Protection-based Distributed Generation Penetration Limits on MV feeders

Using Machine Learning

By

Emmanuel Nxumalo

NXMEMM002

Department of Electrical Engineering
University of Cape Town

Supervised by:

Mrs Kehinde O. Awodele

Department of Electrical Engineering
University of Cape Town

September 2021

Dissertation submitted to the Department of Electrical Engineering at the University of Cape Town in partial fulfilment of the academic requirements for a Master of Science degree in Electrical Engineering

Key Words: Protection Philosophies, Distributed Generation; Network Planning; Photovoltaic (PV) Opportunity Networks; Protection Based Penetration limits; Protection Miscoordination; Recommendation Systems

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the IEEE convention for citation and referencing. Each contribution to, and quotation in, this master's project report from the work(s) of other people, has been attributed, cited and referenced.
3. I declare that all the work in the document, except for the work, which is properly acknowledged, and the normal guidance of my supervisors, is my own work.
4. I have not allowed, and will not allow, anyone, to copy my work or part thereof, with the intention of passing it off as their own work.
5. I know the meaning of plagiarism and declare that all the work in the document, save for that which is properly acknowledged, is my own. This thesis/dissertation has been submitted to the Turnitin module (or equivalent similarity and originality checking software) and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor.

Name: **EMMANUEL NXUMALO (NXMEMM002)**

Signature: Signed by candidate **Date:** 20 September 2021

Acknowledgements

Firstly, I would like to thank my supervisor, Mrs K Awodele for her guidance and knowledge throughout this research project. Thank you for your time, motivation, and for pushing me to my fullest potential.

I would like to gratefully acknowledge the Eskom network planning department and, the University of Cape Town (UCT) Department of Electrical Engineering for providing the support and infrastructure for conducting this research. The help of Mr Sanjian Malapermal in sourcing distribution network data is gratefully acknowledged.

I would also like to thank the UCT Power Research Group; their guidance and positive criticism on the research project shaped my approach and made me a better researcher. Also, I would like to extend my gratitude to Eskom (EPPEI) for partial financial support.

I am equally thankful and forever grateful to my family; Thandiwe Angel Nxumalo, Dineo Nxumalo, Happiness Portia Silinda, Phuloso Silinda, Confidence Silinda, Dion Moeng, and Derrick Khoza for believing in, and supporting me. I am grateful I had a strong woman like my mother who took both parenting roles and did everything she could to make sure I had an opportunity to succeed, pursue my dreams and goals. Your love, hard work, and teachings are what made me the person I am today.

Thank you to all my friends for being there every step of the way; your friendship was in all my achievements at UCT.

Lastly, I would like to acknowledge Analisa Nokwanda Ndebele. Your encouragement, support, and love got me to yet another achievement. Thank you for being forever in my corner and for pushing me to my fullest potential.

Abstract

The rise of disruptive technologies and the rapid growth of innovative initiatives have led to a trend of decentralization, deregulation, and distribution of regulated/centralized services. As a result, there is an increasing number of requests for the connection of distributed generators to distribution networks and the need for power utilities to quickly assess the impacts of distributed generators (DGs) to keep up with these requests.

Grid integration of DGs brings about protection issues. Current protection systems were not designed for bi-directional power flow, thus the protective devices in the network lose their ability to perform their main functions. To mitigate the impact of distributed generation (DG), some standards and policies constrain the number of DG that can be connected to the distribution network. The problem with these limits is that they are based only on overload and overvoltage, and do not adequately define the DG size/threshold before the occurrence of a protection issue (NRS 097-2-3). The other problem with distributed generation is the vast difference in the technology, location, size, connection sequence, and protection scheme requirements which results in future DG network planning inadequacies – The **Network DG Planning Dilemma**.

To determine the amount of DG to connect to the network, a detailed analysis is required which often involves the use of a simulation tool such as DIgSILENT to model the entire network and perform load flow studies. Modelling networks on DIgSILENT is relatively easy for simple networks but becomes time-consuming for complex, large, and real networks. This brings about a limitation to this method, planning inadequacies, and longer connection approval periods. Thus, there is a need for a fast but accurate system-wide tool that can assess the amount of DG that can be connected to a network.

This research aims to present a technique used for calculating protection-based DG penetration limits on MV networks and develop a model to determine medium voltage opportunity network maps. These maps indicate the maximum amount of DG that can be connected to a network without the need for major protection scheme changes in South Africa. The approach to determining protection-based penetration limits is based on supervised machine learning methods. The aim is to rely on protection features present in the distribution network data i.e. fault level, Inverse Definite Minimum Time (IDMT) curve, pick-up current settings, Time Multiplier Settings (TMS), calculated relay operating times and relay positions to see how the network responds at certain DG penetration levels ('actual' relay operating times). The dataset represents carefully anonymized distribution networks with accepted protection philosophy applied. A supervised machine learning algorithm is applied after nontrivial data pre-processing through recommendation systems and shuffling. The planning dilemma is cast into three parts: the first part is an automated pattern classification (logistic regression for classification of protection miscoordination), the second part involves regression (predicting operating time after different levels of DG penetration), and the last part involves developing a recommendation system (where, when and how much photovoltaic (PV) DG will be connected).

Gradient descent, which is an optimisation algorithm that iterates and finds optimal values of the parameters that correspond to the local or global minimum values of the cost function using calculus was used to measure the accuracy of each model's hypothesis function. The cost function (one half mean squared error) for the models that predict 'actual' relay operating times before DG penetration, at 35%, 65%, and 75% DG penetration converged to values below 120, 20, 15, and 15 seconds^2, respectively, within the first 100 iterations. A high variance problem was observed (cross-validation error was high and training error was low) for the models that used

all the network protection features as inputs. The cross-validation and training errors approached the desired performance of 0.3±0.1 for the models that had second-order polynomials added.

A training accuracy of 91.30%, 73.91%, 82.61%, and a validation accuracy of 100%, 55.56%, 66.67% was achieved when classifying **loss of coordination**, **loss of grading and de-sensitization, respectively**. A high bias problem was observed (cross-validation error was high and training error was high) for the loss of grading classification (relay positions eliminated) model. When the models (horizontal network features) were applied to four MV distribution networks, loss of coordination was not predicted, the loss of grading model had one false positive and the de-sensitization model had one false negative. However, when the results were compared to the vertical analysis (comparing the operating times of upstream and downstream relays/reclosers), 28 points indicated a loss of coordination (2 at 35%, 1 at 65% and 25 at 75% DG penetration). Protection coordination reinforcements (against loss of grading and desensitization) were found to be a requirement for DG connections where the MV transformer circuit breaker TMS is between **0.5 and 1.1**, and where the network fault level is between **650 and 800A.**

Distribution networks in affluent neighbourhoods similar to those around the Western Cape- Somerset West area and Gauteng- Centurion area need to be reinforced to accommodate maximum DG penetration up to the limit of 75% of the After Diversity Maximum Demand (ADMD).

For future work, the collection of more data points (results from detailed analytical studies on the impact of DG on MV feeders) to use as training data to solve the observed high variance problem is recommended. Also, modifying the model by adding upstream and downstream network features as inputs in the classification model to solve the high bias problem is recommended.

Contents

Figures

Tables

Acronyms

ARC	Auto-Recloser
ANN	Artificial Neural Network
ADMD	After Diversity Maximum Demand
CB	Circuit Breaker
DER	Distributed Energy Resources
DG	Distributed Generation/Generator/s (Both singular and plural form depending on context)
DOCR	Directional Overcurrent Relay
DRIVE	Distribution Resource Integration and Value Estimation
EG	Embedded generation
EPRI	Electric Power Research Institute
HB	Herman-Beta
HC	Hosting Capacity
kWp	Kilowatt peak
LV	Low Voltage
MAIFI	Momentary Average Interruption Frequency Index
MBI	Model-Based Interface
MVON	Medium Voltage Opportunity Network
NRS	National Rationalised Specification
NDP	Network Development Plan
OCR	Overcurrent Relay
RES	Renewable Energy Sources
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SCC	Short Current Capacity
SSEG	Small Scale Embedded generation
SBDG	Synchronous-based DG
SVM	Support Vector Machine
TCC	Time-current coordination
TCCC	Total Clearing Characteristic Curve
TRF	Transformer

1. Introduction

1.1. Background of the problem

The rise of disruptive technologies and the rapid growth of innovative initiatives have led to a trend of decentralization, deregulation, and distribution of regulated/centralized services [1]. Good examples of resource distribution include Uber and Airbnb, one of the biggest taxi companies and one of the biggest holiday home companies in the world, respectively. Both these companies do not own most of the cars used as taxis or the property used as holiday homes. This trend brings about the need to investigate and study current de-centralized, deregulated, smart power systems (grids).

Also, the decrease in the demand for electricity in certain areas due to off-grid distributed generation (DG) installations and the increase in the magnitude and complexity (power factor correction) of the demand of electricity in most areas brings about the need to adapt transmission and distribution networks. Power utilities like the South African state-owned company (Eskom), spend substantial amounts of money in the expansion of transmission and distribution facilities. The process of building transmission lines takes years due to servitude acquisition, public participation on route selections, and the decreasing supply of steel. For urban areas, clearances, and aesthetics affect the construction of HV transmission networks. It can be argued that these facilities may be underutilized in the future due to the shift towards decentralized power systems with DG. This shift is a result of the global shift towards sustainable development including reductions in carbon footprint, energy cost, and load shedding impacts, as well as government incentives to increase renewable energy penetration, the decline in the price of solar PV cells, and the need for high electricity supply reliability and quality.

