification is that the results would become more accurate and satisfactory to determine the fault classification and location of the fault. By providing more satisfactory results, it provides higher levels of confidence to dispatch fault identification information to the field personnel. This process would force each of the trained neural networks to have a simpler and more direct prediction task. The information needed to train each neural network would be less diverse. The discussed process within this chapter can be visualized by the provided flow diagram shown in Figure 42.
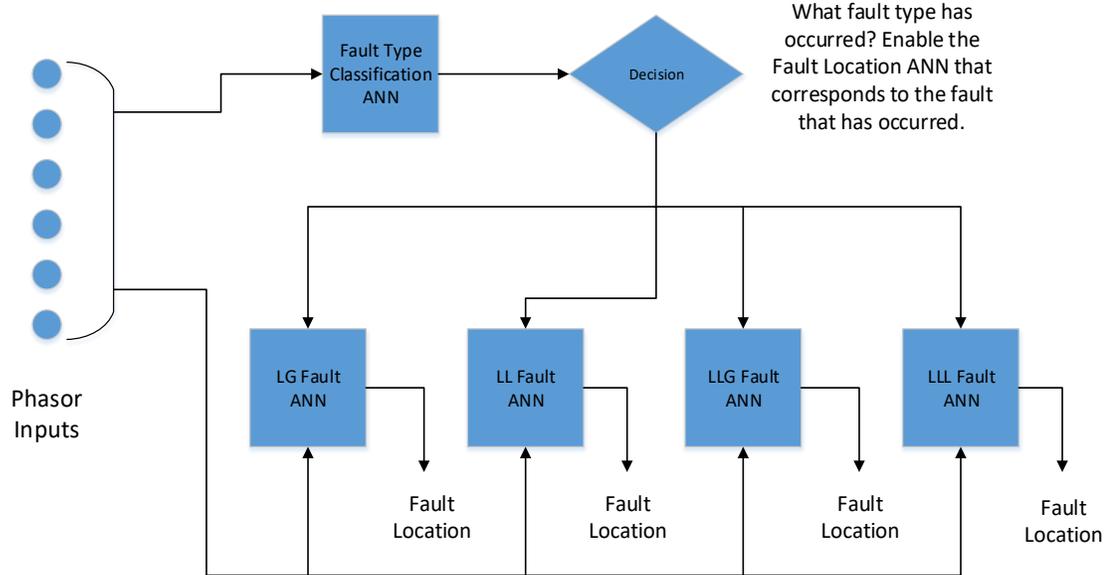
Figure 42 - Fault Identification (Phase 2) Flow Diagram

Within this flow diagram the inputs, which would be the recorded phasor measurements, are used as the inputs into the into the five different neural networks. The process will begin by analyzing the input fault training data to determine the fault classification. The fault classification ANN will describe the fault type as one of the ten different fault types. Based on the ANN output for the fault classification, a decision is made to determine which generic fault category line to ground (LG), line to line (LL), double line to ground (LLG), or a three-phase fault (LLL) the fault classification falls into. Once one of the four generic fault categories have been predicted by the fault classification ANN, an enable bit is set on one of the four fault location ANNs that is associated with the predicted fault category. Once this fault location ANN has been enabled, the neural network can then output the predict fault location.

4.1 Two-Terminal Single Transmission Line Model for Phase 2

The transmission line topology that is being studied in this phase of the research is the

two-terminal single transmission line. The same transmission line model that was used

within the phase 1 was also used during this phase of the research. There were some

changes/modifications to the model which will be discussed during this section. All

details related to the development of the transmission line model that is not discussed in

this chapter should have been covered in chapter 3 under section 3.1.

The first adjustment related to the transmission model was the version of the MATLAB

and Simulink software that was used. Phases 2 and 3 of this research will utilize the

2019a version of the MATLAB and Simulink software. It was discovered that there were

some limitations with the 2016a student version of MATLAB and Simulink that was

being used. While reviewing the transmission line model with the newer version of

MATLAB and Simulink software, it was identified that there was an updated version of

the Fourier analysis block. The transmission line model was updated to include the new

Fourier analysis block.

While performing some initial testing on the single transmission line model using the

new version of MATLAB and Simulink, it was identified that applying a fault close to

either substation A or substation B caused some harmonic frequencies to be imbedded

into the voltage and current waveforms from the time the fault was applied to the

transmission line at 2nd cycle. This imbedded harmonic content is still present in the

waveforms out past the fifth cycle of simulation time. As an example, Figure 43 shows a

voltage waveform with the induced harmonic content. This waveform was simulated

using the transmission line model with an LG fault applied to the transmission line at 3.5

km from substation A while applying the fault with a fault resistance of 1 Ω.
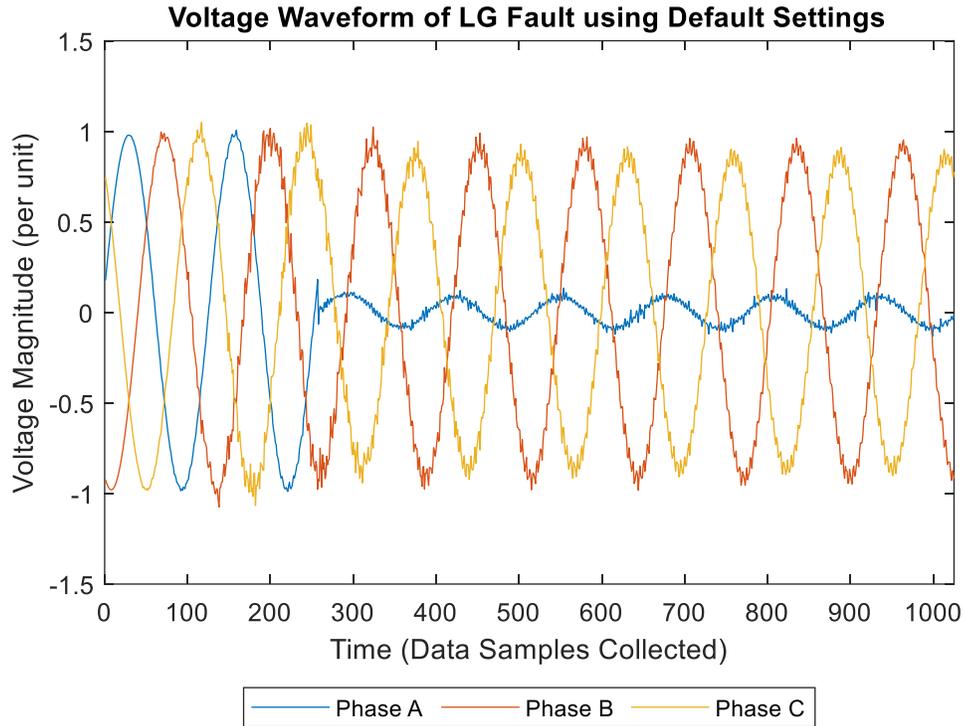
Figure 43 – Single Transmission Line Voltage Waveform for LG Fault using Default
Settings

This voltage waveform should have a smooth sixty (60) hertz waveform to retrieve an

acceptable faulted value at 5 cycles after the fault. Figure 44 provides a similar voltage

waveform that is simulating a phase A to ground (LG) fault. This fault has also been

applied at 3.5 km from substation A with a fault resistance of 1 Ω. The difference in this

waveform has the relative tolerance setting adjusted from the default value of 1e-4 to 1e-

7. Figure 44 shows that the harmonic content is still imbedded in the faulted waveform, but eventually dies out of the signal by the fifth cycle of simulation.

Voltage Waveform of LG Fault using Relative Tolerance = 1e-7
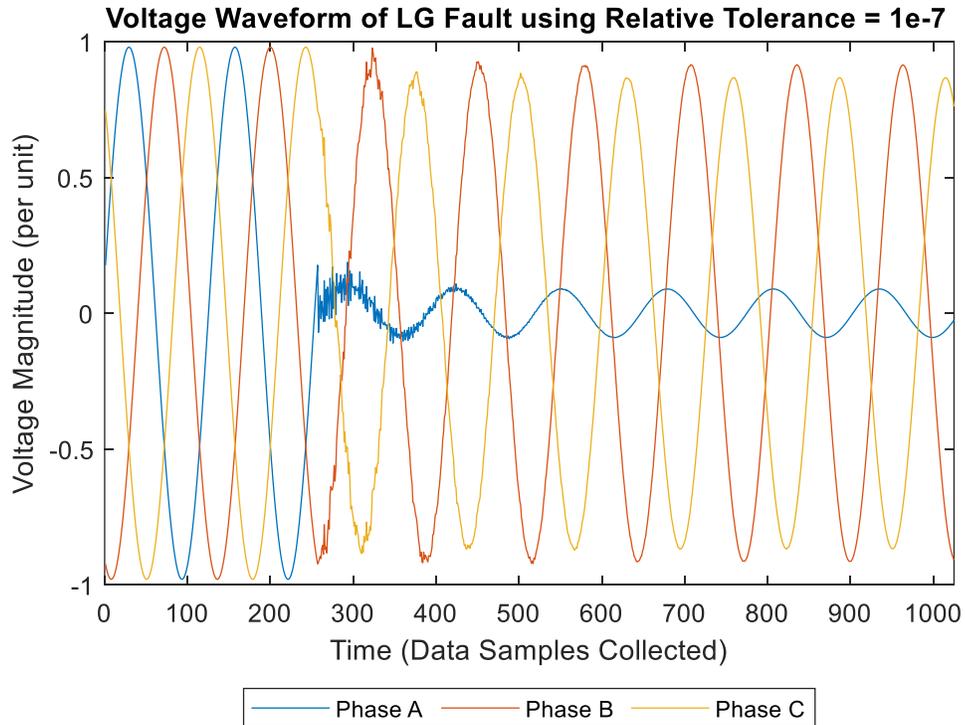
Figure 44 – Single Transmission Line Voltage Waveform for LG Fault using Relative Tolerance = 1e-7

Since this research concentrates on using the data point at the beginning of the 5th cycle of simulation time this became a data concern. This caused enough concern when it was identified that the ANN would have a hard time predicting the fault classification or fault location that investigation on how to fix the waveform data was researched. In order to evaluate the harmonic imbedded content, it was identified that the relative tolerance of the Simulink simulation should be evaluated for faults applied to the transmission line.

The relative tolerance is a parameter located in the Simulink software within the "Configuration Parameters". The relative tolerance specifies the largest acceptable solver error, relative to the size of each state during each time step. If the relative error exceeds this tolerance, the solver reduces the time step size [36]. For the transmission line model, the relative tolerance default value is set to 1e-4 or 0.01%. The relative tolerance value used within this setting was adjusted during the development of the input and target training data process before the fault was simulated on the transmission line. The adjustments of the relative tolerance value will be discussed during section 4.2.

4.2 Development of Input and Target Training Data for Single Transmission Line using Multiple ANN Approach

The process of deriving the input and target training data was not much different than the process that was described within chapter 3. Phase 2 still simulates faulted conditions on the Simulink transmission line model to obtain the faulted current and voltage phasor data. But instead of varying the step size that moves the fault down the transmission line as in phase 1, the faulted conditions were moved down the transmission line at a static step size of 0.1 km.

The Simulink transmission line model was still simulated for a total of eight cycles where the fault was applied to the model at beginning of the second cycle. One of the modifications of collecting the training data during this phase of the research was that the initial and final simulated fault locations were adjusted. The initial fault location was set

to 3 km from substation A and the final fault location was set to 97 km from substation

A. During the evaluation process for the harmonic content imbedded into the

measurement waveforms, no clear (smooth) waveform could be found when the fault got

closer than 3 km to either substation A or substation B. Due to this limitation, the training

input and target data along with the testing data was limited to faults no closer than 3 km

from either substation.

While evaluation of the harmonic content within the voltage and current waveforms,

faults were applied to the transmission line model with different fault locations, different

fault resistances, and different fault types to review the waveforms by ranging the relative

tolerance value. The scope block was used to review the voltage and current waveforms

while adjusting the relative tolerance value. This analysis was completed by applying

faults to the transmission line with the ten different fault types and ranging the fault

resistance values between 1 Ω to 50 Ω by taking steps of 10 Ω within that specified

range. Faults were applied to the transmission line at different locations of increments of

5 km and start applying different relative tolerances to the model. This analysis began

with the relative tolerance at the default setting of 1e-4 and was ranged to 1e-7. A trial

and error approach continued until a recommendation combination of fault location to

relative tolerance could be meet. Table 12 provides a match list for the range in fault

location from substation A and the relative tolerance setting in Simulink.

Table 12 - Relative Tolerance Settings for Single Transmission Line Model

Relative Tolerance Settings for Single Transmission Line Model	
Fault Location Distance	Relative Tolerance Setting
3 km to 5 km	1e-7
5 km to 10 km	1e-6
10 km to 18 km	1e-5
18 km to 82 km	1e-4

After determining the modified initial and final fault locations that would be used for the input and target training data, the equation that calculates the number of simulated fault location was adjusted to compensate for the change. Again, the number of simulated fault locations is the equation that defines the number of transmission faults to be applied for each faulted scenario. This adjusted number of transmission faults to be applied to the transmission line is presented in equation 4.1.

$$Number\ Faults\ per\ Fault\ Scenario =$$

$$floor\left(\frac{Final\ Fault\ Location - Initial\ Fault\ Location}{Step\ Size}\right) + 1 \qquad (4.1)$$

Since, it is known that the final fault location is 97 km, the initial fault location is set to 3 km, and the step size of moving the fault down the transmission line is set to 0.1 km, the result of equation 4.1 becomes 941 simulated fault locations per faulted scenarios. This assumes that the transmission line being modeled is a 100 km transmission line.

The fault resistance value that was set in the three-phase fault block that was applied to the transmission model did change with this phase. Instead of only using a fault

114

resistance range of 1 Ω to 50 Ω where the trained fault resistance values used multiples of

10 Ω between the limits of the fault resistance range, the multiples of the fault resistance

values were varied in this phase of the research. There were three different sets of fault

resistance values that were used during the training for all fault identification ANNs.

These three sets of fault resistances stepped between the limits of the fault resistance

range (1 Ω to 50 Ω) with multiple of 10 Ω, 5 Ω, and 2.5 Ω. This process had a similar

purpose to varying the step size of the fault moving down the transmission line used in

phase 1, which increases the size of the input and target training data. Increasing the

number of training data provides the ability to analyze how the training data sets improve

the ANN predictability for fault identification.

Since the equation determining the number of faults that are applied to the transmission

line per fault condition was modified from phase 1, equation 3.3 needs to be modified to

correctly calculate the total number of input and target training sets that contains all fault

types.

$$Number\ of\ Training\ Data\ Sets = (\#\ of\ Parallel\ Lines)\ *$$
$$(\#\ of\ Fault\ Classifications)\ *\ (\#\ of\ Fault\ Resistances)\ *$$
$$(\#\ of\ Fault\ Applied\ to\ Transmission\ Line) \hspace{3cm} (4.2)$$

Knowing the different sets of fault resistances, the number of faults applied to the

transmission line per faulted scenario, and knowing that the simulation will be provide

input and target training data for the ten different fault types, the number of training input

115

and targets data sets can be determined. The number of training data sets within the input

and target training data for each fault data setups is presented in Table 13.

Table 13 - Number of Training Data Sets Used for ANN Training (Phase 2)

Number of Training Data Sets Used for ANN Training (Phase 2)		
Simulated Fault Resistances	Fault Location Step Size	Number of Training Data Sets
1Ω, 10Ω, 20Ω, 30Ω, 40Ω, 50Ω	0.1 km	56,460
1Ω, 5Ω, 10Ω, 15Ω, 20Ω, 25Ω, 30Ω, 35Ω, 40Ω, 45Ω, 50Ω	0.1 km	103,510
1Ω, 2.5Ω, 5Ω, 7.5Ω, 10Ω, 12.5Ω, 15Ω, 17.5Ω, 20Ω, 22.5Ω, 25Ω, 27.5Ω, 30Ω, 32.5Ω, 35Ω, 37.5Ω, 40Ω, 42.5Ω, 45Ω,47.5Ω, 50Ω	0.1 km	197,610

Related to collecting the input data, the last modification for the second phase of this

research was looking at the time constraint it took to gather the input and target training

data. The time parameters within the automation MATLAB code, recorded the amount of

time it took the automation to collect the training data in terms of the number of seconds,

number of minutes, and the number of hours it took for the automated program to

simulate all faulted scenarios. Table 14 correlates the number of training data sets that

were collected for each faulted scenario (Table 13) to the amount of time it took for the

program to simulated all the faulted scenarios.

Table 14 - Time Requirement to Collect Training Data (Phase 2)

Time Elapsed to Collect Training Data Sets (Phase 2)	
Number of Training Data Sets	Time Elapsed to Collect Data (Hours)
56,400	43.0717
103,400	80.4317
197,400	152.8447

The first thing that was observed from the amount of time it took to gather input and target data was that collecting this input and target training data is time intensive. From the data shown in Table 14, the data collection took just shy of two days and up to just over six days to complete the data collection. This time constraint is proportional to the number of faulted scenarios that the user wants to collect data from.

The training input and target data will be collected as described within chapter 3. But just to recap this data collection process, electric faults were applied to the single transmission line model for different combinations of fault types and fault resistances. These fault combinations began at an initial location on the transmission line and moved down the transmission line until the final fault location was reached. All the fault data was then collected into five different matrices that correspond to voltage magnitudes, voltage phase angles, current magnitudes, current phase angles, and the corresponding target (actual) fault data. These five matrices are the full set (containing all fault types) of fault data. At this point the target data will be collected as described in chapter 3.

For phases 2 and 3 of this research, the full set of data is to only be used to train the fault classification ANNs. The fault classification ANNs are going to take the fault measurement data and predict the type of fault that has occurred. Based on the ANN classification output, only one of the four different fault location ANNs would be enabled

to perform the fault location prediction. Since the fault location will be predicted based on the fault type, the full set of data cannot be used to train the four different fault location ANNs. Recall that the four distinct fault location ANNs that are used within this phase will relate to the four different fault classification categories (LG, LL, LLG, LLL). In order to train the fault location ANNs, the full set of input and target training data was separated into fault category sets of data. This will result in four additional sets of data being available for voltage magnitude, voltage phase angles, current magnitudes, current phase angles, and associated target data based on LG, LL, LLG, and LLL fault categories. This process will be known as parsing the fault training data.

4.3 Training the Multiple ANN Approach for the Single Transmission Line Model

Now that the process for parsing the fault training input and target data has been completed, the next step in the process of identifying fault identification is designing and training the fault classification and different fault location ANNs. It was decided that in phase 2 of this research that fault classification and fault location ANNs would only use single hidden layer networks. With each network attempting to solve a more direct problem as compared to the first phase of this research it was assumed that the neural network designs could be simpler in design. Figure 45 and Figure 46 provide the high-level layouts or visual representations for the fault classification and fault location ANN structures. The fault classification and fault location ANNs can be differentiated by

evaluating the output layer. Fault location ANNs will only contain one output neuron as compared to the fault classification ANN what will contain four output neurons.
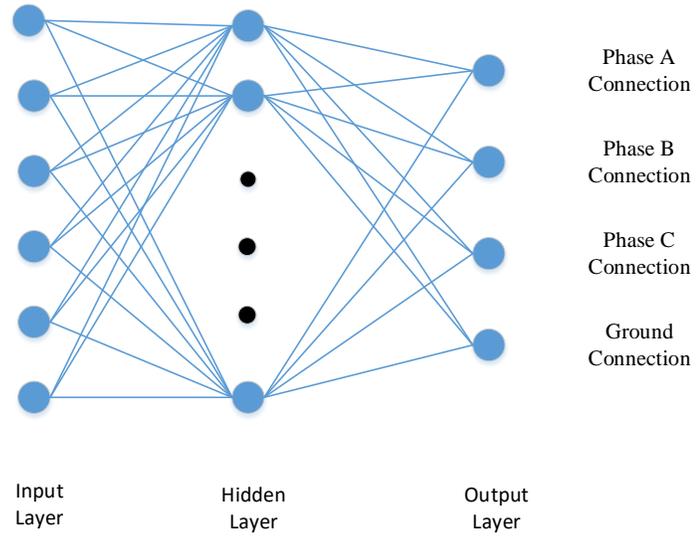
Figure 45 – Fault Classification Artificial Neural Network Structure (Phase 2)
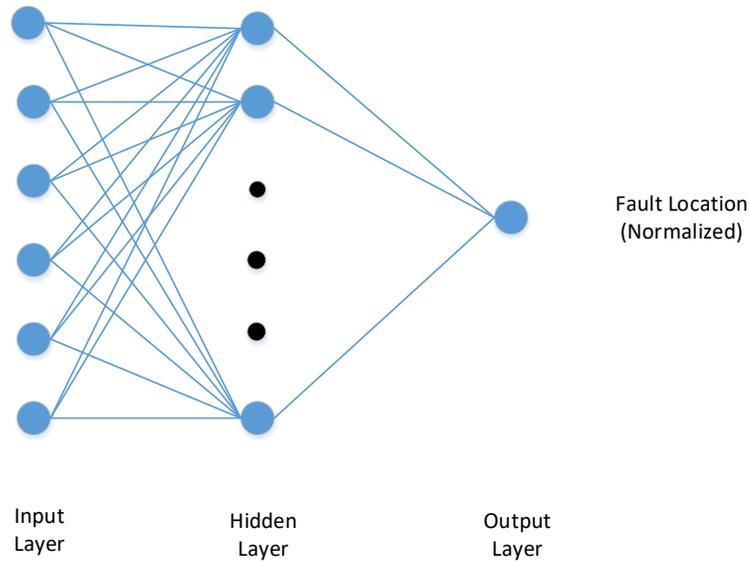
Figure 46 – Fault Location Artificial Neural Network Structure (Phase 2)

The development of the multiple ANNs used in this phase will follow a similar design approach as discussed in chapter 3. The input layer will only contain data entry points into the developed neural network. When a neural network is trained, assuming the MATLAB command "train" is used, the neural network will configure the number of inputs and outputs within each neural network structure. This configuration process is based on the input and target training data that has been presented to the network while performing the training process. This phase of the research will also be using the same measurement configurations as described in chapter 3.

Related to the multiple ANNs used within this phase of the research, each of the ANNs will be using the same number of network inputs that was presented to the network in chapter 3, but for clarity the number of inputs based on the collected measurement configuration being used to train the networks are presented in Table 15.

Table 15 - Number of ANN Inputs based on Type of Phasor Measurement (Phase 2)

Number of ANN Inputs per Type of Phasor Measurement	
Phasor Measurements with Orientation	Number of ANN Inputs
Current (I) @ Substation A	6
Voltage (V) @ Substation A	6
Voltage (V) and Current (I) @ Substation A	12
Voltage (V) and Current (I) @ Substations A and B	24

It is worth bringing up again that each current and voltage measurement used as the input training data contains a magnitude value in per unit and a phase angle value measured in degrees for each phase of the transmission line. Since the transmission lines are designed with three phases this results in six data points for each type of data measurement collected at each substation. Knowing that the overall single transmission line model used within phase 2 of this research was not changed from phase 1, it would be expected that the maximum and minimum voltage and current phasor values would not change. The maximum and minimum values for the full set of data extracted from the models used in phase 2 are shown in Table 16.

Table 16 - Maximum and Minimum Values for Voltage and Current Phasors (Phase 2)

	Max Value (per unit)	Min Value (per unit)
Voltage Magnitude	1.0163	0.0428
Voltage Angle	179.99	-180
Current Magnitude	66.551	3.1921
Current Angle	180	-180

The hidden layers of the fault classification and fault location ANNs will both contain only one hidden layer of neurons. This single hidden layer of neurons will again test a range of neurons to determine the best ANN structure that can have the highest accuracy of predicting any fault identification that has taken place on the transmission system. The range of neurons that phase 2 will attempt to test will utilize 6 to 36 neurons. The different ANN structures will step through the range of neurons by multiples of 3 neurons until all hidden layer neuron structures have been trained and developed. Each neuron used in the hidden layer will incorporate the use of the hyperbolic tangent sigmoid transfer function.

Finally, the output layer will be configured when training the network from the training data. If the training data was presented to the network as is, then the neural network would contain five ANN outputs since the fault location and fault classification target training data has not been separated. Therefore, before any training can take place the training target data will need to be separated so that the fault classification and fault location ANNs use the correct target data. Fault classification will use the first four rows of data in the target training data. These four rows of data will contain a discrete value of either a 0 or 1 that describe the fault connection or no-connection status of that phase or ground connection. Equation 4.3 provides the target training data used for the fault classification ANNs.

$$Fault\ Classification\ Target\ Data = \begin{bmatrix} Phase\ A \\ Phase\ B \\ Phase\ C \\ Ground \end{bmatrix} \qquad (4.3)$$

As for the ANNs that will predict the fault location there will only be one ANN output. The fault location target data will be imbedded in the target training data in the fifth row. This fifth row will need to be the only row used when training the fault location ANNs. This fault location value will be an actual floating-point value of the fault location from the reference substation.

Before training the fault classification and fault location ANNs there needs to be some pre-processing of the input and target training data. This was the same pre-processing steps that took place in phase 1. The input training magnitude values need to be normalized to the maximum magnitude value recorded in the training data set. Phase angle values were normalized to the maximum (positive) phase angle recorded in the training data. The input training data is common for both fault classification ANNs and fault location ANNs. Therefore, the same normalization will take place for both types of ANNs. As for the target training data, the fault classification target data will only contain discrete values of either a 0 or 1. There will be no normalization with these values. This is not the case for the fault location. The fault location was normalized to the total length of the transmission line. In theory, the fault location can be no longer then the length of the transmission line.

The last parameter to set before training the ANN is to select the training function or learning rule that will be used to train the neural networks. For all phases of this research, it was elected to use the Levenberg-Marquardt training function algorithm. Training these ANNs can possibly become a time constraint concern. When ANNs begin to contain numerous input and output data points, adjusting these weights between epochs can take

a lot of time. This should be a factor of thought when considering designing a neural network.

4.4 Development of Testing Data for Single Transmission Line Using Multiple ANN Approach

There were two different sets of testing data that was developed to evaluate the ANNs predictability for fault identification. For each phase of this research, different testing data sets were developed since each phase had small additions or modifications to the models and ANN development process. The models that were used to generate the input and target training data for training the ANNs were the same models that were used to generate the input and target testing data. The target testing data is used to validate the results and to check the ANN predictability error.

While developing the testing data sets, it was decided that the parameters of the faulted scenarios should not be changed drastically all at once. It was the thought that if the testing data was changed drastically and the ANN predictions were poor then it would be hard to understand why and when the results began to diverge from the actual fault classification and location of the fault. Instead the testing data was developed by only changing one fault scenario parameter at a time. The first set of testing data used a MATLAB random generator function (rand) to select fault locations on the transmission line that were different then fault locations used in the training data. Since the step size was so small (0.1 km) when moving the fault down the transmission line, the fault

locations that were used for testing incorporated fault location values to four decimal points of accuracy. For the first set of testing data, the fault resistances were not changed from the training data. Recall, that this phase of the research used fault resistances within the fault resistance limits of 1 Ω to 50 Ω, while varying the steps between the limits at multiples of 10 Ω, 5 Ω, and 2.5 Ω. When only modifying the fault locations for the 10 Ω separation, the testing input and target data was obtained with 110 different fault locations. For the ANNs that used the 10 Ω separation, the same fault resistances were used with the first testing data set. But for 5 Ω and 2.5 Ω separation, fault resistance values of 5 Ω multiple separation was used for evaluation with ANNs trained with both 5 Ω and 2.5 Ω separation training data. For the 5 Ω and 2.5 Ω separation testing data sets, only 60 random fault locations were simulated. This data results in the entire testing data sets to contain 6,600 testing data points for data sets containing only modified fault locations.

The second set of training data added some more complexity to the first testing data set. Along with the fault locations being varied, the second data sets also varied the fault resistances using the random generator in MATLAB. The only stipulation for varying the fault resistances was the random values of the fault resistances had to be between the trained fault resistance limits of 1 Ω to 50 Ω. The fault resistance values were generated by the random generator function in MATLAB (rand). When randomizing both the fault locations and fault resistance values, the full set of testing input and target data contained 25 random fault locations and 26 random fault resistance values. The full testing data set results in a total of 6,500 testing data points.

4.5 Results for the Multiple ANN Approach using Single Transmission Line Model

This section will be presenting an overview on the performance of fault identification using the multiple ANN approach. This multiple ANN approach will be using one ANN to identify the type of the fault that has occurred on the transmission line. But the developed approach will be using that fault classification ANN to enable one of four different ANNs to identify the accurate location of the fault. The idea behind separating the fault classification and fault location tasks into different ANNs is hoping that the developed networks have a simpler and more direct problem to solve. This phase of the research only looked at single hidden layer networks in hopes that simpler networks could be used to identify fault identification using multiple ANNs. Recall from earlier sections that it was believed that the threshold value that was set forth in phase 1 for fault location of this research was set relatively high to provide high levels of confidence for information which would be provided to field personnel. But if the threshold was set lower and the number of instances of faulted test scenarios over the threshold value became larger than the results might be misleading to believe that the single ANN approach had poor fault identification predictions. In this phase of the research, since the networks are believed to be a simpler architecture design, then it is possible that the threshold value for absolute error could be lowered. It was decided that the threshold value in phases 2 and 3 of this research would be set to a value of 1 km. This threshold converts into 0.621371 miles, roughly over a half of a mile. Other than the absolute error threshold value, the same performance metrics used in phase 1 will be used in phase 2 as well.

These performance metrics for the fault classification ANN were evaluated using maximum absolute error, minimum absolute error, and the average absolute error. The results presented in this section will be looking at the maximum absolute error to determine how high the absolute error gets when trying to predict the phase and ground connections of the fault. The minimum absolute error for all trained ANNs for fault classification and fault location had maximum error values near zero error. The performance metrics for fault location evaluated the performance based on maximum absolute error, minimum absolute error, average absolute error, and the number of instances that provide absolute error over the threshold value of 1 km. The results provided in this section will concentrate on the maximum absolute error and the number of instances over the threshold value. It was believed that these two sets of metrics can provide sufficient evidence to describe the performance of the ANNs ability to predict fault identification. Fault classification absolute error is the difference between the actual discrete value of 0 or 1 for the fault connection or no-connection algorithm versus the ANN prediction output that corresponds to that connection. Using absolute error for fault classification would be unitless. But when absolute error is used for fault location, which is comparing the actual fault location from a reference bus versus the ANNs fault location prediction, both parameters of this comparison will be in units of km to make the absolute error calculation be in units of km.

Each ANN trained in this phase of the research was trained with the four measurement configurations that have been discussed throughout this dissertation. Each one of these measurement configurations will describe the results within its own a sub-section. The training data used for training the ANNs used 0.1 km fault steps for the fault data, but the

fault resistance step sizes between the low and high fault resistance limits were varied by

10 Ω, 5 Ω, and 2.5 Ω. Testing data was presented in two sets of data which represent data

that only modified the fault locations to containing random fault locations that were not

in the original training data. The second testing data set was developed by modifying

both the fault locations and fault resistances with the values selected at random between

the original fault location and fault resistance limits. Each ANN that was tested during

this phase of the research concluded that when both, the fault location and fault

resistance, were modified then the absolute error contained the worst-case values.

Therefore, the data presented in the following sections will present data with random

fault locations and random fault resistances.

4.5.1 Fault Identification Results using Current from Substation A (Phase 2)

The first set of results that will be presented will represent the measurement configuration

of current phasors being available from one substation. The first ANNs that were tested

were trained with fault resistances that used steps of 10 ohms between the low and high

fault resistance limits. The maximum error produced on each phase by this set of ANNs

for fault classification for most of the hidden layer neuron structures produced low errors.

Figure 47 provides the performance trend of maximum absolute error for the fault

classification ANN with the ANN structures being trained with fault resistances of 10 Ω

multiples.
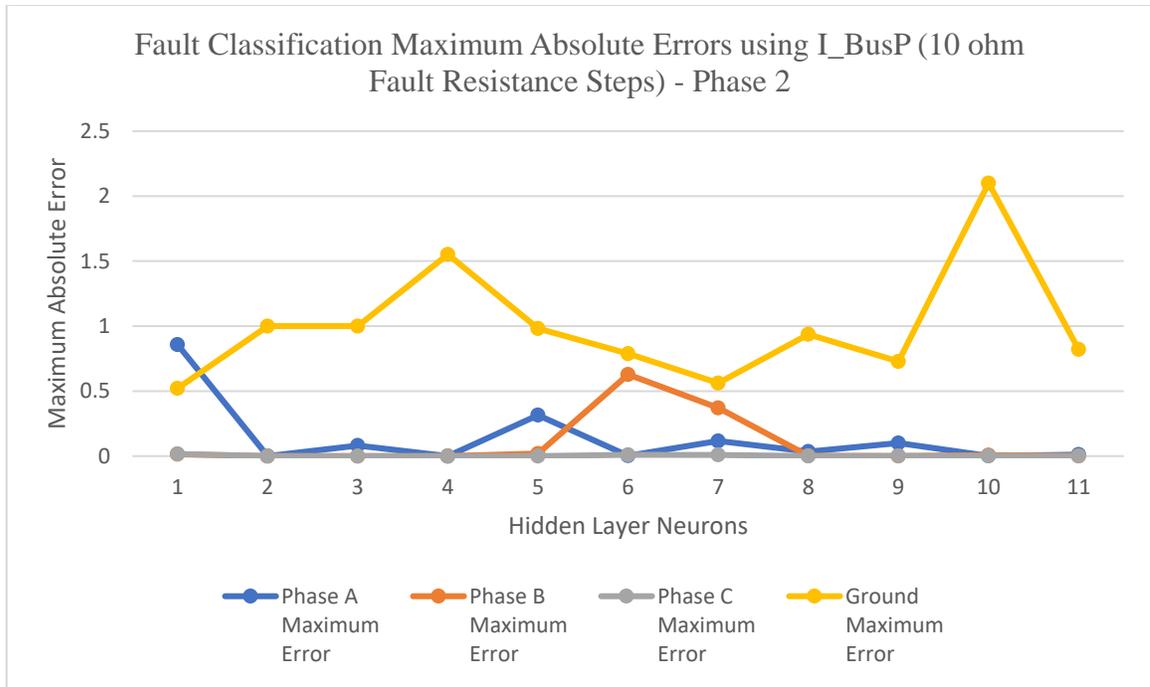
Figure 47 - Fault Classification Maximum Absolute Error using I_BusP (10 Ω Fault Resistance Steps) - Phase 2

But the ground connection has extremely high errors that creep above 2. Since this is a connection or no connection status the values should be between 0 or 1. Therefore, the ground connection of the classification ANN is predicting this fault value outside of the 0 or 1 range for most ANN structures. By changing to ANNs that were trained with fault data containing fault resistance of 5 Ω and 2.5 Ω, the ground connection errors became less than 100 percent maximum error which would conclude that the ANN is predicting an output value in the tolerable range. By using training data containing 2.5 Ω fault resistance steps, the data shows that the best performing ANN structure predicting maximum error with 18 neurons in the hidden layer produce ground connection errors near 1 percent.
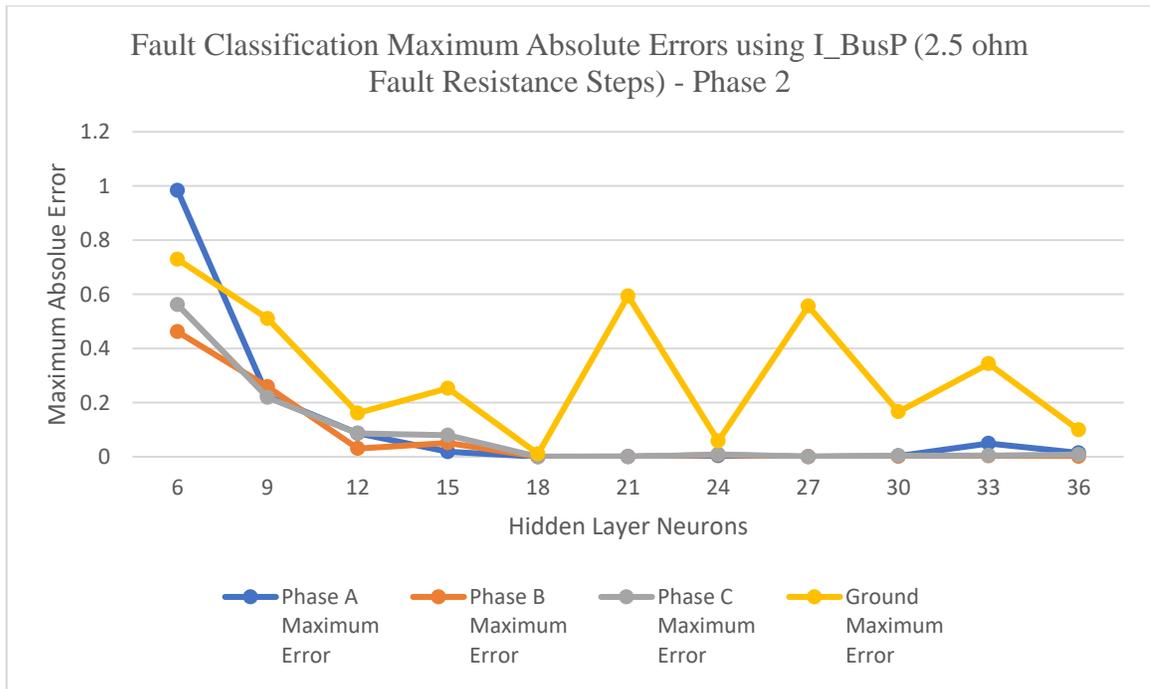
Figure 48 - Fault Classification Maximum Absolute Error using I_BusP (2.5 **Ω** Fault
Resistance Steps) - Phase 2

The figure still shows that the ground connection is still a little unstable during adjacent

neuron structures. This could mean that training the ANN structures again with the

random training values being selected could produce a different outcome.

Table 17 provides a fault classification error comparison that defines the number of

faulted scenarios that generated errors greater than 10 percent for ANNs trained with 2.5

Ω separation training data.

Table 17 - Fault Classification Error Comparison using I_BusP for 2.5 Ω Training Data

Fault Classification; I_BusP Data; Fault Distance = 0.1 km; Fault Resistance 1 - 50 ohm, Training Data Separated by 2.5 ohm								
	Phase A		Phase B		Phase C		Ground	
Hidden Neuron	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%
15	1.86E-02	0	5.10E-02	0	7.96E-02	0	0.2529	42
18	1.90E-04	0	1.00E-05	0	1.00E-05	0	1.08E-02	0
21	9.00E-04	0	1.07E-04	0	7.44E-05	0	0.5935	2
24	3.77E-03	0	6.84E-03	0	8.79E-03	0	5.81E-02	0
27	1.67E-04	0	1.36E-04	0	9.20E-04	0	5.56E-01	18
30	2.84E-03	0	1.22E-03	0	3.94E-03	0	0.16639	4
33	4.93E-02	0	3.41E-03	0	4.73E-03	0	0.3436	14
36	1.44E-02	0	1.59E-03	0	7.86E-03	0	9.93E-02	0

Table 17 shows that all phase conductors have less than 10 percent error for all neuron structures from 15 to 36 neurons. But this is not the case for all the ground connection.