A solution to un-constraining constrained distribution networks while minimizing the cost and delays that come with the construction of HV transmission networks is to integrate DG close to load centres. DG penetration will result in distribution networks progressively becoming complex due to bidirectional power flow. Current electrical distribution systems are based on a topology whereby electricity/energy flows radially from the substation to the consumer/loads. The protection schemes currently in use are not adapted to the existence of power injections distributed throughout distribution networks [2]. The increase in the requests from consumers to connect small renewable energy generators to distribution networks, lead to the need for a fast but accurate system-wide DG penetration impact analysis tool/technique. Hence, there is a need for a fast evaluation of DG source connection points with minimum impact on current electrical distribution networks, to satisfy utility and societal expectations.

Protective systems are designed such that lower short circuit levels are experienced by downstream devices compared to the short circuit levels that are upstream (unidirectional power flow). The original overcurrent, earth fault, and sensitive earth fault protection coordination designs were not performed with DG penetration considerations. DG penetration leads to an increase in short circuit levels that result in a change in the short circuit levels experienced by upstream and downstream protective devices relative to the DG. Therefore, for certain levels of DG penetration (size and DG type), alteration of the system short circuit current characteristics result in loss of protection coordination [3], [4]. In these situations, the protection system does not operate as expected.

Another problem with DG is the vast difference in the technology, location, size, connection sequence, and protection scheme requirements which result in future DG network planning inadequacies – The **Network DG Planning Dilemma**. To determine the amount of DG to connect to the network, a detailed analysis is required: which often involves the use of a simulation tool such as DIgSILENT to model the entire network and perform load flow studies. Modelling networks on DIgSILENT is relatively easy for simple networks but becomes time-consuming for complex, large, and real networks. Thus, this method is limited and could cause planning inadequacies and longer connection approval periods. There is a need for a fast but accurate system-wide tool that can assess the amount of DG that can be connected to a network.

This dissertation presents a machine learning technique used for calculating protection-based DG penetration limits on MV/LV networks, and a model to determine medium voltage opportunity network maps which indicate the maximum amount of DG that can be connected without the need for major protection scheme changes in South Africa.

1.2. Purpose of the study

This research project investigates the effects of network protection parameters on relay operating time after DG penetration, develops machine learning models to determine protection miscoordination groups after DG penetration on MV distribution networks, and recommend where, when and how much PV will be connected. The results are used to create a protection-based DG penetration limit Medium Voltage Opportunity Network (MVON) guide: referring to current network protection parameters that determine the protection-based DG penetration level. The MVON guide aims to eliminates the cost involved and the need to change the current protection settings/devices every time a DG is introduced to the distribution network.

The study will also be used to develop an adaptive planning approach for DG penetration which will aid in planning and expansion studies for the Network Development Plan (NDP).

1.2.1. Objectives of this study

The following objectives were drawn:

1. Analyse and identify the impact of PV DG connection scenarios on distribution networks protected using Eskom's MV/LV protection philosophy.
2. Identify all the network parameters and variables that contribute to DG penetration limits.
3. Develop and train a machine learning model to calculate relay operating times before and after DG penetration and classify resulting protection miscoordination groups that determine protection-based DG penetration limits.
4. Identify a methodology/model for assessing feeders with high DG penetration limits.
5. Establish protection-based DG penetration limits for networks protected by Eskom's MV protection philosophy standard (considering past and future DG connections).
6. Identify improvements to network planning standards to account for changing DG installation scenarios.

1.2.2. Hypothesis

Machine learning value estimation and recommendation techniques are used in this study. To ensure that the method is correctly implemented into a commercial Power system planning tool: the underlying penetration limit, basis, criteria, performance are rigorously checked and tested.

Therefore, to achieve the aim of this research project, the following hypothesis is tested:

Machine learning techniques – value estimation and recommendation systems – can be modified, trained, and applied as a tool to determine/analyse protection-based penetration limits on distribution networks protected using Eskom's MV protection philosophy.

1.2.3. Research questions

The following research questions will be used to assess the validity of this hypothesis:

1. What is the impact of DG on feeders with protection systems based on Eskom's MV protection philosophy?
2. What are the factors that affect DG planning and connection approvals?
3. What are the tools/concepts used to establish a network's DG hosting capacity?
4. What combination of protection features determine different protection coordination problems?
5. Where, when, and how much PV DG is/will be connected?
6. What are the protection-based DG penetration limits for MV networks protected by Eskom's MV protection philosophy?

1.3. Scope and limitations

This research focuses on the impact of increasing DG penetration on relay/recloser protection coordination. Three phase-fault data from a detailed DG penetration analysis study were used to train the machine learning value-estimation model. The model was applied to data from four real power distribution networks. PV installation, date, size, and location data up to early 2019 were used in the recommendation system. Although the study of DG penetration limits on power system includes control, management, implementation, laws, voltage, loading, and more, this project focuses only on the total power that can be injected by a DG on a system protected according to Eskom's protection philosophy and DG network planning. The analysis was carried out on DIgSILENT Power Factory, Octave, Jupyter notebooks, and Excel.

The extent of the work done as building blocks to the topic investigated in this research project is as follows:

- Validation of the model success criteria and classes by the Electric Power Research Institute (EPRI) system-wide DG assessment protection category.
- The use of results from a detailed DG analysis study with a variety of test scenarios to train the model.
- The use of polynomials, elimination, and the combination of both (adding polynomials and elimination) techniques to select the best hypothesis function (a line that fits the data well).

1.4. Dissertation outline

Chapter 1 begins with a description of the research background and motivation: the increase in the requests from consumers to connect small renewable energy generators to distribution networks, and the need for a fast but accurate system-wide DG penetration impact analysis tool/technique. It then states the purpose of the research project, objectives, hypothesis and the research questions.

Chapter 2 reviews the literature on current LV and MV network protection philosophies, future/proposed network protection philosophies, current tools used to determine DG penetration limits, and DG network planning methods. Applications of machine learning in power systems are also reviewed.

Chapter 3 presents the method used in this research project, introduces the MV distribution network case files, PV data, and global horizontal irradiation map of South Africa.

In **Chapter 4,** Eskom's MV protection philosophy and the steps of applying the philosophy to a distribution network including network reduction and modelling are comprehensively described. Chapter 4 also introduces the models and parameters of four real distribution networks used in this study. It then focuses on the application of the MV protection philosophy to these network models. The training data model is described in this chapter.

Chapter 5 presents the application of recommendation systems on the PV and load data. The most common PV size, time of day where PV generation is at maximum, location with the highest PV, time of day with a correlation to a time where PV is at maximum generation and load is at a minimum, area PV installation growth and most common PV sizes per postal code are recommended. The implications of these recommendations in DG network planning are discussed.

Chapter 6 describes in detail the development of the value estimation model. The section begins with the definition of the model success criteria and classes. It then focuses on the use of linear regression for an operating time predictor and the model's learning curves. Logistic regression is then used to classify protection miscoordination. Classifier confusion matrixes are then presented. The performance of the best models with regards to DG protection penetration limits is discussed.

Chapter 7 presents the results of the application of the best model to the four real networks. Feeder/relay locations with the highest and lowest penetration limit are identified.

Chapter 8 contain the conclusion of this study; the extent to which the investigation and modelling done in this research project answers the research questions and validates the hypothesis.

In **Chapter 9,** recommendations are presented.

Appendix A lists the publications from the research**.**

2. Literature review

This chapter presents the literature review of this research project. It briefly describes current LV and MV network protection philosophies, the impact of DG on these protection philosophies, future/proposed LV and MV network protection philosophies, tools/algorithms used to establish DG hosting capacity on distribution networks and network planning for DG.

2.1. Introduction

A protection philosophy for LV/MV distribution feeders is designed for optimal performance in protecting lives and equipment, optimizing quality and reliability of supply, minimizing life cycle costs and minimizing fire risks due to incandescent particles from clashing conductors [5]. A typical protection scheme consists of overcurrent, earth fault, and sensitive earth fault protection. Earth fault protection operates on residual current, the sum of the three-phase currents, and it is insensitive to load current. High impedance faults caused by live conductors touching the ground or trees are known as sensitive earth faults [6]. Sensitive earth fault protection is mainly for the prevention of public fatality/injury in the case of contact with live equipment [6]. In most cases, MV networks are resistively earthed at the substation using the NEC/R, the earth fault current contribution per neutral earthing point is 360 A for rural networks and 960 A for urban networks [7]. The functions of any protection scheme on a power system are to detect and clear the faulted section selectively in a short period. From the coordination requirements, it is not always possible for the protection system to operate fast enough to safeguard life in the case of direct contact with live lines.

2.2. Current LV and MV network protection philosophies

Multiple protection schemes exist; two are commonly used globally i.e. overcurrent relays and recloser-fuse schemes [2]. MV and LV network feeders are mostly protected by non-directional inverse overcurrent relays. inverse overcurrent relays are selective, reliable, and cost-effective, hence their wide use. They consist of two main settings: time multiplier setting (TMS) and plug setting multiplier (PSM) which determine the operating time of the relay [8]. LV and MV network protection philosophies can be grouped into five groups:

2.2.1. Fuse-fuse coordination

Fuses provide a relatively inexpensive and effective method for clearing fault currents. Fuses are normally applied as single-phase devices, only one fuse blows in the event of a phase to phase fault. The devices downstream remain live because of phase to phase connections of MV/LV transformers downstream via phases whose fuses remain intact (dead-side return situation) [6]. Fuse coordination involves two main characteristics of a fuse: minimum melting (MM) - the time required for the fuse to begin melting and total clearing time (TC). A typical fuse-fuse radial distribution feeder is shown in Figure 2-1.

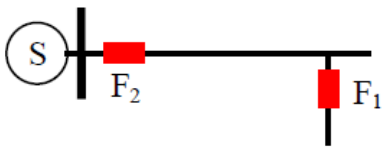


Figure 2-1: Fuse-fuse arrangement on a radial distribution feeder [9].

For typical fuse-fuse coordination (Figure 2-1), Fuse 2 protects the main feeder, and fuse 1 protects the lateral feeder. The system is coordinated such that the desired selectivity is achieved, for a fault on the lateral feeder, fuse 1 must melt before fuse 2. The TC characteristics of fuse 1 must be below the MM characteristics of fuse 2. This configuration works well when the fault levels are within the coordination range (Figure 2-2) [9].

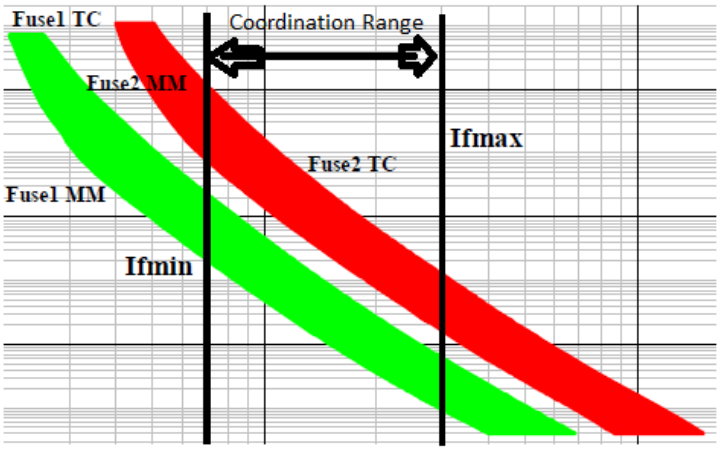


Figure 2-2: Fuse-fuse coordination range [9]

This type of coordination only works with short feeders. For long feeders, it becomes difficult to find fuses with the MM and TC curves that will allow correct grading between the fuses. The grading of fuse-fuse coordination is jeopardized by the replacement of fuses with the incorrect type following the blowing of a fuse. Fuses degrade over time due to exposure to light which may lead to incorrect operation. Expulsion fuses often cause fires due to the expelled molten metal that causes ignition. Also, fuses bring about a dead-side return situation and cause voltage dips due to slow operating times at low fault currents [6]. Fuses also affect the System Average Interruption Duration Index (SAIDI) and the System Average Interruption Frequency Index (SAIFI) due to the time required to replace a blown fuse.

2.2.2. Auto-recloser fuse coordination

For this type of protection scheme, the recloser protects the main feeder while the fuse is used to protect lateral feeders (Figure 2-3). A recloser operates according to two curves, a fast and a slow response (Figure 2-4).

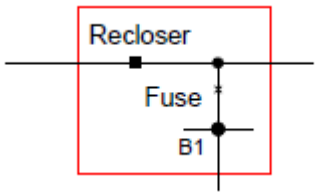


Figure 2-3: Recloser – fuse arrangement on a radial feeder [9]

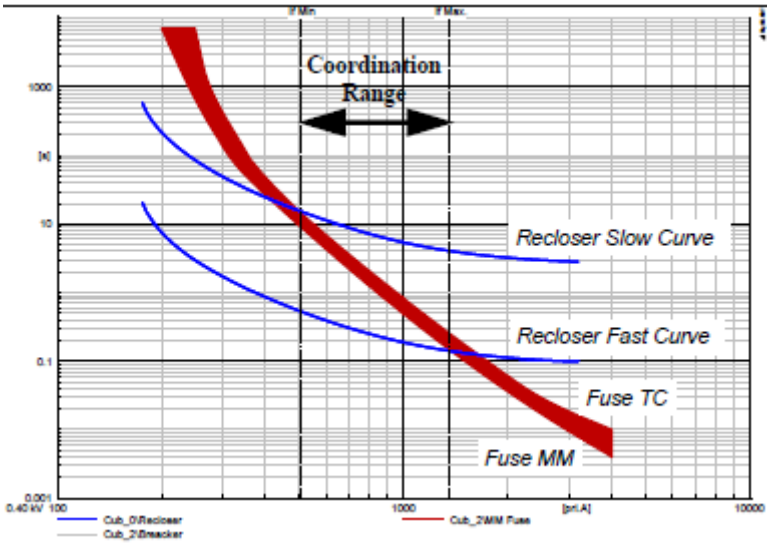


Figure 2-4: Recloser -fuse coordination range [9]

For an auto-reclose protection philosophy, safeguarding life is guaranteed if the recloser locks out in the shortest possible time when closing onto a permanent fault. The dead times of the auto-recloser affect the SAIDI indices. The SAIDI and SAIFI indices determine the average duration and frequency of sustained interruptions experienced by customers. Returning a line to service from a transient fault must happen within 2 minutes (Eskom standard) or 5 minutes (NRS) so that it does not count towards the SAIDI and SAIFI indices. The multi-shot reclosing of an auto-recloser largely affects the Momentary Average Interruption Frequency Index (MAIFI), which concerns the number of momentary interruptions between 3 seconds to 2/5 minutes experienced by customers [6].

The coordination range for the recloser-fuse arrangement is shown in Figure 2-4. Within this range, the recloser will operate first before the fuse is damaged. If the fault is permanent, the fuse will then blow. This ensures that the load does not get disconnected for temporary faults, improves reliability, and deters prolonged power discontinuity due to temporary faults. This is known as the Fuse-Saving strategy; it helps reduce fuse replacement costs and extended power interruptions due to temporary faults that make up 70% to 80% of faults on distribution networks [2], [9], [10]. Recloser-fuse coordination also becomes complex for long feeders with multiple laterals due to fuse selection and coordination. Fuse saving strategy is sometimes not possible on MV distribution systems, for example, a 20 K speed fuse begins melting in less than 80 milliseconds (ms) which is the fastest auto-recloser operating time for currents above 300 A.

The use of fast and slow protection curves applied to fuse saving strategies requires zone sequence coordination which affects the reliability of the auto-reclose function. In most cases, MV faults display instability characteristics during their early stages. The protection will pick-up and reclose several times within the fault current transient stage, timing out, and tripping the auto-recloser on its slow protection curve even though it is the first auto-recloser sequence [6].

2.2.3. Recloser-recloser coordination

For this protection scheme, selectivity is achieved by coordinating the reclosers in series such that the downstream recloser which is electrically close to the fault trips first and isolates the section.

2.2.4. Overcurrent relay coordination

Overcurrent relays provide versatility; a relay allows a wide range of configurations, thus providing coordination for various scenarios [2]. In addition, the remote-control functionality of overcurrent devices makes them ideal for modern power systems. For correct grading, the time delay is selected sequentially to ensure that the overcurrent relay closest to the fault trips first. Typical characteristic curves for relays in a typical overcurrent relay coordination is shown in Figure 2-5. Each relay must have different PSM and TMS to ensure selectivity.

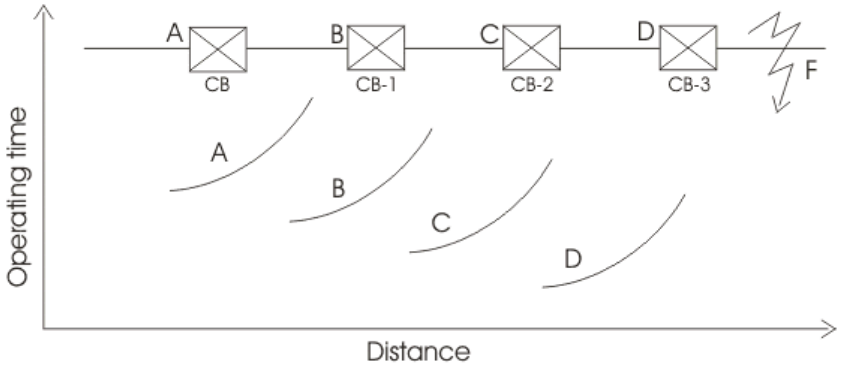


Figure 2-5: Inverse minimum time-delay protection for radial feeder [11]

2.2.5. Overcurrent relay recloser-fuse coordination

This type of coordination involves the use of an overcurrent relay at the beginning of the feeder, reclosers along the feeder and laterals, and fuses for pole mounted MV/LV transformers (less than 2 MVA). The fuses on the distribution transformer are not for the protection of the transformer but to promote continuity of supply on the remaining network and assist in fault finding (rapid identification of transformer winding/surge arrestor that failed) [6]. Hence, fuse saving is not applicable in this coordination method because the faults in these fuse zones are expected to be permanent in nature. Because of the current transformation in the MV/LV distribution transformer, the fuse sizes on the transformer primary side are small and make it impractical to save/coordinate with upstream auto-recloser (ARC). However, the MV fuse must grade with protective devices on the LV side of the transformer and should be capable of detecting phase to neutral and three-phase faults [6]. The following example (Figure 2-6) illustrates the fault current seen by the primary side of the MV/LV transformer because of an LV (secondary side) 3 phase/ phase-neutral fault.

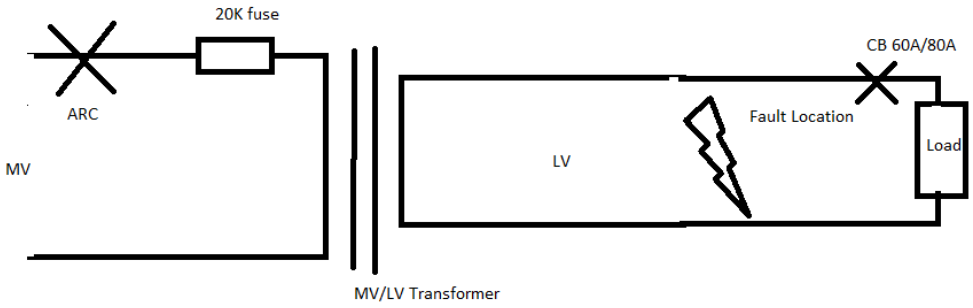


Figure 2-6: Transformer example

- Given a 20 K fuse (minimum operating current of 50 A, at 11 kV and apparent power of 950 kVA),
- An ARC 80ms (ARC for 300 A)
- A 200kVA transformer (11/0.4 kV),
- LV Ph-N fault of 121 A, 3-phase fault of 210 A.