It was observed that very similar behavior occurred when using the multiple ANNs to predict fault location based on their projected fault type. These results for fault location assume that the fault classification predictions are 100 percent correct and at the correct fault location ANN will be enabled. This allows the results to be studied independently. When ANNs were used to predict fault location using 10 Ω fault resistance data, then nearly every fault type except for line to line faults had maximum errors close to 40 km. Using 33 neurons in the hidden layer produced the lowest maximum error or around 20 km for all fault types, which also results in LG faults have 73 instances out of 1950 tested scenarios and LLL having 123 instances out of 650 tested scenarios over the 1 km threshold. These results need to be lowered before this type of approach can be used for

fault location. Figure 49 provides the fault location trend with ANNs trained with 10 Ω

fault resistance steps.
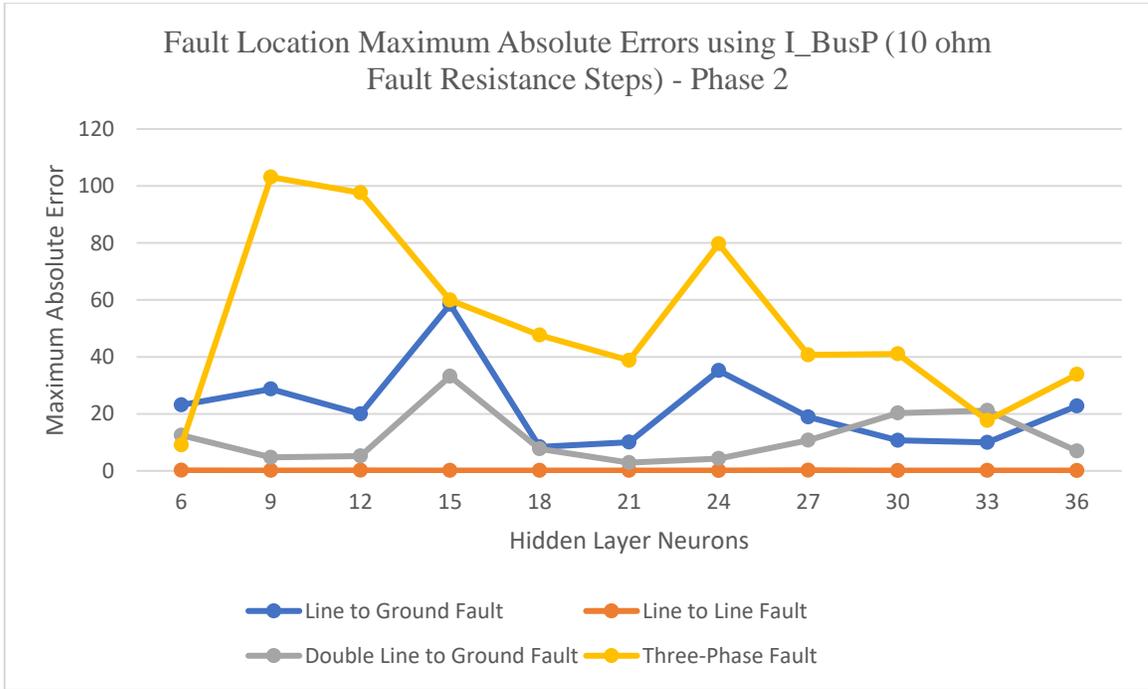
Figure 49 - Fault Location Maximum Absolute Error using I_BusP (10 **Ω** Fault Resistance Steps) - Phase 2

When the ANNs that have been trained with data containing 2.5 Ω fault resistance steps,

it was identified that having 15 to 36 neurons in the hidden layer produce maximum

absolute errors less than 0.5 km (0.3107 miles). The best performing ANN structure for

fault location maps to 24 neurons in the hidden layer. This produces maximum error of

0.3 km of error for LLL faults. Figure 50 provides the fault location trend with ANNs

trained with 2.5 Ω fault resistance steps.
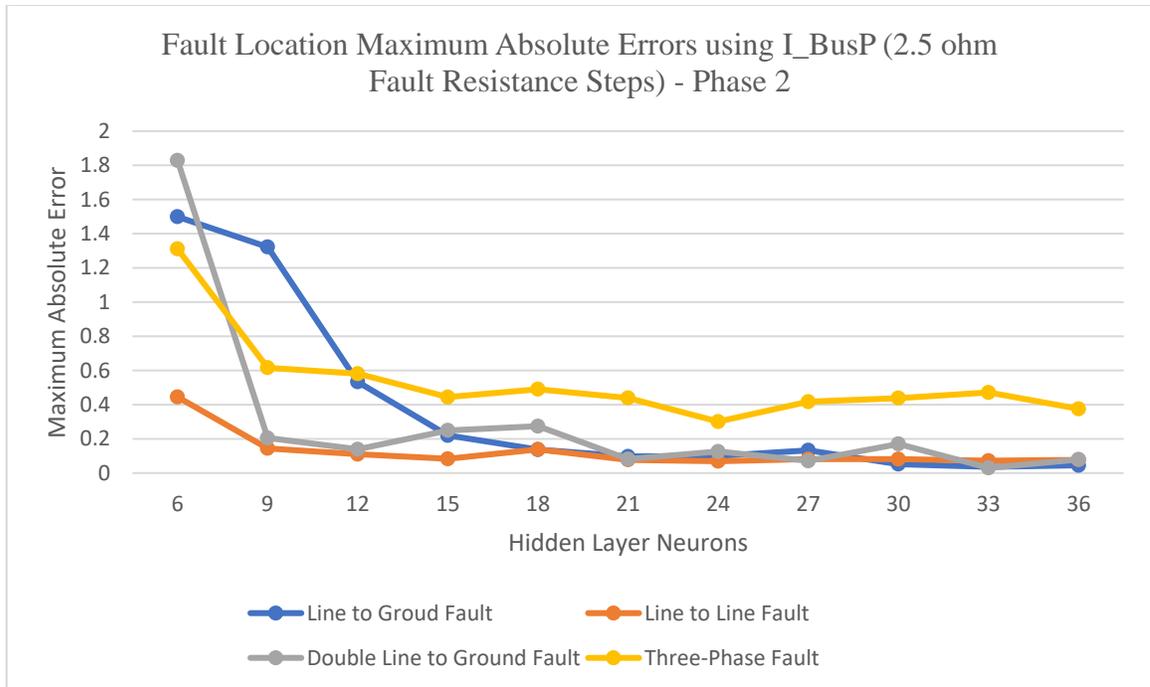
Figure 50 - Fault Location Maximum Absolute Error using I_BusP (2.5 ohm Fault
Resistance Steps) - Phase 2

4.5.2 Fault Identification Results using Voltage from Substation A (Phase 2)

Evaluation of fault identification continues with using voltage from one substation that is

connected to the transmission line. As results are continued to be presented for phase 2, a

trend in the results becomes evident. ANNs that were trained with 10 Ω fault resistance

steps produce results containing higher values of errors then when using 5 Ω or 2.5 Ω

fault resistance steps. When using voltage from one bus, the maximum errors stay within

the acceptable ranges, but ground connection has nearly 100% maximum error for some

tested scenarios using ANN structures that contain less than 18 neurons in the hidden

layer. The maximum error for the ground connection then begins to decrease to the

lowest ANN structure of 33 neurons in the hidden layer at nearly 20 percent error. Other

than the ground connection having high errors the phases for nearly half of the structures

seem to predict the fault classification with very low error rates.
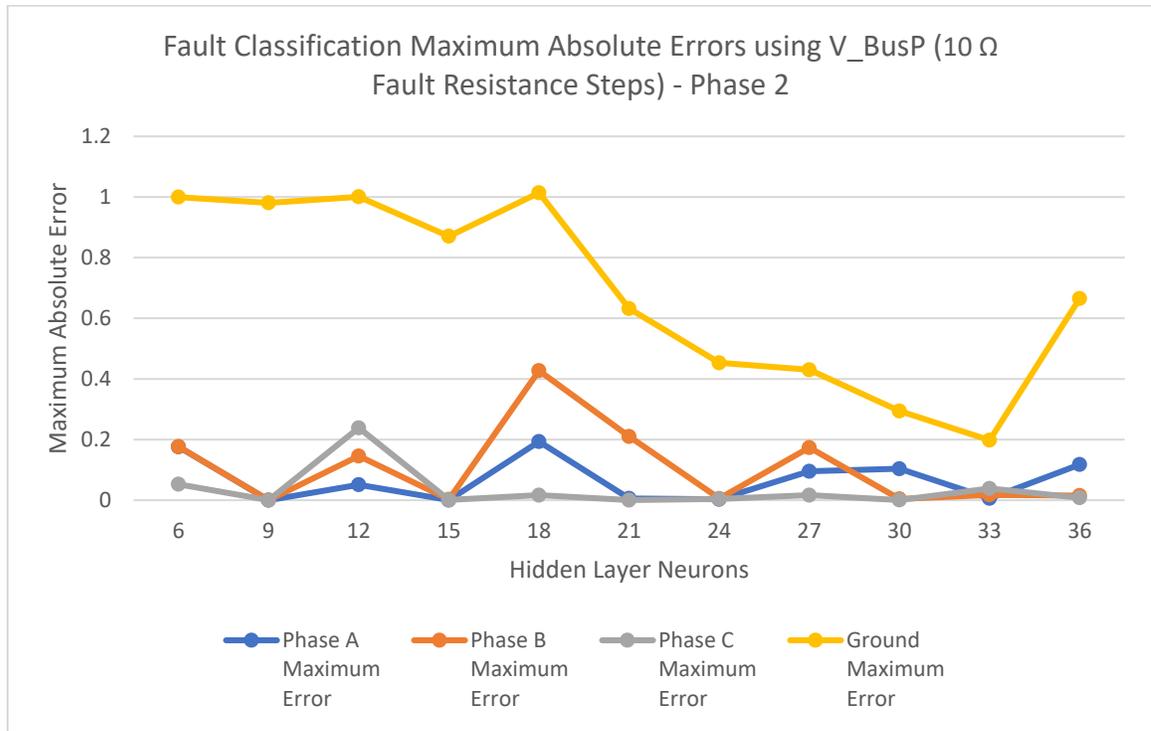
Figure 51 - Fault Classification Maximum Absolute Error using V_BusP (10 Ω Fault
Resistance Steps) - Phase 2

Once ANNs that have been trained with 5 Ω or 2.5 Ω fault resistance steps, the fault

classification maximum absolute error for the ground connection begins to improve

drastically between 15 to 36 neurons. Between this range of 15 to 36 neurons the

maximum error peaks around 10 percent to well below 1 percent. For fault classification

the best performing neuron structure would be 24 neurons in the hidden layer to produce

minimum error for all phase and ground connections. Table 18 provides the maximum

number of fault classification instances that exceeded 10 percent for each neuron

structure between 15 to 36 neurons.

Table 18 - Fault Classification Error Comparison using V_BusP for 2.5 Ω Training Data

Fault Classification; V_BusP Data; Fault Distance = 0.1 km; Fault Resistance 1 - 50 ohm, separated by 2.5 ohm							
	Phase A		Phase B		Phase C		Ground	
Hidden Neuron	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%
15	7.98E-02	0	1.43E-04	0	3.71E-02	0	0.15953	3
18	1.38E-05	0	1.46E-05	0	4.00E-05	0	6.44E-01	6
21	1.62E-01	2	4.00E-04	0	2.07E-02	0	0.28015	5
24	3.81E-04	0	3.90E-04	0	5.12E-04	0	2.73E-03	0
27	6.89E-03	0	1.17E-03	0	3.19E-03	0	4.41E-02	0
30	1.24E-03	0	2.82E-03	0	3.47E-03	0	0.10126	1
33	3.71E-03	0	9.20E-04	0	2.95E-03	0	2.14E-02	0
36	3.53E-04	0	3.95E-04	0	1.24E-03	0	2.32E-03	0

Fault location predictions are very poor using ANNs trained with fault resistance steps of

10 Ω. It was observed that the three-phase fault and the line to ground faults had high

prediction errors. These errors are peaking over 80 to 100 km for three-phase faults and

nearly 60 km for line to ground faults. This was consistent with the behavior that was

seen when using current from one bus. Line to line faults had near perfect fault location

predictions. Results continue to drastically improve if the test scenarios are placed in an

ANN that has been trained with 5 Ω or 2.5 Ω fault resistance steps. If the ANNs used for

fault location predictions are trained with 2.5 Ω fault resistance steps, then the results

begin to average around 0.5 km for each fault type after 18 to 36 neurons in the hidden

layer. This is deemed a very acceptable error when attempting to locate fault location.

The best performing ANN structure to locate fault location was 36 neurons in the hidden

layer. All fault location errors have no instances over the threshold value that was set for

this phase. Figure 52 provide the results for fault location maximum absolute error using

voltage from one substation using ANNs trained with 2.5 Ω fault resistance step data.
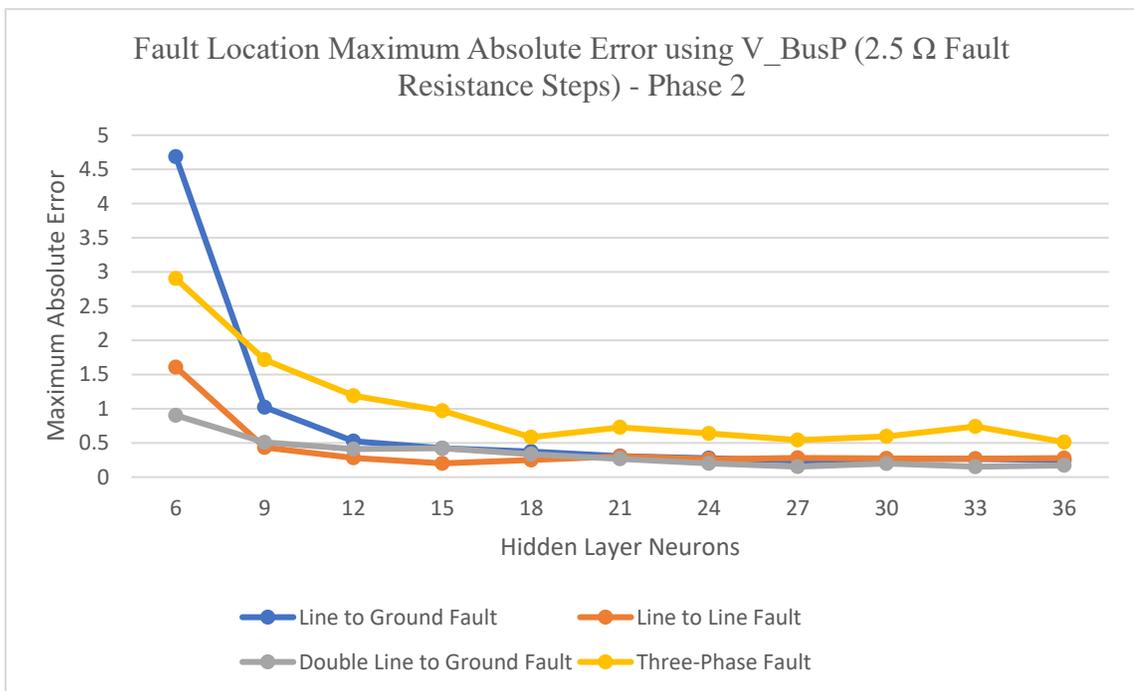
Figure 52 - Fault Location Maximum Absolute Error using V_BusP (2.5 Ω Fault
Resistance Steps) - Phase 2

4.5.3 Fault Identification Results using Voltage and Current from Substation A (Phase 2)

When voltage and current are combined for identifying fault identification, it would be expected that the results would improve since there is more data available. This was not the case. When the ANNs that are used to detect the fault classification that are trained with 10 Ω fault resistance steps the errors are just as high as using on voltage or current phasors from one substation. The results are again showing that the phase A and ground connections are having trouble detecting fault classification with errors reaching between 50 to over 100 percent. Phase C and phase B connections are having maximum errors near 10 percent.

If the ANNs are trained with 2.5 Ω fault resistance steps, the ANN projections show major improvements for fault classification. With 21 to 36 neurons being in the hidden layer maximum errors decrease just over 10 percent. The best ANN structure that was used to predict fault classification with the smallest maximum errors was 30 neurons in the hidden layer. Figure 53 provides the fault classification results for maximum absolute error using a 2.5 Ω fault resistance step trained ANN with voltage and current measurement configuration from one substation.
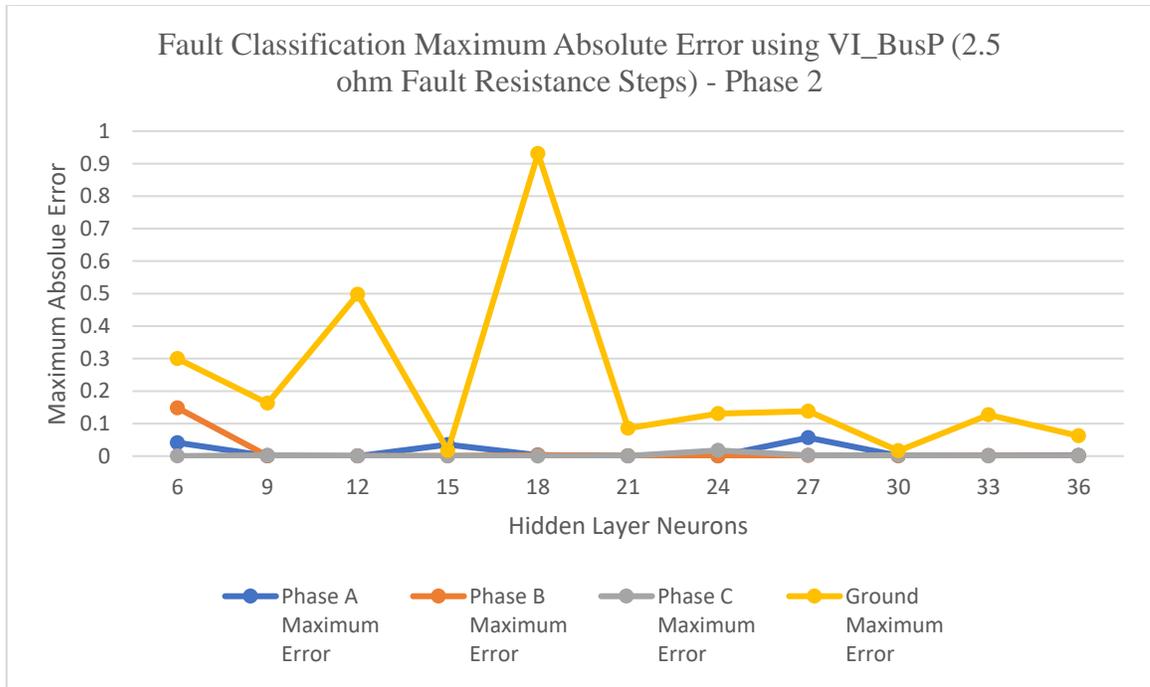
Figure 53 - Fault Classification Maximum Absolute Error using VI_BusP (2.5 Ω Fault Resistance Steps) - Phase 2

Table 19 provides the maximum number of fault classification instances that exceeded 10 percent for each neuron structure between 15 to 36 neurons. The ground connection still has the worst performance of all connection points, but the number of test points exceeding 10 percent is still positive.