For a 3-phase fault and a Ph-N fault on the transformer secondary side, the primary side fault current are 7.64 A and 4.4 A, respectively. Given the fault current magnitudes on the transformer primary side, the fuse only operates for transformer faults greater than 50 A. The earth fault protection on the feeder circuit breaker is time graded with the series auto-reclosers and the pole transformer fuses for fault current exceeding 100 A on the MV side due to the LV earth electrode resistance value used [6].

2.3. Impact of distributed generation on current protection philosophies

DG provides several benefits to power systems such as ancillary services (reactive power compensation) and reduction of transmission losses (DG connects at distribution level). However, DG can affect the distribution network negatively since current protection systems were designed with the assumption of unidirectional current [12]. Traditional inverse-time protection schemes that are based on overcurrent fail to protect distribution networks with DG when the fault current is almost equal to the pick-up current setting. This is usually the case for faults with high impedance [8]. DG connection may lead to blinding of protection, false and sympathetic tripping, overreach, and unintentional islanding. Therefore, the introduction of DG results in the loss of protection coordination among protective devices. Loss of protection coordination can be cast into three groups:

2.3.1. Loss of coordination

Loss of coordination occurs when secondary relays have shorter operation times than primary relays as a result of an increase in the short circuit level.

2.3.2. Loss of grading

Loss of grading occurs when the operation time difference between primary and secondary relays is greater or less than 300 ms.

2.3.3. De-sensitisation

De-sensitization (slow operating times) occurs when there is a reduction in the fault current level drawn from the mini substation due to DGs contributing to the fault current.

The type of DG plays a role in determining the impact of DG on the protection system. Synchronous based distributed generators have an uncontrolled current during faults as opposed to asynchronous based distributed generator whose fault current depends on the control method rather than the type of generator [13]. Inverter-Based Distributed Generators (IBDG) are decoupled from the grid, the input voltage to the inverter stays the same during a fault. During a fault, the inverter detects the voltage drop on its output terminals, and a new current setpoint is calculated based on the new voltage. Synchronous generators can produce a fault current up to 5-10 times their rated current while inverter-based DG are limited to 1-2 times their rated current by the power electronics [13]. Due to the limited fault current capacity and rapid response of inverter-based DG control, the impact of inverter-based DG has not been explored fully. Although the fault current contribution of inverter-based DG is low, the connection of multiple IBDG to the grid can lead to fault levels that are too low and require high protection sensitivity [13]. In [12], it was shown that inverter-based DG at low penetration can cause miscoordination between the

fuse and recloser in fuse-recloser coordination (fuse saving). However, the impact of inverter-based DG on protection philosophies other than fuse-recloser has not been explored.

2.4. Future/proposed LV and MV network protection philosophies

As discussed in the previous section, the increase in the number of DG within distribution networks has an impact on the coordination, sensitivity, and selectivity of protection schemes. Therefore, the previous assumptions, methods, protection, and control strategies need to be reconsidered when designing a protection system with DG. The new protection design methodology must ensure that the possible multiple current paths that feed a fault because of DG do not jeopardise the operation of the protection system [2].

The South African distribution code states that it is the responsibility of the utility to develop a protection guide for connecting DG to the distribution system for safety and operation reliability [7]. DG units are supposed to disconnect upon the detection of unintentional islanding. In MV networks, the neutral earth point of DGs are connected to the utility neutral point to avoid issues of earth fault desensitization and circulating zero-sequence or triplen harmonic currents between distant earth connections [7]. An isolation transformer is required for the connection of a DG unit to the utility MV network [7].

Several network protection philosophies and methods have been used to mitigate the impact of DG. In [12], the proposed methods are classified into 5 categories:

a) Limiting the DG capacity to the maximum capacity before the protection system is jeopardised.
b) Modifying the current protection system by introducing extra breakers, directional relays, distance relays, and reconfiguring the distribution network. At Eskom, the typical requirements of a DG unit depend on the connection type i.e. connection to the utility busbar or radial line tee-in. The requirements include direct transfer trip, directional protection, auto reclose dead time, SEF definite time characteristics, overcurrent and earth fault time characteristics, and anti-islanding.
c) The use of adaptive protection system relays capable of communicating with other relays, accessing remote measurements, and dynamically changing their settings using processing units.
d) Lastly, the use of fault current limiters.

The modification of the current protection system is costly. It can be argued that limiting the DG capacity is not a desirable solution since it limits the growth of DG connections. However, knowing the DG penetration limit allows power utilities to accept DG connections up to the specified limit without the need to perform major load flow studies. These load flow studies often delay the process of DG connection approval and in turn limit the growth of DG connections in distribution networks.

2.5. Network planning for DG

System planners worldwide are faced with a growing challenge of integrating an increasing number of Renewable Energy Source (RES) DG on the MV/LV network, expansion of HV networks to connect large scale RES generators, and network stability due to the intermittent and variable nature of RES generators [14].

Before the connection of a DG unit, the fault level contributions of the DG plant need to be considered. The owner of the DG unit is required to provide all the relevant information required to model the generator and its contribution to fault current over time. Sometimes, the fault current contribution from the DG is a small fraction of that supplied by the grid and decays rapidly with time. Therefore, the current protection coordination applied to radial MV networks is not affected by DG [7]. Planning studies and impact assessment of the connection of large DG units can be performed for planned sites, however, these studies take time to perform and are only practical for large DG units. In addition, the actual DG plans can only be finalised once preferred bidders are announced (illustrated in Figure 2-7) [15]. The above-mentioned problems buffer the process of DG connection approval by the utility. Even though the planning studies for small DG units can be performed, the date when the customer installations are connected will be unplanned, frequent, and random [14]. This brings about the need for an adaptive and system-wide DG planning approach together with the need to calculate the threshold of DG penetration before protection coordination changes are applicable. Figure 2-7 emphasises the uncertainty of planning for large DGs by showing the typical differences between project timelines for a transmission line and a Renewable Energy Independent Power Producer (REIPP) project.

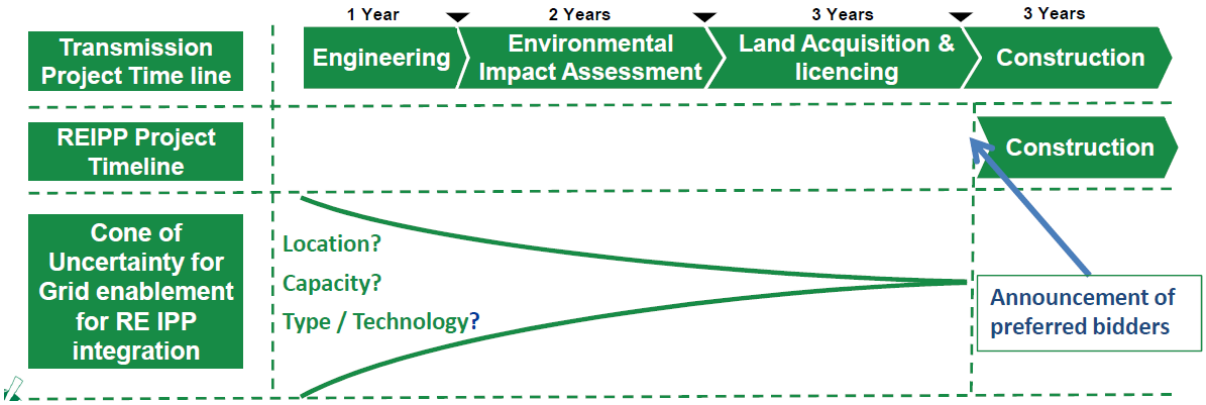


Figure 2-7: Typical project timelines for a transmission line and renewable energy independent power producer [15]

As illustrated in Figure 2-7, planning decisions for DG installations are affected by the capacity, type, and locations of the DG. A DG planning study could focus on one Active Network Management (ANM) scheme where DG capacity is increased, and the optimal DG limit is found [14]. However, the study is only valid if the exact DG limit is connected, which is highly unlikely. In [14], it is argued that most DG penetration studies within the literature have an underlying assumption that DG penetration will increase and distribute within the network accordingly until the limit is reached, and DG planning is based on a one-time "plan and forget" approach. A planning approach that takes into consideration the randomness of numbers, size, location, and connection times was proposed. This was done by adaptively changing customer DG installation scenarios while increasing DG penetration.

The control of future smart grids should consist of ANM schemes that control the utilisation of network assets to maximise DG (including RES) integration. The controlled network assets include on-load tap changers, capacitor banks, reactive power compensators, breakers, battery storage systems, inverters, electric vehicles, demand-side management systems, power factor control, and DG active power management systems. Hence the need for a network DG planning approach which maximises DG capacity, while minimizing integration costs, operation costs, network losses, and improving the voltage profile [14].

2.6. Tools/algorithms used to establish DG hosting capacity on distribution networks

The impact of DG units on protection varies with their size, type, and location; as a result, it is difficult to assess their impact. This brings about challenges when planning or establishing the hosting capacity of a distribution network. Hosting capacity is the amount of DG that can be connected at any location on any feeder on the network without any issues [16]. The literature on penetration limit calculation varies among researchers. In some studies, the penetration limit is calculated in relation to the capacity of the transformer, or as a ratio of installed DG to the feeder maximum/minimum load and based on voltage variations. The inclusion of the relationship between the actual load and actual DG provides a more comprehensive approach to penetration limit calculation [17]. Figure 2-8 shows the relationship and contribution of Renewable Energy Sources (RES) (in the form of hourly average capacity factor) to the South African load profile.