Table 19 - Fault Classification Error Comparison using VI_BusP for 2.5 Ω Training Data

Fault Classification; VI_BusP Data; Fault Distance = 0.1 km; Fault Resistance 1 - 50 ohm, separated by 2.5 ohm								
	Phase A		Phase B		Phase C		Ground	
Hidden Neuron	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%	Max Error	Error > 10%
15	3.51E-02	0	4.70E-04	0	4.67E-04	0	0.01714	0
18	3.13E-03	0	3.03E-03	0	6.00E-05	0	9.31E-01	9
21	5.20E-04	0	6.10E-04	0	5.10E-04	0	0.0857	0
24	2.40E-04	0	1.45E-04	0	1.74E-02	0	1.31E-01	8
27	5.66E-02	0	1.70E-03	0	2.29E-03	0	1.38E-01	2
30	9.00E-04	0	9.38E-04	0	2.73E-03	0	0.0163	0
33	1.19E-03	0	8.40E-04	0	5.54E-03	0	1.27E-01	2
36	1.95E-03	0	7.40E-04	0	4.83E-04	0	6.23E-02	0

The ANNs predicting fault location does show some improvements while using voltage and current from one substation. Using ANNs that are trained with 10 Ω fault resistance steps causes the line to ground fault errors to decrease within the range of 10 to 20 km maximum error. But the line to ground fault still experiences about 70 to 80 instances out of 1,950 tested scenarios that exceed the 1 km threshold. Therefore, there is still lots of room for improvement. This improvement exists when more data is used to train the ANNs. When ANNs trained with 2.5 Ω fault resistance steps, the results show that the maximum error for unsymmetrical faults are less than 0.1 km. But the three-phase fault is not improving as much but can produce maximum errors as low as 0.3 km. This is still a very promising result. Figure 54 provides the fault classification results for maximum absolute error using a 2.5 Ω fault resistance step trained ANN with the voltage and current measurement configuration.
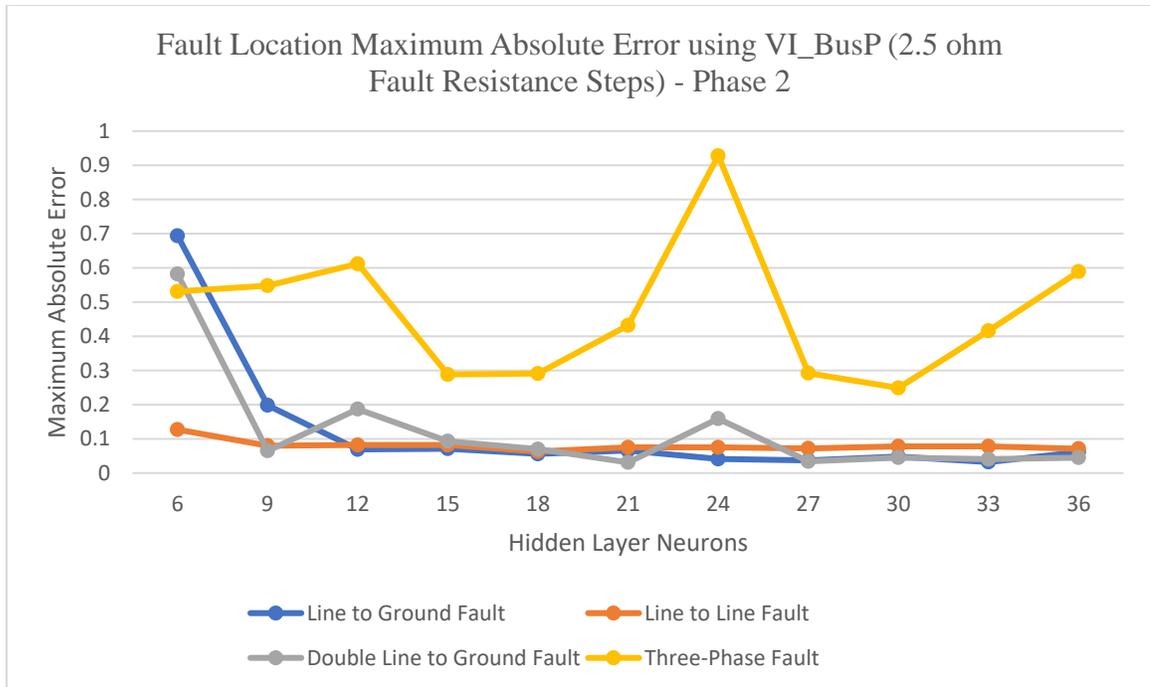
Figure 54 - Fault Location Maximum Absolute Error using VI_BusP (2.5 Ω Fault Resistance Steps) - Phase 2

4.5.4 Fault Identification Results using Voltage and Current from Substation A and B (Phase 2)

The last evaluations for phase 2 was reviewing the fault identification performance using voltage and current phasors from two substations that are connected to each other via the transmission line. Using this measurement configuration helped the ground connection maximum error performance by decreasing near an average of 50 to 60 percent with ANNs trained with 10 Ω fault resistance steps. The other phases for all evaluated ANN structures had errors near 20 percent. For the ground connection, the results need to

improve drastically since data containing 50 percent error could not be used in the algorithm proposed in this phase.

When the ANNs are modified to include 2.5 Ω fault resistance step training data, which results into more unique training data, the ANN output predictions results are well improved between 15 to 36 neurons in the hidden layer. All phases and ground connections are at or below 10 percent maximum error. The best performing structure for fault classification would be 27 neurons in the hidden layer. Figure 55 provides the fault classification maximum errors using ANNs with 2.5 Ω fault resistance step training data.
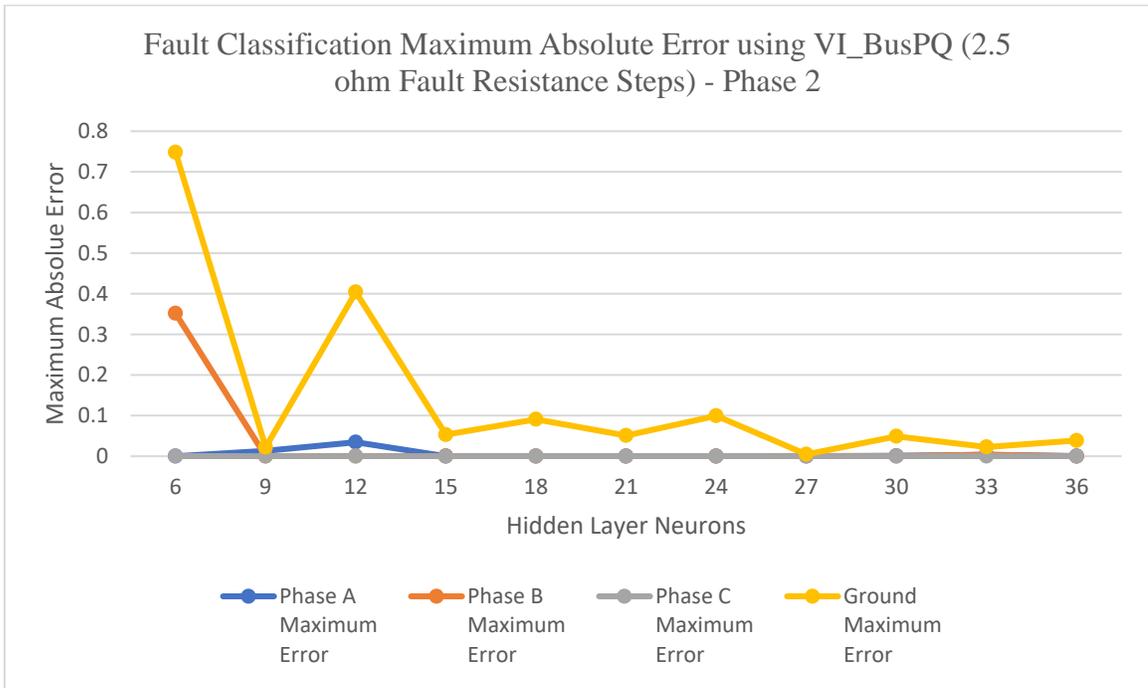
Figure 55 - Fault Classification Maximum Absolute Error using VI_BusPQ (2.5 Ω Fault Resistance Steps) - Phase 2

As for fault location when the testing data is supplied to the ANNs trained with 10 Ω fault resistance step data, the ANN predictions contain less than 10 km errors for all tested neurons in the hidden layer except for 36 neurons. These results can still be improved by supplying the testing data to an ANN with more trained data as shown in this dissertation. When the testing data is supplied to ANNs with 2.5 Ω fault resistance step training data, its observed that the data presents maximum fault location error less than 0.1 km for all tested neurons except for 33 neurons. The best performing fault location ANN structure would be 27 neurons in the hidden layer. Figure 56 provides the fault location maximum errors using ANNs with 2.5 Ω fault resistance step training data.
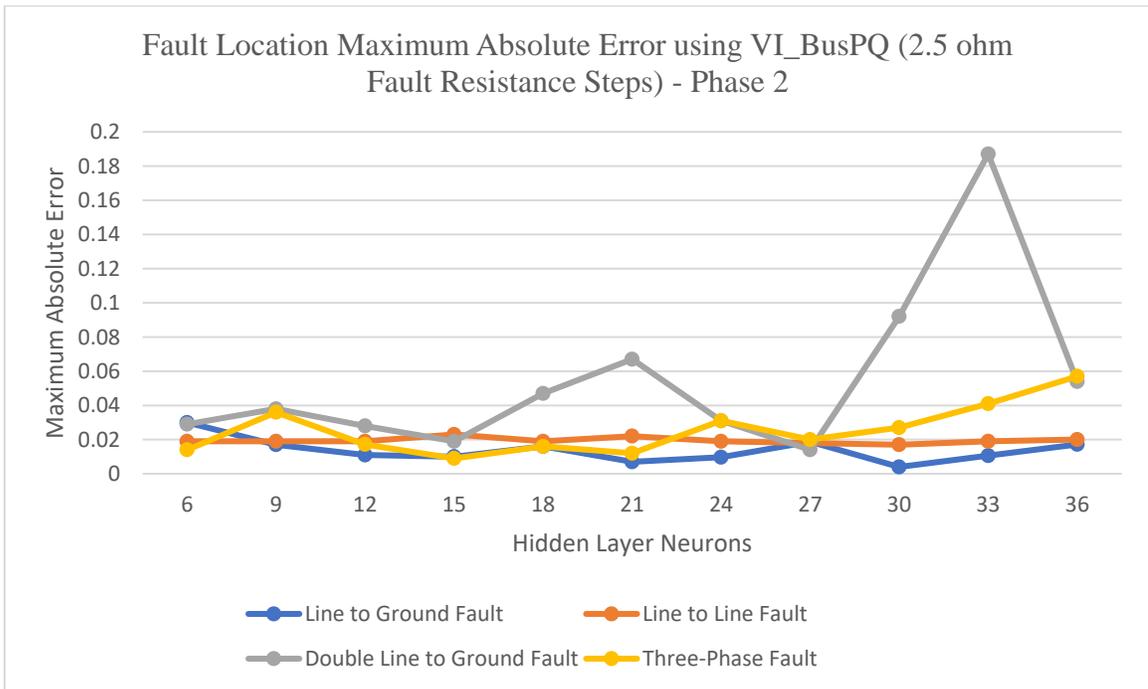
Figure 56 - Fault Location Maximum Absolute Error using VI_BusPQ (2.5 Ω Fault Resistance Steps) - Phase 2

Chapter 5 – Fault Identification with Parallel Transmission Lines using Multiple ANN Approach

This chapter will discuss the last phase of this research. During this final phase, the transmission line model topology was changed to analyze transmission fault identification using the parallel transmission line configuration. A parallel transmission line configuration consists of two or more similar or distinct transmission lines that are located within close proximity to each other. These multiple transmission lines may operate at different voltage levels, and power may flow the same or opposite directions depending on system conditions. For transmission lines to be considered parallel transmission lines they must be either located on the same transmission towers or on separate transmission towers within the same right of way easement. There are many exceptions to the definition of parallel transmission lines. Most of these parallel transmission lines will only follow the same path of the other transmission line for only a portion of the distance from one substation before diverting into a different direction to connect to another substation.

The parallel transmission line model is going to require a few set up changes to the Simulink models and development of the input and target training data algorithms along with the amount of ANNs that have been trained related to phase 2. The following sections of this dissertation will point out all the modifications that were made in order to analyze fault identification on parallel transmission lines. This portion of the research attempted to follow a same ANN development approach that was discussed in chapter 4.

5.1 Two-Terminal Parallel Transmission Line Model

The last transmission configuration that will be studied within this research is the parallel transmission line. This configuration will contain two or more transmission circuits located within close proximity to each other or share the same right of way easement. The most interesting change that is introduced with the parallel transmission line topology is the effect of mutual coupling. NERC defines mutual coupling as the electromagnetic interaction between two or more transmission lines that are routed in parallel for a substantial distance [37]. As shown in chapter 1, mutual coupling can play a role with parallel transmission line topologies when a fault occurs on one of the parallel circuits and can cause an increase of current in the other non-fault circuit. This effect can be greater in faults that encompass a faulted connection with ground which is related to the zero-sequence current that is induced in the healthy, non-faulted, circuit.

This phase of the research will be evaluating a parallel transmission line model that contains two parallel transmission line that follow the same entire path from substation A to substation B. This entails that there will not be any break in the mutual couple modeling within the Simulink model. The development of the parallel transmission line model began using the single transmission line model used in phase 2. The only two sections of the single transmission line model that was not modified in developing the parallel transmission line model was the generation and three-phase mutual impedance blocks. These two blocks are still serving the same purpose as they did with the single transmission line model in phase 1 and 2. To begin creating the parallel transmission line model, the changes began with adding the second transmission line using the distributed

parameter line blocks. Recall from earlier in this dissertation, chapter 3, that within the

distributed parameter line blocks that adding more phases to the modeling block will in

add another transmission line in parallel to the original single transmission line. To model

both transmission lines, the distributed parameter line blocks had the number of phases

modified from three to six. Since each transmission line contains three phases and

changing the number of phases to six, this places two transmission lines in parallel. If

more transmission lines were modeled in parallel the number of phases would change by

multiples of three. Now that there are two transmission lines in the model and they are

considered close to one another, there needs to be some type of mutual coupling

impedance added to the distributed parameter blocks. Table 20 provides the impedance

values that were used in the parallel transmission line model.

Table 20 - Distributed Parameter Line Model Block Impedance Details (Phase 3)

Distributed Parameter Line Model Block – Parallel Transmission Line Model			
		Distributed Parameter Line (Bus P to Fault)	Distributed Parameter Line (Bus Q to Fault)
Phases		6	6
Positive Sequence	R1	0.15476 Ω/km	0.15476 Ω/km
	L1	9.707e-4 H/km	9.707e-4 H/km
	C1	1.209e-8 F/km	1.209e-8 F/km
Zero Sequence	R0	0.37417 Ω/km	0.37417 Ω/km
	L0	3.0e-3 H/km	3.0e-3 H/km
	C0	7.495e-8 F/km	7.495e-8 F/km
Zero Sequence Mutual	R0m	0.36287 Ω/km	0.36287 Ω/km
	L0m	1.89e-3 H/km	1.89e-3 H/km
	C0m	4.505e-9 F/km	4.505e-9 F/km

There will be two developed versions of the parallel transmission line model. One of these models will place the three-phase fault block on the first transmission line and the other model will place the three-phase fault block on the second transmission line.

Since the model now contains two transmission lines, both lines at both end points of the line need to monitor bus voltage and line currents. Theoretically, there could be only one bus voltage measurement for both lines since both lines terminate at the same substation bus. But for simplicity the transmission line model will capture bus voltage and line current at both ends on the transmission lines. This capture of bus voltage and line current was achieved by using a second three-phase V-I measurement block at each end of the added transmission line. Once all the V-I measurement blocks have been added, the model should contain four V-I measurement blocks used within the model. The two new V-I measurement blocks should be replicas of the original two V-I measurement blocks with different signal labels to differentiate between the different V-I measurement blocks at either substation or on either transmission line.

This will complete all the transmission line topology model changes. Figure 57 provides the complete high-level transmission line modeled topology within Simulink.
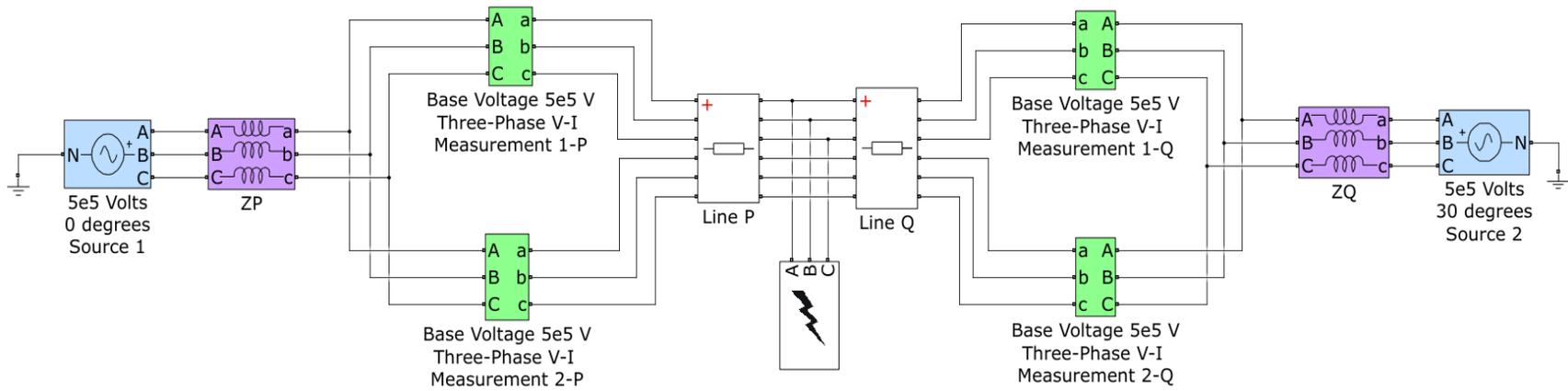
Figure 57 - Parallel Transmission Line Simulink Model (Phase 3)

Since two V-I measurement blocks were added to the Simulink model, there needs to be two conversion blocks added to the Simulink model to convert the instantaneous voltage and current measurements that are outputs from the new V-I measurement blocks into phasor values. Again, these blocks are exact replicas from the masks used in phase 2 of this research. This completes all modifications to the Simulink parallel transmission model so that the model can now go through faulted condition simulations to obtain input and target training data.

5.2 Development of Input and Target Training Data for Parallel Transmission Line using Multiple ANN Approach

The approach that acquires the input and target training data from the parallel transmission model is discussed in this section. This approach will closely follow the approach used in phase 2. This phase of the research will also be using the multiple ANN structure approach to solve both the classification of the fault and the fault location. By introducing the parallel line configuration, the fault classification prediction become a little more useful then describing just the type of fault. The fault classification ANN will provide a prediction for not only the type of fault that has occurred but also providing which transmission line contains the faulted condition. This information alone can be very useful for individuals in the field trying to isolate the fault or identify the cause of the fault.

This process begins by setting up the simulation automation for the Simulink model that was described in section 5.1. As stated in section 5.1, there will be two similar parallel transmission line models used to perform the faulted simulations. The difference in the two models is the three-phase fault block will be placed on one of the transmission lines in the first model and then the second model places the fault on the second transmission line. Each transmission line model will again be simulated for a total of eight cycles with the faults being applied to the transmission line models at the second cycle. This places the fault on the line for a total of 6 cycles and allows the user to see the effects that the fault has on the line if needed. Before any recorded faulted simulations were performed, the faulted voltage and current waveforms were reviewed to see how the relative tolerance within each simulation needed to be adjusted. This analysis to determine how to set the relative tolerances was performed in the same manner as performed within phase 2. This relative tolerance analysis was completed by first applying faults to the transmission line that encompassed the ten different fault types. Within these different fault types, different fault resistance values between 1 Ω to 50 Ω by taking steps of 10 Ω within that specified range were applied to each location. Faults that were applied to the transmission line were placed at different locations on the transmission line in increments of 5 km while applying different relative tolerances to the model. This process was repeated until the voltage and current waveforms became relatively smooth (without any imbedded harmonics). This analysis ranged the relative tolerance at the default setting of 1e-4 to 1e-7. This process was a trial and error approach (manual process) which was continued until a recommendation combination of fault location to relative tolerance could be meet. Table 21 provides a match list for the range in fault location from

substation A and the relative tolerance setting in Simulink. The reason for setting the

small relative tolerance is for the simulation to converge especially for close in fault

scenarios. These relative tolerance settings were determined to be same values as used

within phase 2 that used the single transmission line topology.

Table 21 - Relative Tolerance Settings for the Parallel Transmission Line Simulations
(Phase 3)

Relative Tolerance Settings for Parallel Transmission Line Model Simulations	
Fault Location Distance	Relative Tolerance Setting
3 km to 5 km	1e-7
5 km to 10 km	1e-6
10 km to 18 km	1e-5
18 km to 82 km	1e-4

This phase of the research will be collecting bus voltage and/or line current measurement

data during faulted transmission line conditions at both substations on both lines for the

initial sample of the fifth cycle of simulation time. There has been discussion within this

dissertation that describes all the possible measurement configurations that are possible

with voltage and/or current measurements that are contained within a two-bus

transmission line topology. As with phases 1 and 2, the following are the measurement

configurations that were used to collect data to analyze the predictability of fault

identification with parallel lines.

- Faulted line current measurements from substation A

- Faulted bus voltage measurements from substation A

- Faulted bus voltage and line current measurements from substation A

- Faulted bus voltage and line current measurements from both substation A and B

It has been implied that when the automation simulated the faulted conditions on the transmission line models all voltage phasors, current phasors, and target data is contained with their own data set. When each of the neural networks need to be trained on a specific measurement configuration the full set of training data will be parsed to obtain only the corresponding data as needed.

Faults were applied down the transmission line in steps of 0.1 km. With the initial and final fault locations being set at 3 km and 97 km (3 km from each substation) from substation A. To determine the number of faulted conditions that will be applied to each faulted scenario, the result is the same as in phase 2 which is 941 faulted conditions. But since there are now two transmission lines in the model, this number of fault conditions are doubled to 1,882 total faulted conditions per fault scenario. As stated, many times in this dissertation, the goal of this research was to have the ability to provide more training data to the neural networks and determine if more data provides better ANN predictability results. Since the step size of moving the fault down the transmission line was kept constant the only way to have variable fault data sets was to modify the fault resistances values. This was the approach also used in phase 2 and will be the approach used within this phase. To have a good comparison of results between the single transmission line versus the parallel transmission line, the same data sets of fault resistance values were used as well. To recap, there are three different sets of fault resistance values that were used. These three sets of fault resistances stepped between the limits of the fault resistance range (1 Ω to 50 Ω) at multiple of 10 Ω, 5 Ω, and 2.5 Ω.

Since the step size of moving the fault down the transmission line is set at 0.1 km, the number of different fault types, and the different sets of fault resistances are known, the total sets of faulted measurement contained in the training data sets can be calculated. Evaluating the parameters of the data used in this phase and phase 2 its identified that the only difference is there is a second transmission line that all these faulted conditions should be applied. Therefore, all the training inputs for the parallel transmission line model should be multiplied by the number of parallel lines being studied. The number of training data sets within the input and target training data, representing all fault types, is presented in Table 22.

Table 22 - Number of Training Data Sets for ANN Training (Phase 3 – Parallel Transmission Line Model)

Number of Training Data Sets Used for ANN Training (Phase 3 – Parallel Transmission Line Model)		
Simulated Fault Resistances	Fault Location Step Size	Number of Training Data Sets
1Ω, 10Ω, 20Ω, 30Ω, 40Ω, 50Ω	0.1 km	112,920
1Ω, 5Ω, 10Ω, 15Ω, 20Ω, 25Ω, 30Ω, 35Ω, 40Ω, 45Ω, 50Ω	0.1 km	207,020
1Ω, 2.5Ω, 5Ω, 7.5Ω, 10Ω, 12.5Ω, 15Ω, 17.5Ω, 20Ω, 22.5Ω, 25Ω, 27.5Ω, 30Ω, 32.5Ω, 35Ω, 37.5Ω, 40Ω, 42.5Ω, 45Ω,47.5Ω, 50Ω	0.1 km	395,220

It is important to consider the time commitment it took to obtain the training data for the modeled parallel transmission topology. Table 23 provides the recorded time commitment it took to obtain the training data sets. These simulations where performed on an Intel Core I5 – 8400 CPU. The data set containing 395,220 data sets of faulted

152

conditions were simulated in separate simulations separated by the different line models. Then the data was concatenated after the data was collected for the two models. This was an attempt to speed up the data collection process. All other data collection tasks were performed as one simulation run.

Table 23 - Time Elapsed to Collect Parallel Transmission Topology Training Data (Phase 3)

Time Elapsed to Collect Training Data Sets (Phase 3 – Parallel Transmission Topology)	
Number of Training Data Sets	Time Elapsed to Collect Data (Hours)
112,920	125.8298
207,020	233.0367
197,400	567.164

It was obvious from phase 2 that the time it was going to take to gather input and target data was going to be time intensive. From the data shown in Table 23, the data collection took just over 5 days up to just over 23 days to complete the data collection if the simulations were performed serially as presented. For the parallel configuration, the time constraint becomes proportional to the number of faulted scenarios that the user wants to collect data from and the number of transmission lines that are in parallel.

The process of collecting the training input and target data will follow the same approach as discussed within chapters 3 and 4. But there will be some differences in terms of the outputs of the input and target data collection automation. Faults were applied to the transmission model for the different combinations of fault types and fault resistances at determined fault locations on the transmission line model. These fault locations will

begin at an initial location on the transmission line and moved down the transmission line until the final fault location was reached. These faulted simulations were then performed on the second parallel transmission line. All the fault data was then collected into five different data matrices that correspond to bus voltage magnitudes, bus voltage phase angles, line current magnitudes, line current phase angles, and the corresponding target (actual) fault data. As for the bus voltage measurements (magnitudes and phase angles) the measurements are taken at the end points of each parallel line at each substation. Since electrically the transmission lines meet at the same substation bus, the bus voltage measurements for both lines will contain the same phasor values. This will provide the flexibility of using less inputs into the neural network during the training process and hopefully speed up training the different ANNs if needed. As for line currents for the parallel lines, each end of the transmission line (at the substation) will contain six magnitude values and six phase angle values. If the voltage phasor duplications are not ignored, the number of input training entries per set of data will be doubled compared to the single transmission line topology. The five output data matrices are the full set (containing all fault types) of fault data. An example of the input data matrix containing voltage and current phasors from one substation is shown in equation 5.1.

$Input\ Data{:}Voltage\ and\ Current\ from\ One\ Substation =$

$$
\begin{bmatrix}
Magnitude\ V_a - Line\ 1 \\
Magnitude\ V_b - Line\ 1 \\
Magnitude\ Vc - Line\ 1 \\
Magnitude\ V_a - Line\ 2 \\
Magnitude\ V_b - Line\ 2 \\
Magnitude\ V_c - Line\ 2 \\
Angle\ V_a - Line\ 1 \\
Angle\ V_b - Line\ 1 \\
Angle\ V_c - Line\ 1 \\
Angle\ V_a - Line\ 2 \\
Angle\ V_b - Line\ 2 \\
Angle\ V_c - Line\ 2 \\
Magnitude\ I_a - Line\ 1 \\
Magnitude\ I_b - Line\ 1 \\
Magnitude\ I_c - Line\ 1 \\
Magnitude\ I_a - Line\ 2 \\
Magnitude\ I_b - Line\ 2 \\
Magnitude\ I_c - Line\ 2 \\
Angle\ I_a - Line\ 1 \\
Angle\ I_b - Line\ 1 \\
Angle\ I_c - Line\ 1 \\
Angle\ I_a - Line\ 2 \\
Angle\ I_b - Line\ 2 \\
Angle\ I_C - Line\ 2
\end{bmatrix}
\qquad (5.1)
$$

The output training target data will also have a change in the output orientation due to the introduction of the parallel line configuration. This change in the training target data is shown in equation 5.2.

$$
Target\ Training\ Data = \begin{bmatrix} Phase\ A - Line\ 1 \\ Phase\ B - Line\ 1 \\ Phase\ C - Line\ 1 \\ Ground - Line\ 1 \\ Phase\ A - Line\ 2 \\ Phase\ B - Line\ 2 \\ Phase\ C - Line\ 2 \\ Ground - Line\ 2 \\ Fault\ Location \end{bmatrix} \tag{5.2}
$$

The first four lines in equation 5.2 are describing the faulted conditions on line 1 of the parallel transmission line topology and the next four data records are providing the faulted condition for the second line. The actual data that will be presented in the training target matrix will follow the algorithm used within the first two phases of this research which is a faulted connection or non-connection classification. If the abnormal faulted condition exists between any phases or ground connections, then that phase or ground entry in the training target data will contain a value of 1. Likewise, the phases or grounding conductors that are not involved with the faulted condition those entries in the training target matrix will contain a value of 0. As an example, for a phase A to ground (LG) fault, see equation 5.3.

$$
AG\ Fault\ on\ Second\ Transmission\ Circuit = \begin{bmatrix} Phase\ A - Line\ 1 \\ Phase\ B - Line\ 1 \\ Phase\ C - Line\ 1 \\ Ground - Line\ 1 \\ Phase\ A - Line\ 2 \\ Phase\ B - Line\ 2 \\ Phase\ C - Line\ 2 \\ Ground - Line\ 2 \\ Fault\ Location \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ FL \end{bmatrix} \tag{5.3}
$$

This full set of data will only be used to train the fault classification ANN. The fault classification ANN is going to take the fault measurement data and predict the type of fault that occurred. Based on the fault classification prediction, it will also be possible to determine which transmission line encompasses the faulted condition. Once the ANN classification determines the type of fault that has occurred, one of the four different fault location ANNs will be enabled to perform the fault location prediction. Once this process is complete the fault identification information will be made available for the field personnel to describe the type of fault that has occurred, which line the fault occurred on, and the predicted fault location based on a reference substation. As for the fault location ANNs, the full set of data can't be used to train the four different fault location ANNs since they will only be used to predict fault location based on the basic type of fault that has occurred. Recall that the four distinct fault location ANNs that are used within this phase will relate to the four different fault location categories (LG, LL, LLG, LLL). In order to train the fault location ANNs, the full set of input and target training data will need to be separated into fault category sets of data. This will result in additional subsets of data for voltage magnitude, voltage phase angles, current magnitudes, current phase angles, and associated target data based on LG, LL, LLG, and LLL fault categories. This process will be known as parsing the fault training data.

5.3 Training the Multiple ANNs for the Parallel Transmission Line Model

Designing and training the fault classification and different fault location ANNs is the next step in providing insights into fault identification with parallel transmission line topologies. The initial intent for this final phase is to analyze fault identification only using single hidden layer ANN as used in phase 2. But as will be shown in section 5.5, not all measurement configurations provide high confidence in predicting fault identification. Therefore, this phase of the research will also provide analysis for using two hidden layer ANNs along with single hidden layer ANNs to predict fault classification and fault location.

Figure 58 and Figure 59 provide the high-level layouts or visual representations for the fault classification ANN structures utilizing a single hidden layer and two hidden layer design respectively. Likewise, Figure 60 and Figure 61 provide the high-level layouts or visual representations for the fault location ANN structures utilizing a single hidden layer and two hidden layer design respectively. As with phase 2, the fault location ANN structure and the fault classification ANN structure are differentiated by evaluating the output layer. Fault location ANNs will only contain one output neuron as compared to the fault classification ANN what will contain eight output neurons since the faulted condition can occur on either one of the parallel lines. The ANN structure figures show only six inputs and should be assumed as only an example for the type of structure. These inputs will increase depending on the measurement configuration used to train the ANNs.
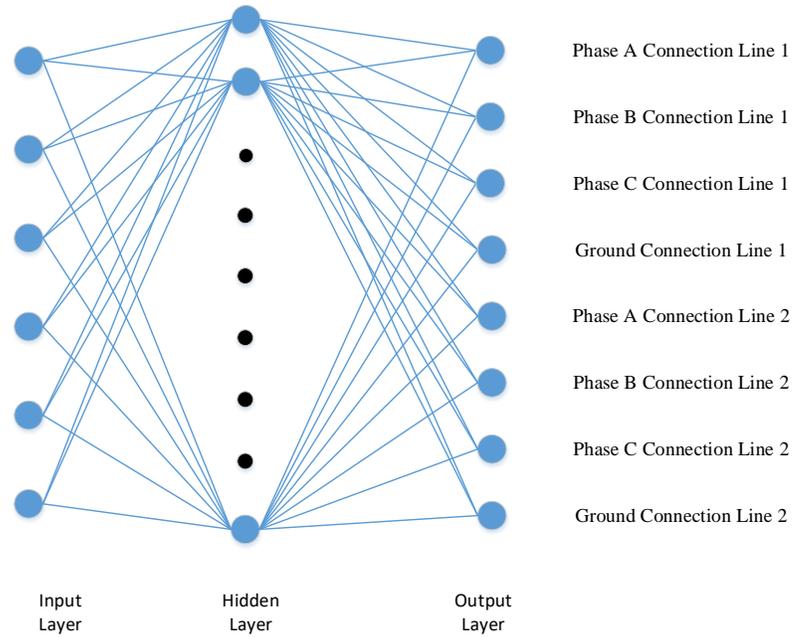
Figure 58 - Fault Classification ANN Single Hidden Layer Structure (Phase 3)
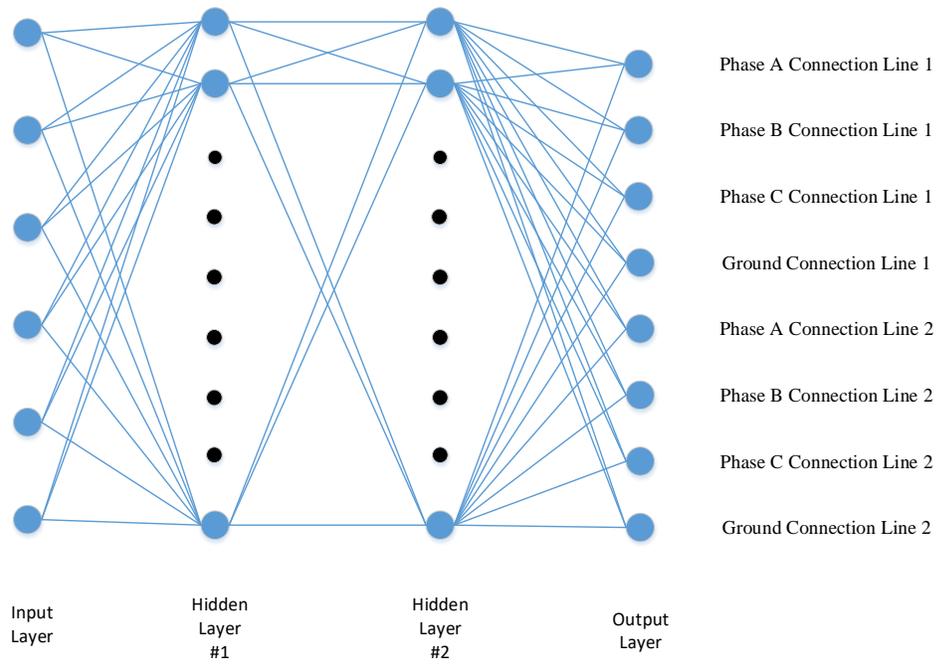
Figure 59 - Fault Classification ANN Two Hidden Layer Structure (Phase 3)
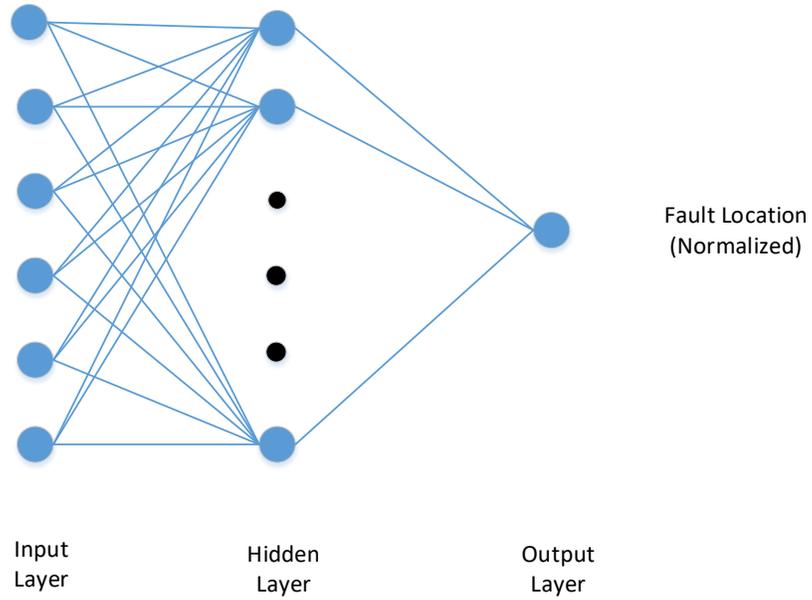
Figure 60 - Fault Location ANN Single Hidden Layer Structure (Phase 3)
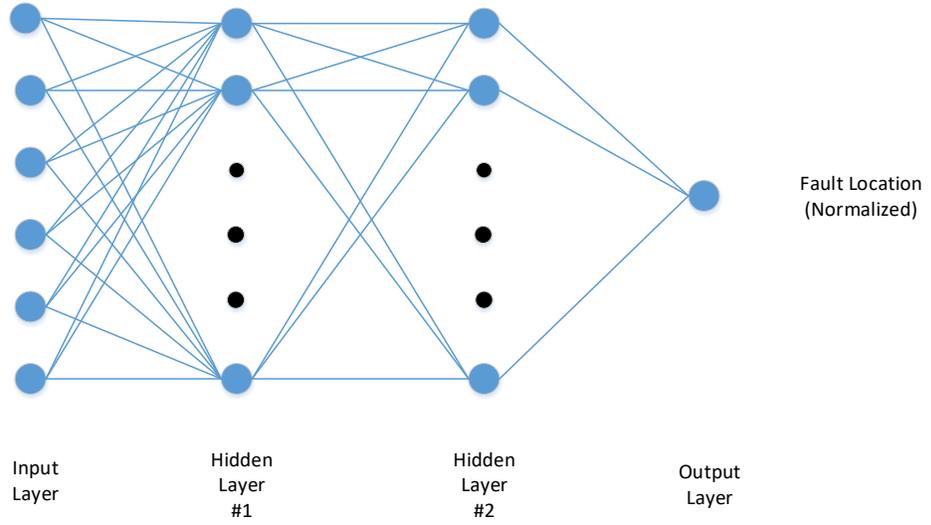
Figure 61 - Fault Location ANN Two Hidden Layer Structure (Phase 3)

The development of the multiple ANNs used in this phase will follow a similar design

approach as discussed in phases 1 and 2. The input layer will only contain data entry

points into each of the developed neural network. When the neural network is trained,

assuming the MATLAB command "train" is used, the neural network will configure the

number of inputs and outputs within each neural network structure which will be based

off the input and target training data presented to the network. This configuration process

is based on the input and target training data that has been presented to the network

during the training process. As discussed in section 5.2 the same measurement

configurations used throughout this research will continue to be analyzed in this phase as

well. But since the transmission line topology contains a second line and the number of

voltage and current measurement are doubled due to the additional transmission line, the

number of inputs per measurement configuration will also double. Table 24 provides the

number of data inputs that will be presented to the ANN during the configuration and

training process.

Table 24 - Number of ANN Measurement Inputs per Measurement Configuration
(Parallel Line Topology)

Number of ANN Inputs per Type of Phasor Measurement for Parallel Line Topology	
Phasor Measurements with Orientation	Number of ANN Inputs
Current (I) @ Substation A	12
Voltage (V) @ Substation A	12
Voltage (V) and Current (I) @ Substation A	24
Voltage (V) and Current (I) @ Substations A and B	48

Since the number of inputs have doubled and the fault classification ANN outputs have

increased, the amount of time to train these networks have increased as well.

The current and voltage phasor measurement that were used as the input training data

contains a magnitude value in per unit and a phase angle value measured in degrees for

each phase. Table 25 provides the maximum and minimum phasor values that were

recorded at the fifth cycle of simulation that is used to train the ANNs. These maximum

and minimum values were extracted from the full set of training data that contained data

from all fault types.