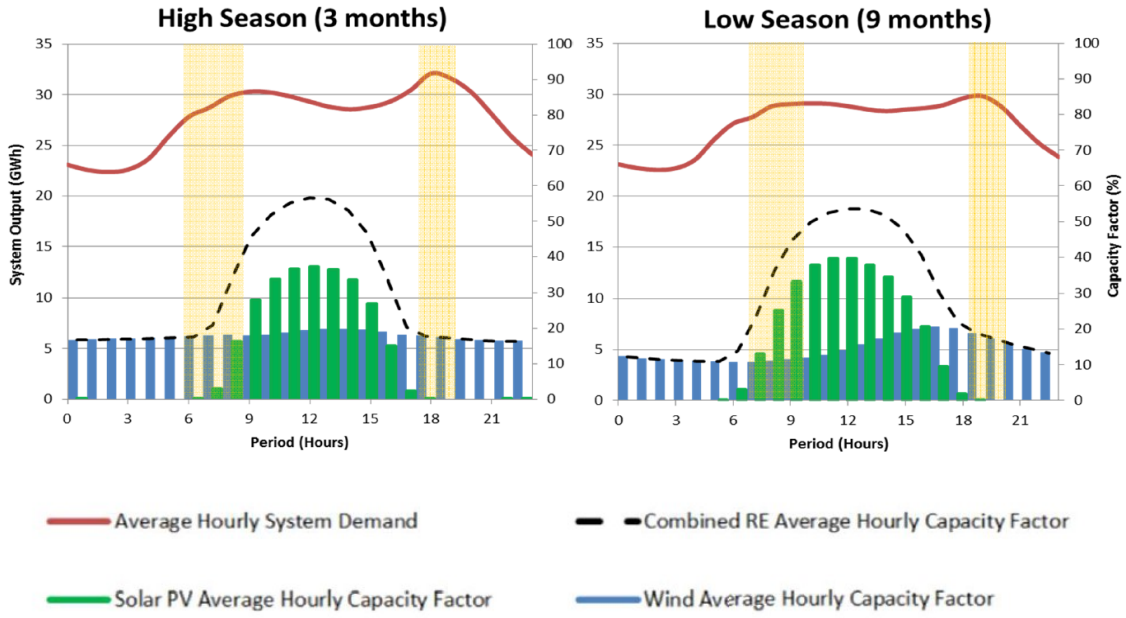


Figure 2-8: Average hourly capacity factor with the time-of-day energy production and load profile [18]

Based on the system demand (illustrated in Figure 2-8), wind contributes more to peak periods than solar, therefore the impact of wind DG on the distribution network is more than that from solar DG.

The factors to consider when calculating the DG penetration limit include, inter alia:

a) Thermal limits
b) Power quality and voltage
c) Protection coordination
d) Reliability and safety

For each of the limiting factors, the DG technology, size, location, and type (Distributed or Centralized) need to be considered. Therefore, hosting capacity is location-specific, feeder dependent, and time-varying.

In [19] and [20], an optimization function was used to calculate the DG penetration limits. The relationship between power from DG and fault currents was found using the sub-transient

reactance of the DG and the short circuit capacity equation. The operating ranges of protective devices (recloser-fuse coordination) were used as constraints to the optimization function to ensure that the fault current lies within the operating range of the protective devices. This method is time-consuming and limited to a few feeders since it relies on conventional power flow equations employing Jacobian matrices.

In [21], the Herman Beta algorithm was modified for voltage calculation in active LV feeders using the negative load approach for DG models to analyse voltage rise constraints. The method was then validated using Monte Carlo Simulations. This method focuses on the voltage rise being the sole determining factor for DG penetration limit. In [22], protection based DG limits were defined and established for DG connected according to the NRS 097-2-3 standard. The results are location-dependent; feeder specific and the protection-based DG penetration limits were only based on overcurrent relay coordination.

The Electric Power Research Institute (EPRI) system-wide DG assessment method is an efficient and accurate method to determine the DG penetration limits on a network at specific locations. It is a streamlined analysis method that is based on learning from a detailed analysis method. A detailed penetration limit analysis involves the use of commercial software (DIgSILENT) to repeatedly perform the load flow, investigate DG scenarios, and observe DG impacts (on the voltage profile and protection coordination). The EPRI streamlined hosting capacity analysis uses a separate Model-Based Interface (MBI) module to add the DG scenarios, to perform the impact assessment, and produce standard hosting capacity results [16]. Although the EPRI drive tool is not a replacement for detailed studies, it gives close approximations of DG impact in less time and is easily replicable across the whole system [23]. Table 2-1 shows the criteria used in the EPRI drive tool for determining the distribution network's hosting capacity.

Table 2-1: The criteria and basis used on the EPRI drive tool in determining the hosting capacity [23]

Category	**Criteria**	**Basis**	**Hosting Capacity Threshold**
Voltage	Primary over-voltage	Feeder voltage limit	Vpu voltage magnitude
	Primary under-voltage	Feeder voltage limit	Vpu voltage magnitude
	Primary voltage deviation	Change in voltage from no DG to full DG	Percentage voltage change
	Regulator voltage deviation	Change in voltage from no DG to full DG at a regulated node	Half bandwidth
Loading	Thermal limit for charging (loads)	Remaining conductor current carrying capacity at peak loading	Percentage of normal current rating
	Thermal limit for discharging (Generation)	Conductor current rating together with minimum loading condition	Percentage of normal current rating
Protection	Additional conductor fault current	Deviation in feeder fault currents	Percentage of fault current increase
	Sympathetic breaker relay tripping	Breaker zero sequence current	Current magnitude
	Breaker relay reduction of reach	Deviation in breaker fault current	Percentage of current decrease
	Reverse power flow	Conductor minimum loading condition	Percentage of current during minimum loading condition
	Unintentional Islanding	Conductor minimum loading condition	Percentage of current during minimum loading condition

The EPRI Drive tool has the following underlying assumptions

- Small and Large DGs are considered as a constant current injection, no change in DG current during induced voltage rise.
- Constant load magnitude during DG voltage change (loads are based on initial load flow).
- Three-phase large DGs are interconnected with transformers that do not pass zero sequence currents (thus not creating a ground source), otherwise all protection hosting capacities are low.
- Fault currents are not based on sub-transient reactance [23].

2.7. Applications of machine learning in power systems

In [24] a machine learning classification technique (support vector machine (SVM)) is used to determine whether a single-phase grid-connected PV system is operating in islanding mode or not. During simulations, islanding was detected with 100% accuracy within a time delay of 20 ms. This method uses the voltage at the point of common coupling (PCC) to obtain all features associated with the simulated fault conditions of a grid-connected PV. Therefore, more computational power is needed as the penetration of DGs increase (increased number of PCC). A PV output power forecasting model was developed in [25] based on the extreme learning machine (ELM) approach. The root mean square error of the PV output power forecasting ELM model ranged between 54.96-90.41% during model validation, with average computational times of 15.3 ms. The power output in this study is from three PV systems in one location and the model cannot work with input data externally originated [25]. A machine learning technique architectural framework for accurate detection and classification of power quality disturbances is defined in [26] as shown in Figure 2-9.

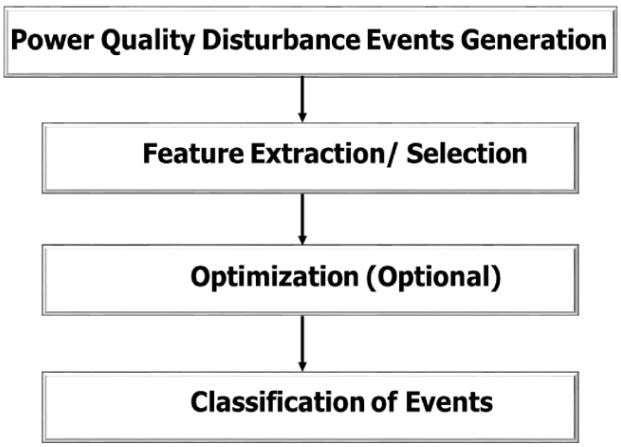


Figure 2-9: Power quality disturbances classification steps [26]

In this research, the aim was to apply an architectural framework similar to Figure 2-9, select a model (linear regression model to calculate relay operating times and logistic regression model to predict protection miscoordination groups) that generalises well and fits the test data set well. To achieve the objective, the following steps recommended in [27] were followed.

- After the implementation of the model, plot learning curves to determine what steps to take to improve the learning.
- Perform error analysis: select the examples in the cross-validation set where the algorithm made errors and spot any systemic trends [27].

2.8. Research approach

The problem that exists with DG is the vast difference in the technology, location, size, connection sequence, and protection scheme requirements which result in future DG network planning inadequacies – the **Network DG Planning Dilemma**. Machine learning techniques provides an opportunity to automatically build computational models of complex relationships (real-world problems that cannot be modelled directly as a closed-form input-output relationship) through the processing of data available and maximising a problem-dependent performance criterion [33]. In this research, the approach to determining protection-based DG penetration limits will be based on supervised machine learning methods. The aim is to rely on protection features present in the distribution network data i.e., fault level, Inverse Definite Minimum Time (IDMT) curve, pick-up current setting, Time Multiplier Settings (TMS), calculated relay operating times, and relay positions to observe how the network responds at certain DG penetration levels (actual relay operating times). Data set representing carefully anonymized distribution networks with accepted protection philosophy applied will be used. A supervised machine learning algorithm will be applied after nontrivial data pre-processing through recommendation systems and shuffling. The planning dilemma will be cast into three parts: the first, an automated pattern classification (logistic regression for classification of protection miscoordination), the second, regression (predicting operating time after different levels of DG penetration), the last recommendation system (where, when and how much PV will be connected).

3. Methodology, MV distribution network case files and PV data

This chapter presents the method used in this research project, introduces the MV distribution network case files, PV data, and global horizontal irradiation map of South Africa.