Table 25 - Voltage and Current Phasor Maximum and Minimum Phasor Values (Parallel
Transmission Topology)

	Max Value (per unit)	Min Value (per unit)
Voltage Magnitude	1.0298	0.0520
Voltage Angle	180	-180
Current Magnitude	81.0210	1.5279
Current Angle	180	-180

Comparing Table 25 from the parallel transmission line simulations and Table 16 from

the single transmission line simulations, the data shows that the parallel transmission line

topology provides larger bus voltage magnitudes and larger line current magnitudes.

As mentioned, the hidden layers of the fault classification and fault location ANNs will

both utilize structures of one and two hidden layers of neurons. From the different ANN

hidden layer structures, a range of neurons will be tested to determine the best ANN

structure that can have the highest accuracy of predicting any fault identification. The

range of neurons for the ANN structure using one hidden layer will utilize neurons

ranging between 6 to 36 neurons. While the different ANN hidden layer structures are

developed using steps of neuron through the identified range in multiples of three until all

hidden layer neuron structures have been developed and trained. As for the ANN structures containing two hidden layers, the ANNs will be trained by using a range of neurons between 12 to 21 neurons. Again, multiple ANNs will be developed by modifying the hidden layer neurons by multiples of three until all neuron hidden layer combinations between the neuron limits have been trained. Each hidden layer neuron used within all ANN hidden layer structures will incorporate the use of the hyperbolic tangent sigmoid transfer function.

Finally, the output layer as earlier stated will be configured when training the network with the target training data. If the data were presented to the network as is, either the full set of data or data that have been parsed by fault type, the neural network would contain 9 ANN outputs. This is because the fault type and the fault location target training data has not been separated. Therefore, before any training can take place the target data will need to be separated so that the correct data entries are used for training the fault classification and fault location ANNs. Fault classification will use the first eight rows of data in the target training matrix. The first eight rows will contain discrete values of either a zero or one that describe the fault connection or no-connection status of that phase or ground connection. This algorithm of setting the fault classification target data was discussed in section 5.2.

The ANNs that predict the fault location will be configured to have only one ANN output. The fault location target data will be imbedded in the target training data in the ninth row of data. The fault location value located in the target training data will contain a floating-point value that represents the distance to the fault from the reference substation.

Before training the fault classification and fault location ANNs there needs to be some

pre-processing of the input and target training data. This was the same pre-processing

steps that took place in phases 1 and 2. The input training magnitude values need to be

normalized to the maximum magnitude value recorded in the training input data set.

Phase angle values were normalized to the maximum (positive) phase angle recorded in

the training data. The input training data will be common for both fault classification

ANNs and fault location ANNs. Therefore, the same normalization will take place for

data used with the fault classification or fault location ANNs. As for the target training

data, the fault classification target data will only contain discrete values of either a zero or

one. There will be no normalization with these values. This is not the case for the fault

location. The fault location will contain a floating-point value of the actual fault location

from substation A as a reference point. This fault location was normalized to the total

length of the line. In theory, the fault location can be no longer then the length of the

transmission line. As with other phases of this research, all ANNs were trained using the

training function Levenberg-Marquardt.

5.4 Development of Testing Data for Parallel Transmission Line Model using Multiple ANN Approach

The development of the fault identification testing data for the parallel transmission line

topology followed the approach used in phase 2. There were two different sets of testing

data that was developed for testing the ANN predictability for fault identification. For

each phase of this research, different testing data sets were developed since each phase

had small additions or modifications to the models and ANN development process. The

models that were used to generate the input and target training data for training the ANNs

were the same models that were used to generate the input and target testing data. The

target testing data is not presented to the model, this data is used to validate the results

and to check the ANN predictability error.

While developing the testing data sets, it was decided that the parameters of the faulted

scenarios should not be changed drastically all at once. It was the thought that if the

testing data were changed drastically and the ANNs predictions were poor then it would

be hard to understand why and when the results began to diverge from the actual fault

identification. Instead the testing data was developed by only changing one faulted

scenario parameter at a time. The first set of testing data used a MATLAB random

generator function (rand) to select fault locations on the transmission line that were

different then fault locations used in the training data. Since the step size of moving the

fault down the transmission line was so small (0.1 km) and used within in all phases of

this research the fault locations used for testing incorporated fault locations out to four

decimal points of accuracy. For the first testing data set, the fault resistances were not

changed from the input and target training data. Recall, that this phase of the research

used fault resistances within the fault resistance limits of 1 Ω to 50 Ω, while varying the

steps between the limits at multiples of 10 Ω, 5 Ω, and 2.5 Ω. When only modifying the

fault locations for the 10 Ω separation, the testing input and target data was obtained with

110 different fault locations. But for 5 Ω and 2.5 Ω separation fault resistance step trained

ANNs, testing data was generated only using fault resistances separated by 5 Ω multiples.

For the 5 Ω and 2.5 Ω separation testing data sets, only 60 random fault locations were created for the first data set. This data collection resulted in the entire testing data sets to contain 6,600 testing data points for data sets containing only modified fault locations.

The second set of training data added some more complexity to the first testing data set. Along with the fault locations being varied, the second data sets also varied the fault resistances using the random generator as with the first testing data set. The only stipulation for varying the fault resistances was the random values of the fault resistances had to be between the trained fault resistance limits of 1 Ω to 50 Ω. The fault resistance values were generated by the random generator function in MATLAB (rand). When modifying both the fault locations and fault resistance for the data separation developed ANNs, the full set of testing input and target data was obtained with 25 different fault locations and 26 different fault resistance values. The full testing data results in the testing data sets to contain 6,500 testing data points.

5.5 Results for the Multiple ANN Approach using Parallel Transmission Line Model

This section will be presenting the performance of fault identification using the multiple ANN approach for the parallel transmission line model. This multiple ANN approach will be using one ANN to identify the type of the fault that has occurred on the transmission line. But the developed approach will be using the fault classification ANN to enable one of four different ANNs to identify the predicted location of the fault. This phase of the research will look at both single hidden layer and multi-hidden layer neural

networks to accurately identify fault identification. All evaluated performance metrics used within chapter 4 will be used to evaluate the ANN performances in this chapter as well. This will relate to the fault classification and fault location ANNs being reviewed on maximum absolute error, minimum absolute error, and average absolute error. The fault location error threshold value will again be evaluated at 1 km. The results provided in this section will concentrate on the maximum absolute error and the number of instances over the fault location threshold value. It was again believed that these two sets of metrics can provide sufficient evidence describing the ANNs ability to predict fault identification. Fault classification absolute error is the difference between the actual discrete value of 0 or 1 for the faulted connection or no-connection versus the ANN output that corresponds to that connection. This use of absolute error would be unitless. But when absolute error is used for fault location, which is comparing the actual location of the fault from a reference bus versus the ANNs fault location prediction, both parameters of the comparison are in units of km then the absolute error calculations will be in units of km.

Each ANN trained in this phase of the research was trained with the four measurement configurations that have been discussed throughout this dissertation. Each one of these measurement configurations will be discussed in the following sub-sections. The training data used for training the ANNs used 0.1 km fault steps for the fault data, but the fault resistance steps between the fault resistance limits were varied by 10 Ω, 5 Ω, and 2.5 Ω. Testing data was presented in two sets of data which represent the data by only modifying the fault location to contain random fault locations that were not in the original training data. The second testing data set was developed modifying both the fault

locations and fault resistances with the values selected at random between the original

fault location and fault resistance limits. Each ANN that was tested during this phase of

the research concluded that when both, the fault location and fault resistance, were

modified then the absolute error contained the worst-case values. Therefore, the data

presented in the following sections will present data with random fault locations and

random fault resistances.

5.5.1 Fault Identification Results using Current from Substation A (Phase 3)

The first results that will be presented will represent the measurement configuration of

current phasors being available from one substation. The first ANNs that were tested

were trained with fault resistances that used steps of 10 ohms between the fault resistance

limit of 1 Ω to 50 Ω using a single hidden layer ANN structure. The maximum error

produced by this set of ANNs for fault classification for most of the hidden layer neuron

structures produced at least one phase or ground connection point exceedingly over 50

percent of maximum error. With possible error reaching 50 percent error, the ANN

structures could not be used in this approach since the fault classification ANNs are

determining which fault location ANN to enable. Therefore, the fault classification error

needs to be reduced. Following the trend in this research would state that adding more

training data would reduce the error. Therefore, ANNs that were trained with fault

resistance data of 2.5 Ω steps were used to provide lower results of maximum error. It

was observed with ANN trained with 2.5 Ω fault resistance steps that at low neurons in

the hidden layer the ANN would still produce 50 to 100 percent maximum errors. But

when the hidden layer starts to contain 30 to 36 neurons the error gets reduced to a

maximum level of 40 percent. Table 26 provides the number of tested scenarios that

generated errors greater than 10 percent for each phase or ground connection in each

ANN tested structure. It is shown that having 24 neurons in the hidden layer produce

strongest results.

Table 26 - Fault Classification Error > 10% for I_BusP Single Layer ANN – Phase 3

Fault Classification Instances > 10& for Parallel Line using Multiple ANN Approach (I_BusP, 2.5 ohm Step Size)								
Hidden Neuron	Line 1 Phase A Num of Errors	Line 1 Phase B Num of Errors	Line 1 Phase C Num of Errors	Line 1 Ground Num of Errors	Line 2 Phase A Num of Errors	Line 2 Phase B Num of Errors	Line 2 Phase C Num of Errors	Line 2 Ground Num of Errors
6	211	2052	2352	3850	294	224	3793	4579
9	855	127	1196	2941	117	1591	0	2098
12	100	20	1060	1553	139	0	1424	918
15	54	0	7	315	49	7	40	371
18	4	17	3	561	6	24	1	479
21	69	33	24	47	26	3	25	114
24	0	0	0	1	0	0	0	5
27	58	0	0	82	13	0	0	163
30	7	13	3	40	7	0	3	242
33	14	8	0	52	15	8	1	2
36	3	0	0	24	19	0	0	39

The ANN structure trained with 2.5 Ω fault resistance steps provides better results, but

still there is room for improvement. The next option available is to begin modifying the

ANN structure itself by adding more hidden layers. Therefore, the next evaluation to

lower the error was to look at a two hidden layer structure. When the two hidden layer

structure ANN that was trained with 10 Ω fault resistance steps was used for fault classification the results were excellent. Other than the 12 neurons in the first layer and 21 neurons in the second layer structure, all tested multi-hidden layer structures produced maximum error results less than 5%.

As for the fault location ANNs, using single layer ANNs that were trained with 10 Ω fault resistance steps produced some maximum errors greater near 4 km for the line to ground fault, but the majority of the ANN structures produced less than 2 km for all faults. If the single layer ANNs are modified to be trained with 2.5 Ω fault resistance steps, as expected the results do improve. For all faults, with 27 to 36 neurons in the hidden layer the maximum fault location error is 0.093 km. Figure 62 shows the fault location trend for maximum absolute error using current phasors from one bus with using single layer ANNs.
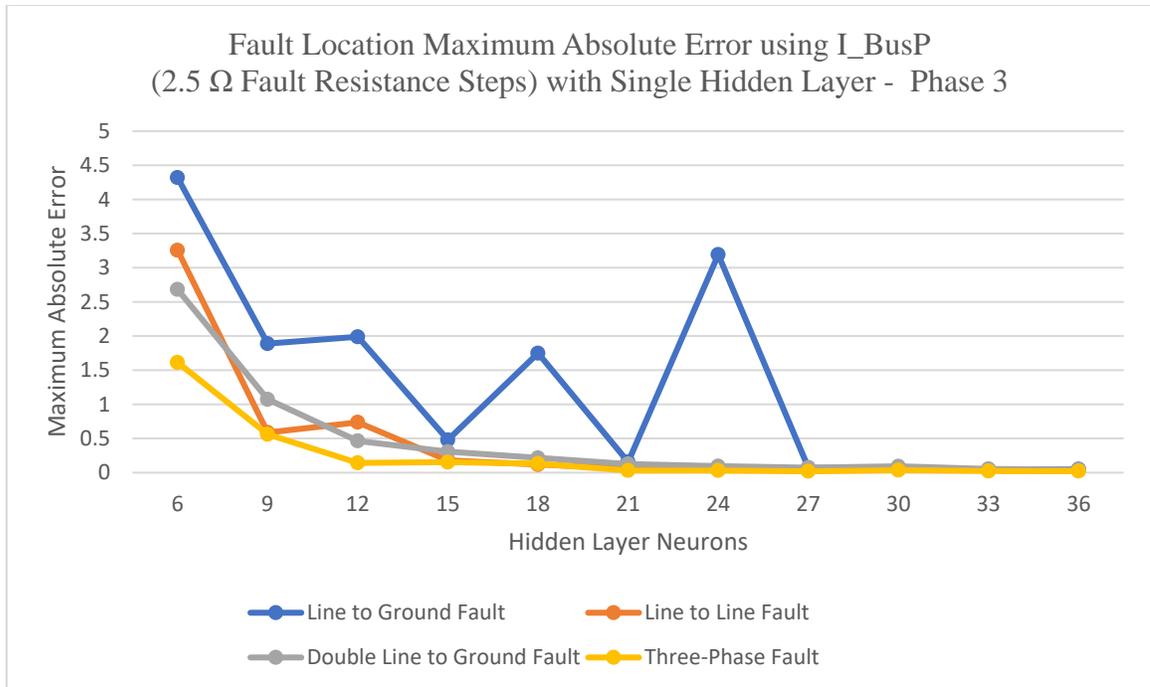
Figure 62 - Fault Location Maximum Absolute Error using I_BusP with Single Hidden
Layer ANN – Phase 3

Using 27 to 36 neurons in the hidden layer produces no tested scenarios over the 1 km

threshold. Even though single layer ANNs have been shown to predict fault location with

small error, these tested scenarios were tested with multi-layer ANNs as well. It was

determined that using multi-hidden layers did not improve the maximum errors. The

results from the multi-hidden layer ANN came out very similar to the single hidden layer

network. The only difference that was identified with using multi-hidden layer ANNs that

have been trained with 2.5 Ω fault resistance steps allowed maximum error under 0.05 km

for every tested ANN structure.

5.5.2 Fault Identification Results using Voltage from Substation A (Phase 3)

The next measurement configuration that will be evaluated on how well ANNs can predict fault identification will be using voltage phasors from one substation. When single hidden layer ANNs that have been trained with 10 Ω fault resistance steps, the maximum error for fault classification hoovers around 60 percent error for 15 to 36 neurons in the hidden layer. This result is very similar to using current phasors from one substation. Which this error is too large to be acceptable and needs to be lowered to use voltage from one bus as an acceptable measurement configuration. When single layer ANNs that have been trained with more training data (2.5 Ω fault resistance steps), the maximum error for fault classification is not drastically improved. The maximum error for 15 to 36 neurons in the hidden layer shifts down to an average of 50 percent. This again is too much error for providing fault classification results to any field personnel. The next option that is available is using multi-hidden layer ANNs to see if that lowers the ANN fault classification maximum errors to a more useful range. As shown in Figure 63 the maximum error doesn't improve with multi-hidden layer either and averages around 53 percent of maximum error. It is concluded that voltage phasors from one bus do not provide accurate fault classification prediction with any tested ANN structure.
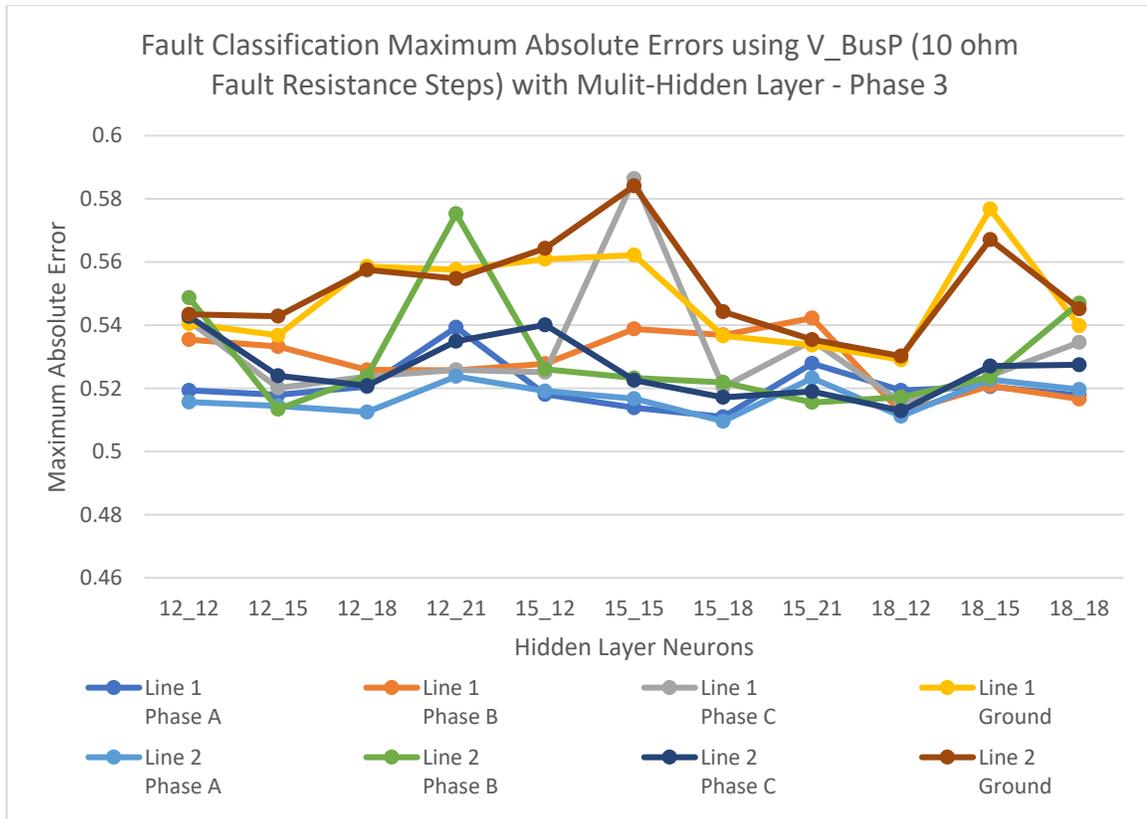
Figure 63 - Fault Classification Maximum Absolute Error using V_BusP with Multi-Hidden Layer ANN - Phase 3

As for using voltage measurements from one substation to predict fault location with ANNs trained with 10 Ω fault resistance steps the results are very interesting. The results show that ground faults provide relatively low maximum errors. The line to ground faults have errors under 4 km for all tested ANN structures and double line to ground faults range between 20 km of error for 12 neurons in the hidden layer to near 0.5 km of error for 15 neurons in the hidden layer. The trends for ground faults can be seen in Figure 64.
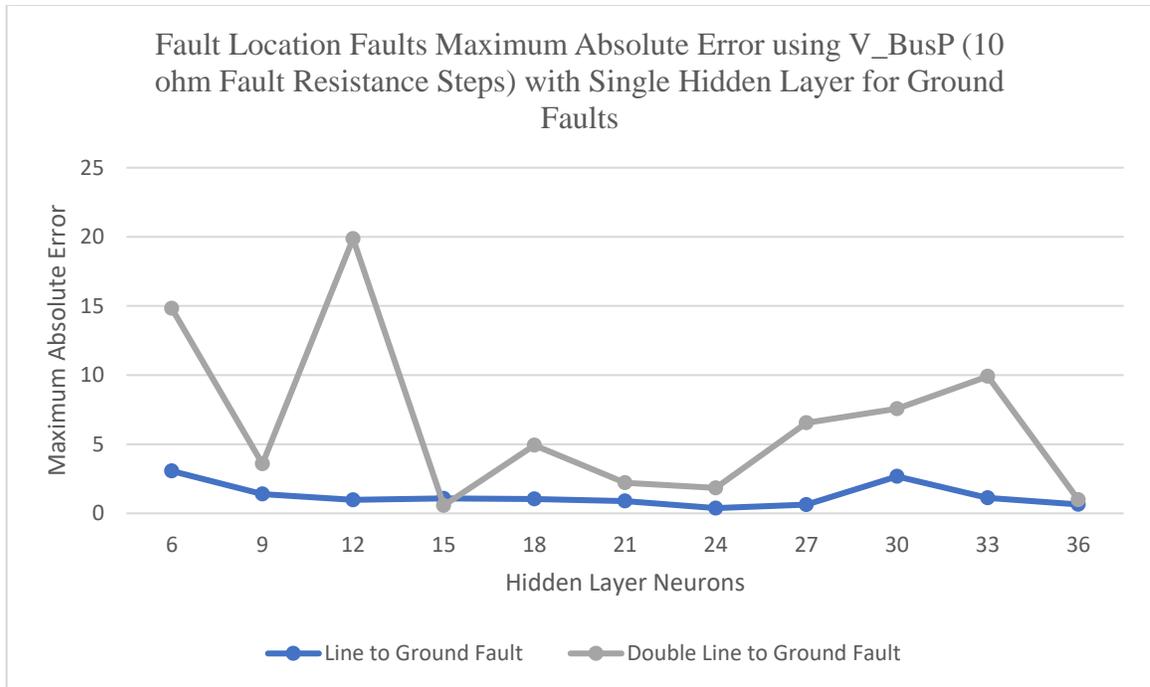
Figure 64 - Fault Location Maximum Absolute Error using V_BusP for Ground Faults
with Single Hidden Layer ANN – Phase 3

Comparing the maximum errors for ground faults versus non-ground faults, the non-ground faults have extremely high error as can be seen in Figure 65. Fault location results are direct outputs the ANN without further process such as limiting it to be between zero and the total line length. This comes back to the point that was made for the fault classification that voltage measurements for parallel transmission lines that share the same substation buses are not suitable for fault identification alone.
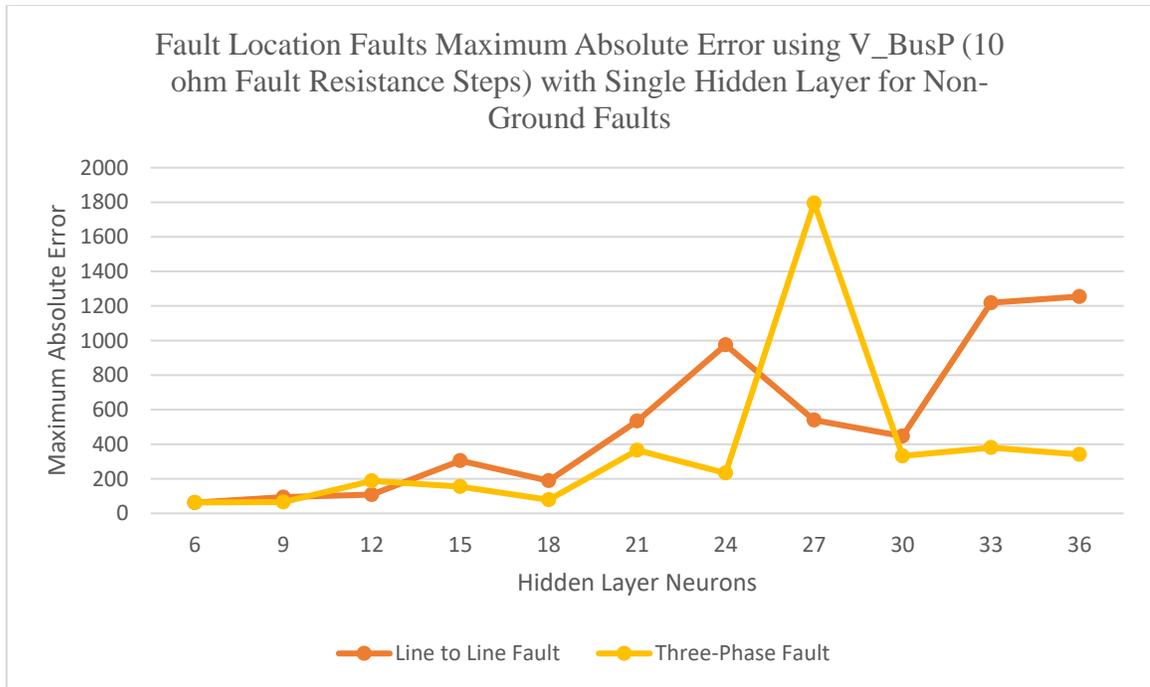
Figure 65 - Fault Location Maximum Absolute Error using V_BusP for Non-Ground
Faults with Single Hidden Layer ANN – Phase 3

Using single hidden layer ANNs and multi-hidden layer ANNs do not improve the results

for non-ground faults. Maximum errors well above the total length of the line are still

recorded.

5.5.3 Fault Identification Results using Voltage and Current from Substation A (Phase 3)

Voltage and current phasors available from only one substation being supplied to the

ANNs is the next measurement configuration to be tested using the parallel transmission

line configuration. When the fault testing data set is applied to the ANNs that were

trained with 10 Ω fault resistance steps, maximum errors that were similar with current

from one substation. The lowest maximum error for all phase and ground connections

occur at 30 neurons in the hidden layer with the performance resulting less than 20

percent error. There are other hidden layer structures that perform just over 20 percent

error.
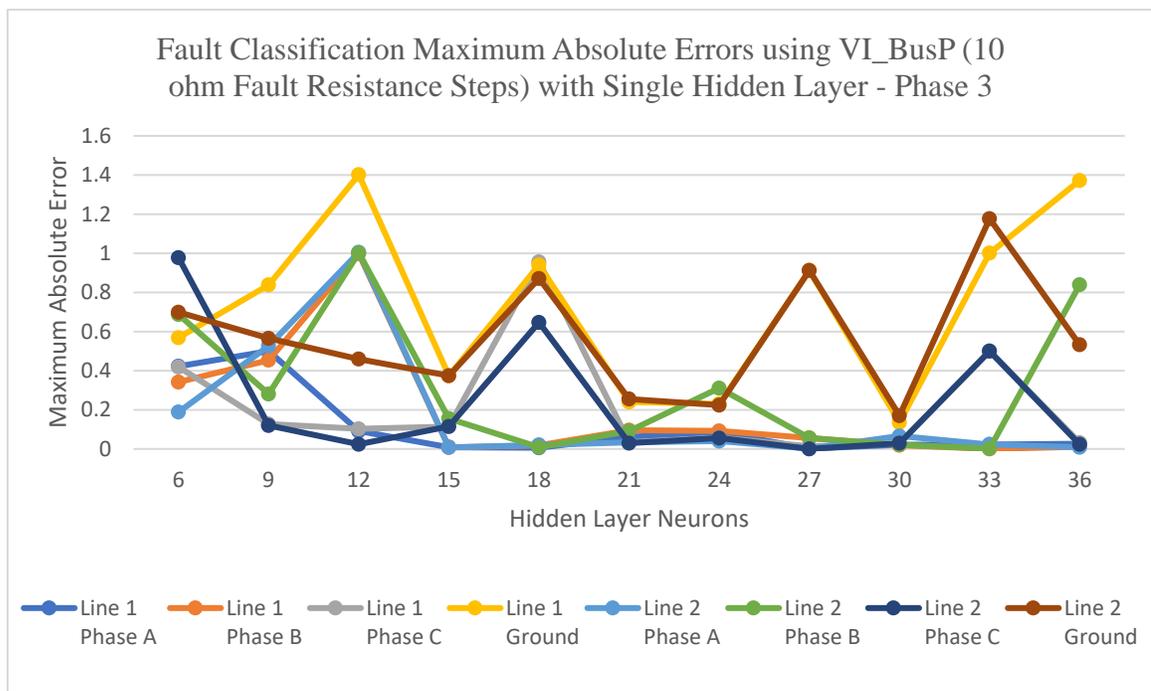
Figure 66 - Fault Classification Maximum Absolute Error using VI_BusP with Single
Hidden Layer ANN - Phase 3

This ANN structure produces maximum errors that are greater than 50 percent. Once the

testing data is presented to the ANNs trained with 2.5 Ω fault resistance steps the

maximum results do improve some. Having 24 or 30 neurons in the hidden layer produce

176

the lowest maximum fault classification errors for all phase and ground connections near 10 percent. As the results are becoming more positive to identifying the fault classification, the question becomes can the results improve. The testing data then was presented to a multi-hidden layer ANN trained with data that contained 10 Ω fault resistance steps. The maximum error results for most of the ANN structures produce extremely low errors. There are two ANN structures that are competing for the best performing fault classification ANN which are 15 neurons in the first hidden layer with 18 neurons in the second hidden layer or 18 neurons in the first hidden layer and 21 neurons in the second hidden layer. These networks are producing errors near 1e-7.

As for fault location using voltage and current measurements from one substation the errors produced by the single layer or multi-layer ANNs continue to be performing at a high level. When single layer ANNs that are trained with 10 Ω fault resistance steps are used to detect fault location for each fault category, the maximum error observed is with the line to ground faults at 5.44 km. But over the full tested spectrum of ANN structures all faults experience around a 2 km error. This results in the ANN predictions having around 10 instances for each type of fault for all ANN tested structures over the 1 km threshold. As the trend continues with this research, more training data or more hidden layers tend to create less errors in the ANN predictions. Therefore, as the test data was presented to the trained ANNs containing 2.5 Ω fault resistance data, the fault location maximum errors for nearly every fault type and ANN structure fall below 1 km. This displays that some simple ANNs can be used to predict fault location with high accuracy. Figure 67 provides the fault location maximum absolute error trend as the number of neurons in the single hidden layer increases.
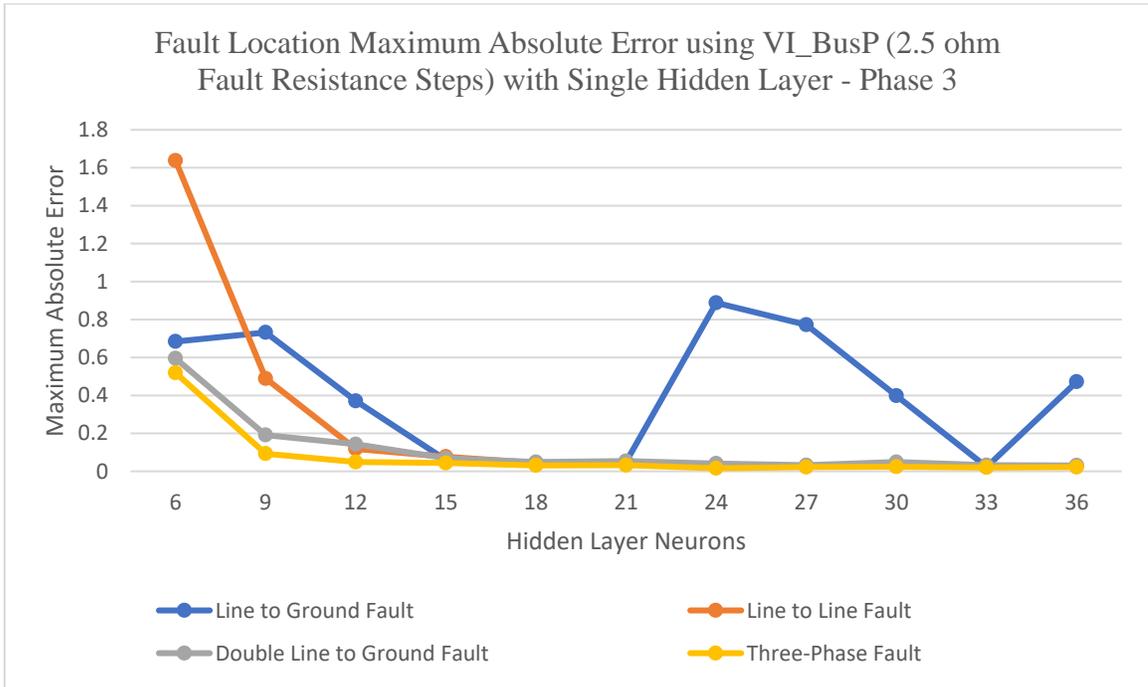
Figure 67 - Fault Location Maximum Absolute Error using VI_BusP with Single Hidden Layer ANN - Phase 3

The best performing ANN structure contains 33 neurons in the single hidden layer of the ANN. This would produce a maximum of 0.03 km error for fault location in all fault types. Using ANNs that contain multi-hidden layer provide results that are very similar to the results presented in Figure 67.

5.5.4 Fault Identification Results using Voltage and Current from Substation A and B (Phase 3)

The last measurement configuration that will be studied with the parallel line topology is using voltage and current phasor measurements from both substations connecting the transmission line. The analysis begins with evaluating the fault classification ANN prediction results. When the test data is presented to the fault classification single hidden layer ANN that was trained with 10 Ω fault resistance step training data, from 15 – 33 neurons in the hidden layer of the ANN produces maximum error for the phase and ground connections near or under 10 percent. There was one ANN structure at 21 neurons in the hidden layer that had maximum error for the second transmission line grounding connection at 50%. This ANN structure only contained 3 instances out of 6500 tested scenarios that exceeded 10 percent error. The single ANN structure trained with 10 Ω fault resistance steps could be used for fault classification ANN predictions and the best ANN structure to best perform would be 18 neurons in the single hidden layer. Even though the ANN structure trained with 10 Ω fault resistance steps contain decent prediction results, the maximum errors or larger range of possible ANNs could possibly be selected for ANNs that have been trained with more training data. Using the ANNs trained with 2.5 Ω fault resistance steps from 21 to 36 neuron produce maximum error results less than 5 percent. The ANN structure of 24 neurons in the hidden layer would provide the best prediction results. Figure 68 provides the overall fault classification maximum errors for each tested ANN structure trained with 2.5 Ω fault resistance steps.
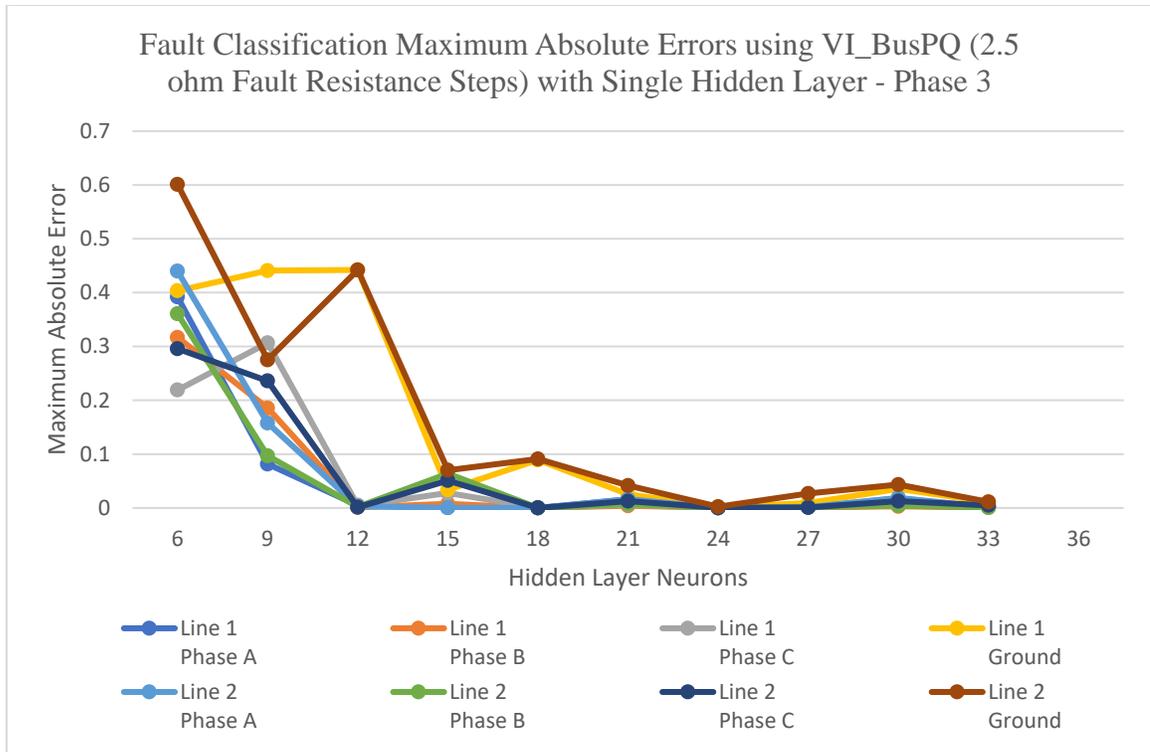
Figure 68 - Fault Classification Maximum Absolute Errors using VI_BusPQ for Single Hidden Layer ANN - Phase 3

From the results that have been provided with the single hidden layer ANN, the simpler network could easily be used to identify the classification of the transmission fault. The only advantage of using a multi-layer ANN would be to reduce the error even further. If multi-hidden layer ANN were used the maximum error could be reduced to 2 percent.

As for fault location, the results that were observed closely resembled the trends for the voltage and current measurements from one bus in section 5.5.3. When ANNs that were trained with 10 Ω fault resistance steps to predict fault location most of the ANN structures had maximum errors less than 0.5 km for all faults except for line to ground faults. Line to ground faults in a few of the ANN structures reached maximum errors over

2 km. But for 9 neurons in the hidden layer the line to ground faults contained maximum

errors under 0.5 km. If the test data was supplied to the fault classification ANN trained

with 2.5 Ω fault resistance steps instead, the maximum errors were reduced for all single

hidden layer ANN structures. ANNs that contained 12 to 21 neurons in the hidden layer

produced maximum errors less than 0.04 km with 15 neurons providing the lowest

maximum errors.
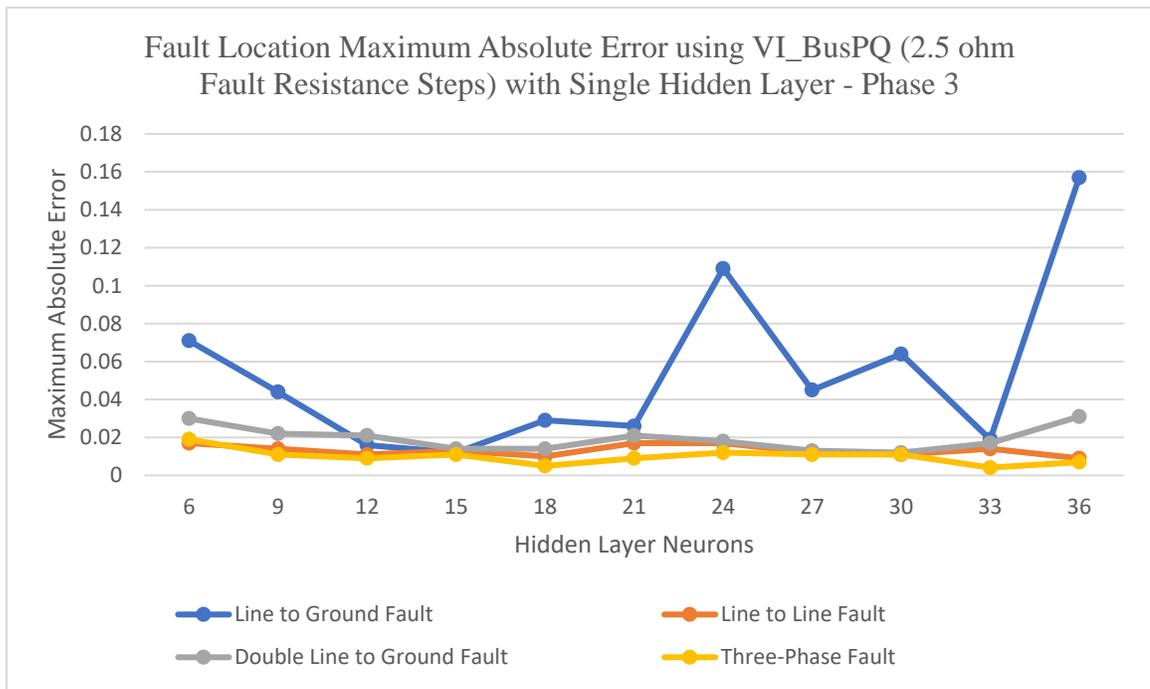
Figure 69 - Fault Location Maximum Absolute Error using VI_BusPQ for Single Hidden

Layer ANN - Phase 3

Using multi-hidden layers did not provide any significant improvements of the maximum

error shown in Figure 69.