3.1. Method

To achieve the objectives stated in section 1.2, the following method is formulated;

1. Collect MV feeder case files from Eskom's network.
2. Incorporate Eskom MV protection philosophy to the case files and on the illustration network.
3. Reduce the network, model the reduced network diagram on a table, export data to CSV, import to Jupyter Notebooks (Anaconda3) and Octave (clean, process, and shuffle the data).
4. Use recommendation systems on the PV, installation, and postal code data (Figure 3-1).
5. Define the DG hosting capacity calculation constraints and considerations -classes and success criteria.
6. Use results from a detailed analysis study on the impact of DG on MV feeders as training data for both linear and logistic regression (Figure 3-2).
7. Training: identify the learning rate, regularization parameter (lambda), polynomial degree and theta (θ) that give an accurate model.
8. Use the best regression model (most accurate/well trained) on the four MV distribution networks to calculate relay or recloser operating times before and after three stages of DG penetration (35%, 65%, and 75% of ADMD).
9. Use the results from the best performing regression model applied to the four MV distribution networks and their parameters as inputs to the most accurate logistic regression model to classify which relay or recloser protection features results in which protection coordination problem (after three stages of DG penetration).
10. Identify the type of feeder with the highest DG penetration limit, the worst-performing feeder, and the accuracy of predictions through comparison (vertical analysis).
11. Study the feeder design and characteristics, recommend the design for DG network planning standards.
12. Identify the feeder's protection parameters (feeder with highest DG penetration limit) and create an MVON for DG (Figure 3-3).

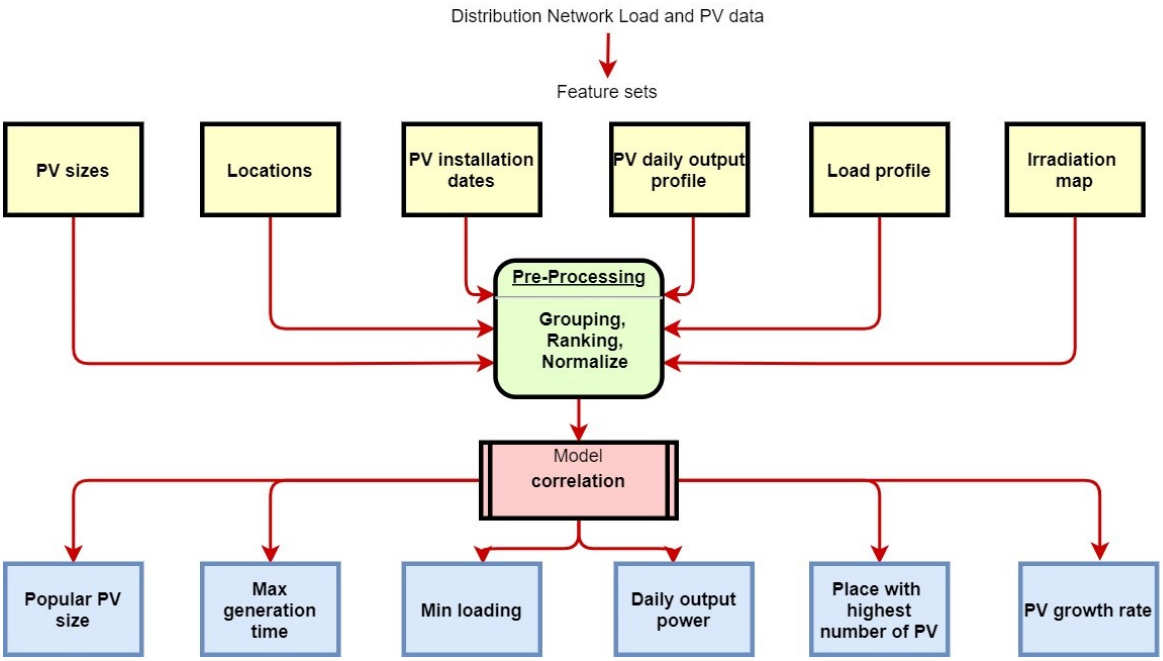


Figure 3-1: Model workflow (Step 3 to 4)

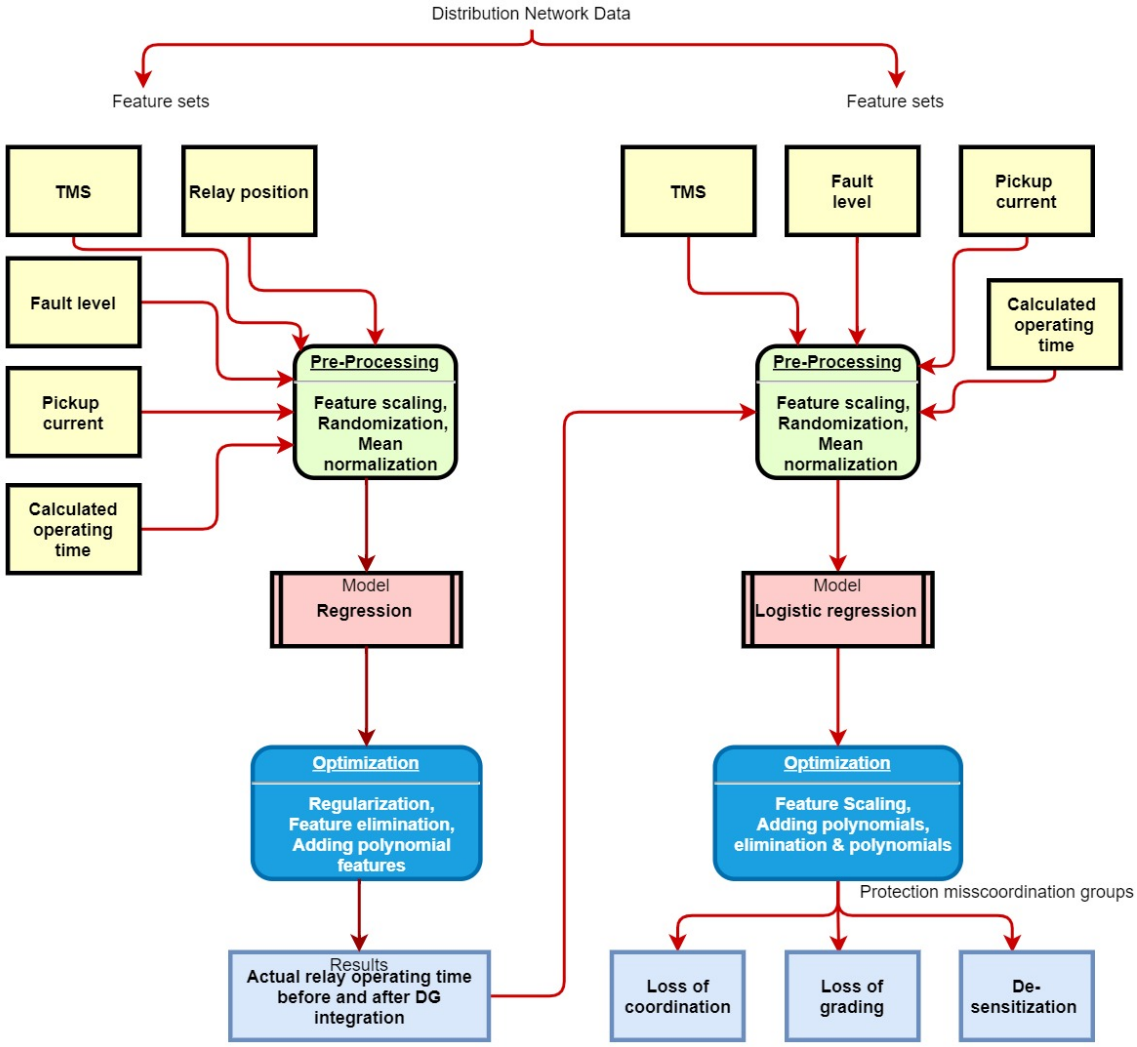


Figure 3-2: Model workflow (Step 6 and 9)

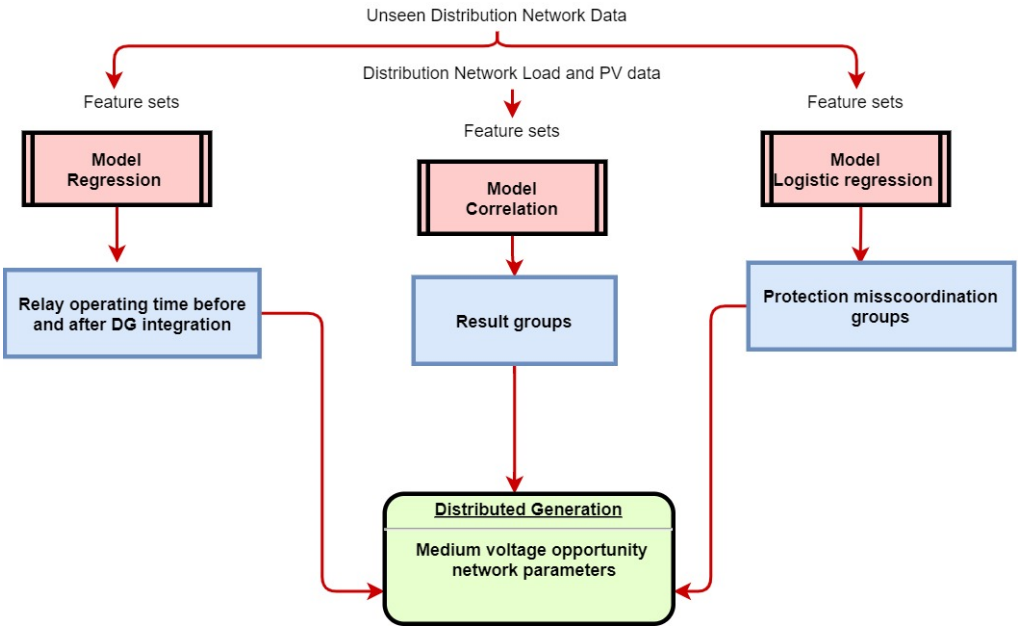


Figure 3-3: Medium voltage opportunity network model (Step 10 and 12)

3.2. MV distribution network case files and PV data

The entire research project work includes three MV distribution network case files, each case file consists of actual half-hourly load data for one year (2015), three-phase and single-phase fault levels and DIgSILENT files with real network parameters. Each distribution network falls into one of the distribution system classifications i.e. radial, ring, and parallel. The DG data consists of; postal code, city, PV name, kW rating, date of installation, and actual power generation of a 5.3 kW PV system recorded in 15-minute intervals over a year (2015) [28]. Lastly, results from a distribution network where detailed studies on the impact of DG penetration on the protection system was done [22]. As per guidelines on Eskom data sensitization, the MV distribution network names from which the data was obtained were labelled A, B, and C. Table 3-1 to Table 3-4 show sample data points from each case file data set.

Table 3-1: Sample 5.3 kW PV output data

Date	Time	Output Power (kW)
14/08/2015	6:45 AM -7:00 AM	0.01 -0.08
14/08/2015	7:00 AM -7:15 AM	0.08 -0.15
14/08/2015	7:15 AM -7:30 AM	0.15 -0.43
14/08/2015	7:30 AM -7:45 AM	0.43 -0.71
14/08/2015	7:45 AM -8:00 AM	0.71 -0.99
14/08/2015	8:00 AM -8:15 AM	0.99 -1.25
14/08/2015	8:15 AM -8:30 AM	1.25 -0.86
14/08/2015	8:30 AM -8:45 AM	0.86 -1.01
14/08/2015	8:45 AM -9:00 AM	1.01 -1.77
14/08/2015	9:00 AM -9:15 AM	1.77 -2.45
14/08/2015	9:15 AM -9:30 AM	2.45 -3.09

Table 3-2: Sample load data (30 min intervals)

Date and Time	Power (kVA)
8/14/2015 5:00	388.02
8/14/2015 5:30	400.70
8/14/2015 6:00	416.08
8/14/2015 6:30	429.73
8/14/2015 7:00	436.55
8/14/2015 7:30	461.50
8/14/2015 8:00	481.76
8/14/2015 8:30	484.17
8/14/2015 9:00	487.93
8/14/2015 9:30	546.13
8/14/2015 10:00	545.31
8/14/2015 10:30	525.45
8/14/2015 11:00	501.22
8/14/2015 11:30	479.87

Table 3-3 shows a sample of the range of postal codes in South Africa (graphically represented in Figure 3-4). Figure 3-4 and Figure 3-5 present a picture of the location of PV systems in relation to the global horizontal irradiation map of South Africa. Table 3-4 presents a sample data of PV systems publicly available on Sunny Portal [28].

Table 3-3: Sample region and postal code range data [29]

Zip Range		Region
1	299	Gauteng—Pretoria/Tshwane
1000	1399	Mpumalanga—northern half
1400	1699	Gauteng—East Rand / Ekurhuleni Metro
1700	1799	Gauteng—West Rand, Mogale City/Krugersdorp, Roodepoort (now part of Johannesburg)
1800	1999	Gauteng—Soweto and Vereeniging/Vanderbijlpark Region (Motsweding)
2000	2199	Gauteng—Johannesburg (original Johannesburg, Randburg, Sandton)
2200	2499	Mpumalanga—southern half
2500	2899	Northwest Province—southern and central
6500	6699	Western Cape—Garden Route and Oudtshoorn area
6700	6899	Western Cape—Klein Karoo
6900	7099	Western Cape—Great Karoo
7100	7299	Western Cape—Area south-east of Cape Town
7300	7399	Western Cape—West Coast
7400	7599	Western Cape—Northern parts of Cape Metropole
7600	7699	Western Cape—Areas East of Cape Town, such as Stellenbosch
7700	8099	Western Cape—Cape Town and Cape Peninsula
8100	8299	Northern Cape—Namaqualand Region
8300	8799	Northern Cape—Eastern Part
8800	8999	Northern Cape—Gordonia Region
9000	9299	Formerly assigned to South West Africa
9300	9399	Free State—Bloemfontein and surrounds
9400	9699	Free State—Northern Free State
9700	9899	Free State—Eastern Free State
9900	9999	Free State—Southern Free State

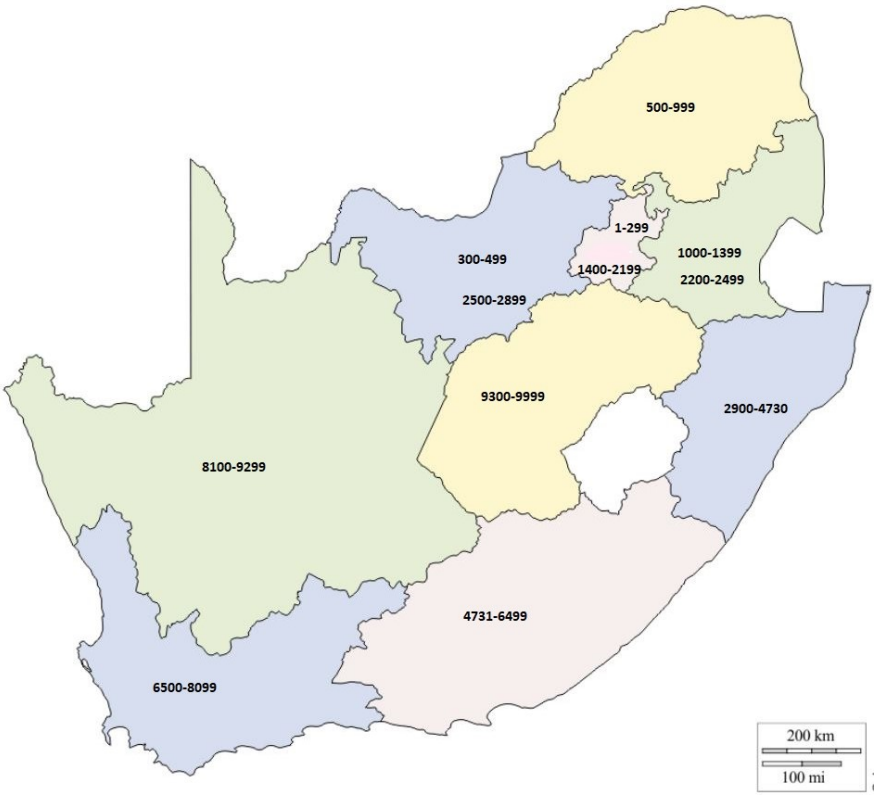


Figure 3-4: Graphical representation of province and postal codes [29] [30]

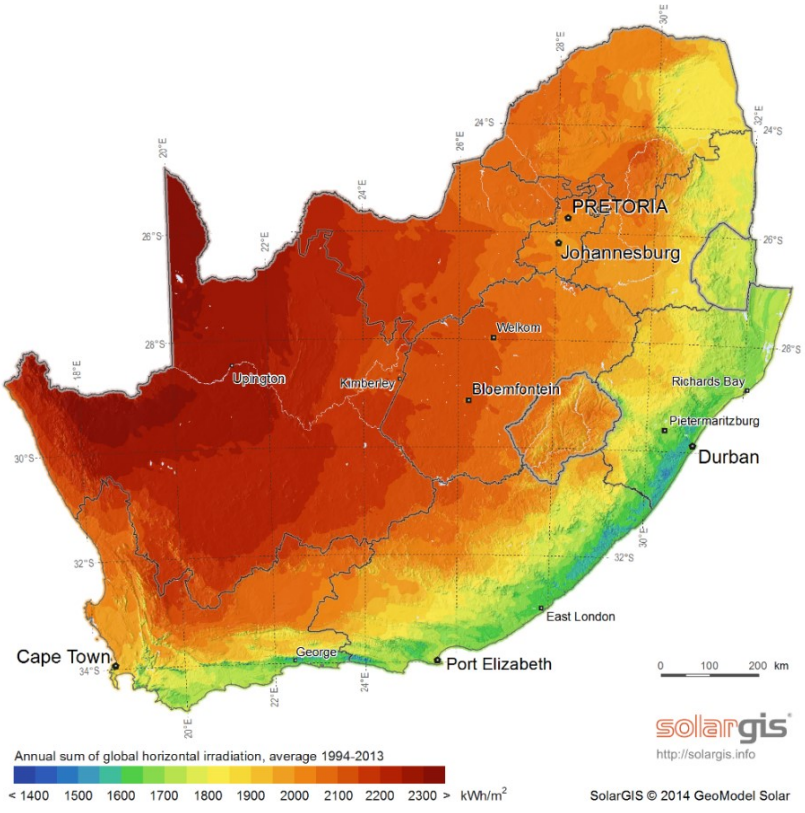


Figure 3-5: Global horizontal irradiation map of South Africa [31]

Table 3-4: PV systems sample data publicly available on Sunny Portal [28]

PV System Name	Country	Zip Code/ Postal code	City	Power (kWp)	Date of installation
Weenen	South Africa	9460	Wesselsbron	25.75	12/04/2015
H9	South Africa	9430	Virginia	5.30	07/03/2018
G. Scholtz 47 Koedoeweg, Woodl	South Africa	9301	Bloemfontein	4.95	22/01/2019
Mr JPC Antunes	South Africa	9301	Bloemfontein	5.60	10/01/2018
FS Botanical Garden	South Africa	9301	Bloemfontein	52.92	
Bram Fischer	South Africa	9300	Bloemfontein	750.00	
Installer	South Africa	8870	Keimoes	20.55	27/07/2017
Kakamas Abattoir	South Africa	8870	Kakamas	102.60	13/10/2014
Upington Toyota	South Africa	8800	Upington	82.10	
Stellaland Begrafnisondernemer	South Africa	8600	Vryburg	10.25	27/11/2017
Mediclinic Gariep Kimberley	South Africa	8301	Kimberley	130.02	
MastetreadsKim	South Africa	8300	Kimberley	122.44	23/05/2017
Hantam Botanical Garden	South Africa	8180	Nieuwoudeville	26.46	28/03/2017
25 GRENSSTREET, VREDENDAL	South Africa	8160	Vredendal	3.00	
Buitelaan	South Africa	8160	Vredendal	4.80	
BUILD IT, VRDENDAL	South Africa	8160	VREDENDAL	22.31	
Aggenbagskraal	South Africa	8135	South Africa	15.95	05/05/2018
House Moser - Soventix ZA	South Africa	8005	Camps Bay	4.94	17/10/2013
Two Oceans Aquarium Solar	South Africa	8001	VandA Waterfront	6.96	
Cape Town DSK - SMA	South Africa	8001	Cape Town	150.00	16/10/2016
Marc Gore	South Africa	8000	Cape Town	5.00	08/07/2015
NF 006 Scholtz - EM	South Africa	8000	Cape Town	6.10	24/05/2016
Red Cross Children's Hospital	South Africa	8000	Cape Town	10.40	
BP Waterfront Office Cape Town	South Africa	8000	Cape Town	67.90	

The complete data points for Table 3-1, Table 3-2, Table 3-3 and, Table 3-4 are in appendix C.