Chapter 6 – Research Conclusion

It is important to understand that ANNs can perform fault identification analysis. The factors to consider will include what data is available to be used for training the ANN, how much time it takes to develop and test the ANN, and how accurate do the ANN predictions need to be. Deriving at a finished ANN product can be very time and labor intensive.

To recap this research, it was devoted to learning how to apply artificial neural networks to identify transmission faults that have occurred on the transmission system. Many of the operational situations that can occur, transmission outages, power transfers leading to transmission congestion, and shifts in generation can all cause transmission elements to be further damaged if the proper fault identification techniques are not performed quickly to begin performing system restoration. This research assumed that there was knowledge that a fault had occurred on a specific transmission line. The physical location and the type of fault needed to be identified to provide field personnel with the fault identification information in hopes that the fault can be removed from the system and the restoration process can begin. There were two transmission topologies that were the main focus, single transmission lines and parallel transmission line. These line topologies are very common within the transmission system. There was two different neural network approach that were used within this research in order to provide the lowest maximum errors between the actual fault identification and the ANN predicted fault identification. One of these approaches placed the fault classification prediction and the fault location estimate within the same single ANN structure. This will place both binary outputs of 0

or 1 and a floating-point value within the same ANN output. The second approach splits

up the fault classification and fault location portions of the fault identification problem

into two different sets of ANNs. The first set is a single ANN that is used to predict the

fault classification. The classification ANN is trained with the complete set of fault data

so that any fault type can be determined. This fault classification ANN output then

enables one of four different ANNs. These four different ANNs represent the four basic

fault types of line to ground, line to line, double line to ground, and three-phase faults.

These ANNs were trained with different sets of training data that correspond to the

different fault types. All the training data sets that were developed fluctuate by the

amount of data they hold. The difference in the data sets contained faulted measurement

data that was collected by setting different fault simulation parameters so that the data

sets increased in the amount fault data available. The idea was to compare how more

training data affected the prediction of fault identification.

This research has allowed the following points to be brought forward, so that this

information may help others in using and designing neural networks to identify fault

identification.

- By comparing the ANN results for all three phases of this research, the first thing

  that can be concluded is that no one neural network can be used to predict every

  fault identification problem. These networks need to be developed by a trial and

  error approach one at a time in order to determine the best performing ANN

  structure.

- ANNs can solve multiple problems within the same single ANN. An example of

  this would be placing fault classification and fault location in the same ANN as

was performed in phase 1 of this research. But by putting the two problems together there is a trade off on performance. If the designer's number one concern is not performance this may be the option to go with. But for this research, performance was a number one concern, because fault identification information cannot be relayed to field personnel if dispatchers are using ANN output predictions when they are not confident in the fault identification.

- The developer should try different transfer function with their ANN development. This research used the hyperbolic tangent sigmoid function in all hidden layers and pure linear function in the output layers, but this does not mean that another available transfer function could produce equivalent results.

- ANNs should be trained with different data sets to determine the correct amount of data need for the trained neural network. It was seen during this research that neural networks that were trained with smaller data sets tended to produce higher absolute errors. To improve the errors, more training fault resistance data was needed or more complex networks needed to be used.

Chapter 3 provided the details of the first ANN approach that was used. This approach used a single ANN to identify both the fault classification and fault location. This analysis was performed on the single transmission line model. The fault identification testing data was evaluated using both single hidden layer networks and multi-hidden layer networks. Table 27 provides the ANN structures for both single hidden layer and multi-hidden layer ANNs that produced the lowest maximum errors. It was identified in phase 1 of this research that using 0.05 km fault steps to move the fault down the

transmission line to obtain the training data would improve the results slightly. But using more training data did not provide drastic improvements. The best improvements were made my using multi-hidden layer networks.

Table 27 - Summery of Best ANN Structures (Phase 1)

Summary of Results for Phase 1 - Best ANN Structures with Maximum Errors					
		Fault Classification		Fault Location	
	ANN Configuration	ANN Structure	Maximum Error	ANN Structure	Maximum Error
I_BusP	Single Hidden Layer	36	< 5%	36	7.278 km
	Multi (Two) Hidden Layer	Any tested ANN Structure	< 5%	18_21	3.46 km
V_BusP	Single Hidden Layer	30	< 10%	30	15.276 km
	Multi (Two) Hidden Layer	18_15	< 5%	12_21	6.682 km
VI_BusP	Single Hidden Layer	33	< 5%	33	6.1 km
	Multi (Two) Hidden Layer	21_21	< 0.1%	21_21	2.326 km
VI_BusPQ	Single Hidden Layer	27	< 1%	27	0.56 km
	Multi (Two) Hidden Layer	21_12	< 0.05%	21_12	0.621 km

Chapter 4 provided the details of the second ANN approach that was used. This approach used a multiple ANN configuration to identify both the fault classification and fault location. This fault classification portion used a single ANN to identify the type of fault that had occurred. Then one of four different ANNs, that were trained by the different fault types, were used to detect the fault location. This fault identification analysis was performed on the single transmission line model using only single hidden layer networks. Table 28 provides the ANN structures for the single hidden layer ANNs that produced the

lowest maximum errors for each measurement arrangement. It was identified in phase 2

of this research that training ANNs with more simulated fault data, which consisted of

more fault resistance data, that the results did improve. Each ANN structure presented in

Table 28 came from ANNs that were trained with data containing 2.5 Ω fault resistance

steps.

Table 28 - Summery of Best ANN Structures (Phase 2)

Summary of Results for Phase 2 - Best ANN Structures with Maximum Errors					
		Fault Classification		Fault Location	
	ANN Configuration	ANN Structure	Maximum Error	ANN Structure	Maximum Error
I_BusP	Single Hidden Layer	18	< 1%	24	0.3 km
V_BusP	Single Hidden Layer	24	< 0.05%	36	0.511 km
VI_BusP	Single Hidden Layer	30	< 1.5%	30	0.249 km
VI_BusP Q	Single Hidden Layer	27	< 0.4%	27	0.02 km

Chapter 5 provided the details of the last ANN approach that was used. This was the

same approach that was used in chapter 4 to identify fault classification and fault

location. The difference in this chapter was the model was changed to a parallel

transmission line model. The fault identification testing data was evaluated using both

single hidden layer networks and multi-hidden layer networks.

Table 29 provides the ANN structures for both single hidden layer and multi-hidden layer

ANNs that produced the lowest maximum errors. It was identified in phase 3 of this

research that training ANNs with more simulated fault data, which contained more fault

resistance data, that the results would improve. It also observed in some cases that

changing the ANN structure to a multi-hidden layer ANN lowered the maximum error.

Table 29 - Summery of Best ANN Structures (Phase 3)

Summary of Results for Phase 1 - Best ANN Structures with Maximum Errors						
		Fault Classification		Fault Location		
	ANN Configuration	ANN Structure	Maximum Error	ANN Structure	Maximum Error	
I_BusP	Single Hidden Layer	30	< 20%	33	0.544 km	
	Multi (Two) Hidden Layer	18_15	<0.5%	15_18	0.3376 km	
V_BusP	Single Hidden Layer	N/A	N/A	N/A	N/A	
	Multi (Two) Hidden Layer	N/A	N/A	N/A	N/A	
VI_BusP	Single Hidden Layer	24	0.37%	33	0.03 km	
	Multi (Two) Hidden Layer	15_18	0%	Any tested ANN Structure	< 0.03 km	
VI_BusPQ	Single Hidden Layer	24	< 0.2%	15	0.04 km	
	Multi (Two) Hidden Layer	Any tested ANN Structure	< 0.2%	Any tested ANN Structure	< 0.04 km	

References

[1] S. Kaplan, "Power Plants: Characteristics and Costs," Congressional Research Service, 2008.

[2] U.S. Energy Information Administration, "Electric Power Monthly with Data for October 2019," U.S. Department of Energy, Washington, DC, 2019.

[3] Office of Electricity Delivery and Energy Reliability U.S. Department of Energy, "United States Electricity Industry Primer," U.S. Department of Energy, July 2015.

[4] North American Electric Reliability Corporation (NERC), "Glossary of Terms Used in NERC Reliability Standards," North American Electric Reliability Corporation (NERC), July 2, 2018.

[5] Natioanal Fire Protection Agency (NFPA), National Electric Safety Code (NESC) C2-2007, New York, NY: Institute of Electrical and Electronics Engineers, Inc. (IEEE), 2006.

[6] North American Transmission Forum (NATF), "NATF TPL-001-4 Reference Document," North American Transmission Forum (NATF), 2018.

[7] Federal Energy Regulatory Commission (FERC), "Managing Transmission Line Ratings," Federal Energy Regulatory Commission (FERC), Washington, DC, 2019.

[8] A. J. Wood, B. F. Wollenberg and G. B. Sheblé, Power Generation, Operation, and Control Third Edition, Hoboken, New Jersey: John Wiley & Sons, Inc., 2014.

[9] System Protection and Control Task Force of NERC Planning Committee, "The Complexity of Protecting Three-Terminal Transmission Lines," North American Electric Reliability Council (NERC), September 13, 2006.

[10] United States Department of Energy, "Annual U.S. Transmission Data Review," United States Department of Energy, Washington, DC, 2018.

[11] J. J. G. a. W. D. Stevenson, in *Power System Analysis*, McGraw-Hill Inc., 1994, pp. 380-530.

[12] IEEE Power and Energy Society, "IEEE Standard for Electrical Power System Device Function Numbers, Acronyms, and Contact Designations," IEEE, New York, 2008.

[13] M. T. Hagan, H. B. Demuth, M. H. Beale and O. D. Jesús, Neural Network Design, 2nd Edition; ebook.

[14] M. H. Beale, M. T. Hagan and H. B. Demuth, "2017a MATLAB Neural Network Toolbox User Guide," MathWorks MATLAB, Natick, MA, 2017.

[15] W. Fan and Y. Liao, "Transmission Line Parameter-Free Fault Locaiton Results Using Field Data," 2018 IEEE Power and Energy Conference at Illinois, Champaign, IL, February 2018.

[16] W. Fan and Y. Liao, "Wide area measurements based fault detection and location method for transmission Lines," Protection and Control of Modern Power Systems, 2019.

[17] W. Fan and Y. Liao, "Fault Identification and Location for Distribution Network with Distributed Generations," International Journal of Emerging Electric Power Systems, May 2018.

[18] W. Fluty and Y. Liao, "Electric Transmission Fault Locatoin Techniques using Traveling Wave Method and Discrete Wavelet Transform," 2020 Power Systems Conference, Clemson, SC, March 2020.

[19] T. M. Haque and A. Kashtiban, "Application of Neural Networks in Power Systems; A Review," World Academy of Science, Engineering and Technology, 2005.

[20] J. Gracia, A. J. Mazón and I. Zamora, "Best ANN Structure for Fault Location in Single and Double-Circuit Transmission Lines," *IEEE Transactions on Power Delivery,* vol. 20, no. 4, pp. 2389 - 2395, October 2005.

[21] K. Lout and R. K. Aggarwal, "A Feedforward Artificial Neural Network Approach to Fault Classification and Location on a 132kV Transmission Line Using Current Signals Only," in *International Universities' Power Engineering Conference (UPEC)*, Uxbridge, Middlesex, United Kingdom, 2012.

[22] E. B. M. Tayeb and O. A. A. Rhim, "Transmission Line Faults Detection, Classification and Location using Artificial Neural Network," in *2011 International Conference & Utility Exhibition on Power and Energy Systems: Issues and Prospects for Asia*, Pattaya City, Thailand, 2012.

[23] M. B. Hessine, H. Jouini and S. Chebbi, "Fault Detection and Classification Approaches in Transmission Lines Using Artificial Neural Networks," in *17th IEEE Mediterranean Electrotechnical Conference*, Beirut, Lebanon, April 2014.

[24] T. Bouthiba, "Fault Location in EHV Transmission Lines using Artificial Neural Networks," *International Journal of Applied Math and Computer Science,* vol. 14, no. 1, pp. 69 - 78, 2004.

[25] E. Bashier and M. Tayeb, "Faults Detection in Power Systems Using Artificial Neural Network," American Journal of Engineering Reserach (AJER), 2013.

[26] A. Jain, A. Thoke, R. Patel and E. Koley, "Intercircuit and Cross-country Fault Detection and Classification Using Artificial Neural Network," 2010 Annual IEEE India Conference (INDICON), Kolkata, 2010.

[27] E. Koley, A. Jain, A. Thoke, A. Jain and S. Ghosh, "Detection and Classification of Faults on Six Phase Transmission Line Using ANN," International Conference on Computer and Communication Technology (ICCCT), Allahabad, 2011.

[28] N. Saravanan and A. Rathinam, "A Comparative Study on ANN Based Fault Location and Classification Technique for Double Circuit Transmission Line," 4th

International Conference on Computational Intelligence and Communication Networks, Mathura, 2012.

[29] A. A. Kale and N. G. Pandey, "Fault Detection and Fault Classification of Double Circuit Transmission Line Using Artificial Neural Network," *International Research Journal of Engineering and Technology (IRJET),* vol. 02, no. 08, pp. 1226 - 1229, Nov-2015.

[30] A. A. Jain, A. S. Thoke and R. N. Patel, "Fault Classification of Double Circuit Transmission Line Using Artificial Neural Network," *World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering,* vol. 2, no. 10, pp. 2363 - 2368, 2008.

[31] A. Jain, "Artificial Neural Network-Based Fault Distance Locator for Double-Circuit Transmission Lines," Hindawi Publishing Corporation, 2013.

[32] A. Jain, A. S. Thoke and E. Koley, "Fault Classification and Fault Distance Location of Double Circuit Transmission Lines for Phase to Phase Faults using only One Terminal Data," in *2009 Third International Conference on Power Systems*, Kharagpur, India, December 2009.

[33] P. Warlyani, A. Jain, A. Thoke and R. Patel, "Fault Classification and Faulty Section Identification in Teed Transmission Circuits Using ANN," *International Journal of Computer and Electrical Engineering,* vol. 3, no. 6, pp. 807 - 811, December 2011.

[34] M. T. Hagh, K. Razi and H. Taghizadeh, "Fault Classification and Location of Power Transmission Lines Using Artificial Neural Network," in *The 8th International Power Engineering Conference (IPEC)*, 2007.

[35] MathWorks, "MathWorks Help Center - Three-Phase Fault," [Online]. Available: https://www.mathworks.com/help/physmod/sps/powersys/ref/threephasefault.html. [Accessed January 2019].

[36] MathWorks, "MathWorks Help Center - Relative Tolerance," [Online]. Available: https://www.mathworks.com/help/simulink/gui/relative-tolerance.html. [Accessed 4 March 2020].

[37] North American Electric Reliability Corporation (NERC), "Consideration of the Effects of Mutual Coupling when Setting Ground Instantaneous Overcurrent," North American Electric Reliability Corporation (NERC), 2015.

[38] North American Electric Reliability Corporation (NERC), "Transmission Availability Data System Definitions," North American Electric Reliability Corporation (NERC), Atlanta, GA, January 14, 2013.

[39] H. Yu and B. M. Wilamowski, "Levenberg-Marquardt Training," in *Industrial Electronics Handbook, volume 5 - Intelligent Systems, 2nd Edition* , CRC Press, 2011, pp. 12-1 to 12-15.

[40] R. Caruana, S. Lawrence and L. Giles, "Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping," Neural Information Processing Systems Conference (NIPS) 2000, Denver, CO, 2000.

[41] C. F. Jensen, "Chapter 2 - Fault in Transmission Cables and Current Fault Location Methods," in *Online Location of Faults on AC Cables in Underground Transmission Systems*, Springer, Cham, 2014, p. 7 to 15.

[42] A. Elmitwally, S. Mahmoud and M. H. Abdel-Rahman, "Fault Identification of Overhead Transmission Lines Terminated with Underground Cables," 14th International Middle East Power Systems Conference, Egypt, 2010.

[43] M. A. Shafi'i and N. Hamzah, "Internal Fault Classification using Artificial Neural Network," 4th International Powr Engineering and Optimization Conference (PEOCO2010), Shah Alam, Selangor, Malaysia, June 2010.

[44] V. D. Andrade, "Typical Expected Values of the Fault Resistance in Power Systems," 2010 IEEE/PES Transmission and Distribution Conference and Exposition, Sao Paulo, 2010.

[45] P. R. Nath and N. V. Balaji, "Artificial Intelligence in Power Systems," IOSR Journal of Computer Engineering (IOSR-JCE), 2019.

[46] C. Darab, R. Tarnovan, T. Antoniu and C. Martineac, "Artificial Intelligence Techniques for Fault Location and Detection in Distributed Generation Power

Systems," 8th International Conference on Modern Power Systems (MPS), Cluj Napoca, Romania, 2019.

[47] J. C. Stacchini de Souza, M. A. P. Rodriques, M. T. Schilling and M. B. Do Coutto Filho, "Fault Location in Electrical Power Systems Using Intelligent Systems Techniques," IEEE Transactions on Power Delivery, Vol 16, No. 1, 2001.

[48] A. Tziouvaras, H. J. Altuve and F. Calero, "Protecting Mutually Coupled Transmission Lines: Challanges and Solutions," 68th Annual Georgia Tech Protective Relaying Conference, Atlanta, Georgia, 2014.

[49] M. Kezunovic, "Use of Intelligent Techniques for Analysis of Faults and Protective Relay Operations," 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, 2007.

[50] C. Asbery and Y. Liao, "Electric Transmission System Fault Identification using Modular Artificial Neural Networks for Single Transmission Lines," in *2020 Power System Conference*, Clemson, SC, March 10-13, 2020.

[51] C. Asbery and Y. Liao, "Electric Transmission System Fault Identification using Artificial Neural Networks," in *2019 International Energy & Sustainability Conference*, Farmingdale, NY, USA, October 17-18, 2019.

[52] Mathworks MATLAB R2016a (Version 9) [Software], [Online]. Available: http://www.mathworks.com/products/matlab/. [Accessed August 2017].

[53] Mathworks MATLAB R2019a (Version 9) [Software], [Online]. [Accessed August 2019].

Educational Institutions

August 2015 – Present

    Ph.D. Student

    Department of Electrical and Computer Engineering, University of Kentucky,

    Lexington, Kentucky, USA

August 2010 – December 2012

    Master of Science in Electrical Engineering

    Department of Electrical and Computer Engineering, University of Kentucky,

    Lexington, Kentucky, USA

August 2005 – December 2008

    Bachelor of Science of Electrical Engineering

    Department of Electrical and Computer Engineering, University of Kentucky,

    Lexington, Kentucky, USA

August 2002 – May 2004

    Associate in Applied Science

    Kentucky Community and Technical College System

    Danville, Kentucky, USA

Professional Positions

February 2014 – Present

    Senior Electrical Engineer

Louisville Gas & Electric and Kentucky Utilities, Louisville, KY, USA

January 2013 – January 2014

Engineering Staff

Honda Manufacturing of America, Inc., Marysville, OH, USA

March 2009 – December 2012

System Engineer

Inter County Energy Cooperative, Danville, KY, USA

<u>Awards and Licensures</u>

Professional Engineering License, State of Kentucky, October 2015

Power Energy Institute of Kentucky (PEIK) Certificate, December 2012

Eta Kappa Nu member