4. MV protection philosophy and the distribution network model

In this chapter, Eskom's MV protection philosophy and the steps of applying the philosophy to a distribution network including network reduction and modelling are comprehensively described. This chapter also introduces the models and parameters of four real distribution networks used in this study. The MV protection philosophy is then applied to these network models and the training data set is presented.

4.1. Reduced distribution network diagram

Consider the 11kV feeder system (Figure 4-1) from the Eskom Distribution network.

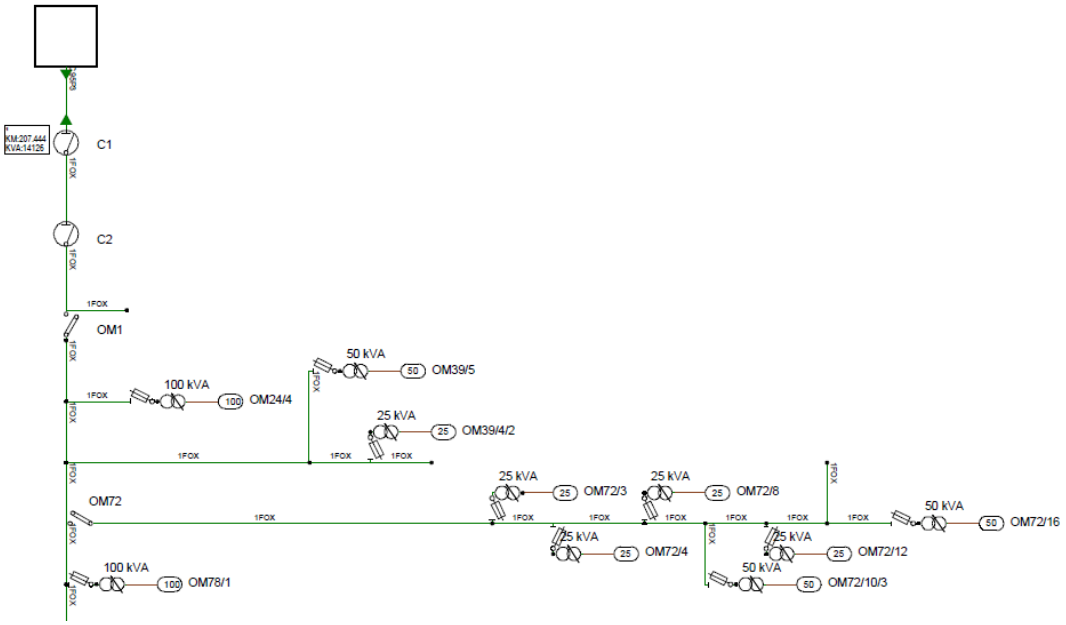


Figure 4-1: Sheet 1 of 19 of an 11kV distribution system from the Eskom MV network

For Figure 4-1 the following data are known:

Table 4-1: Transformer and MV CB parameters

Transformer and MV busbar Circuit Breaker parameters	
Capacity	10 MVA
HV rating	132 kV
MV rating	11 kV
Overcurrent CT ratio	200:1
Earth fault CT ratio	100:1
Transformer circuit breaker operating time	1.50 s
Transformer earth fault pick-up current setting	100 A

The application of the protection settings philosophy to the MV system requires the network diagram to be reduced to a network diagram that only includes specific information such as [6];

- Relay locations, make and types, auto-reclosers, and sectionalisers installed on the feeder.
- Locations of the furthest reach points in every protective zone.

- Conductor types including conductor change points and normally open points.
- Notified maximum demands of bulk supply points, IPPs, and downstream substations.

The network in Figure 4-1can be reduced to a network diagram with five t-offs/branches illustrated in Figure 4-2. It is as assumed that the distance between the nodes/poles is 100 m. Therefore, node 219/132/1 (219*100 m +132*100 m +1*100 m =35200m) is 35.2 km away from the MV CB.

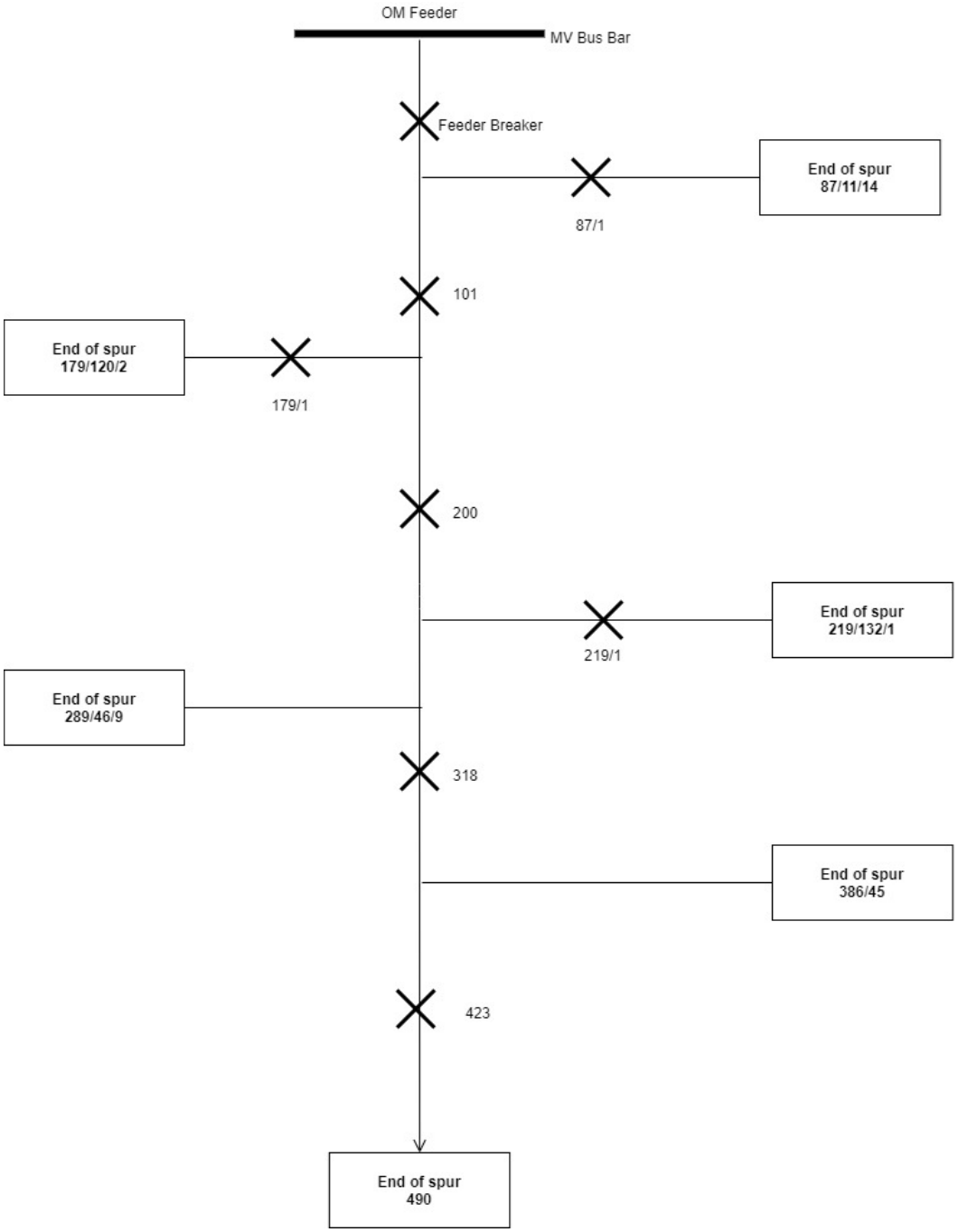


Figure 4-2: Reduced network diagram (Figure 4-1)

4.2. Relay/auto-recloser operating times

Two standard curves are used to achieve fast operating times at higher currents i.e.

- IEC normal inverse curve (Table 4-2 equation 4-1): preferred for applications with long MV feeders because the rise in operating time is slower as the fault current drops down the feeder.
- IEC extremely inverse curve (Table 4-2 equation 4-2): preferred curve for low source impedance (high fault current), relatively short feeders. Provides fast operation at high currents and has a shape that closely matches fuses (useful in grading with transformer fuses) [6].

Table 4-2: IEC60255 Inverse Definite Minimum Time (IDMT)curves

Standard characteristic curves	Formula	
Normal inverse	$t = 0.14 \times \frac{TMS}{((\frac{I_{Fault}}{I_{PU}})^{0.02} - 1)}$	(4-1)
Extremely inverse	$t = 80 \times \frac{TMS}{((\frac{I_{Fault}}{I_{PU}})^{2} - 1)}$	(4-2)

Where,

TMS: time multiplier setting
I_{Fault}: measured fault current magnitude.
I_{PU}: pick-up current setting
t: operating time

The time multiplier is chosen such that coordination with downstream devices (MV side fuses on MV/LV transformers) is optimised. Also, the protection should operate slower than the total clearing time of the fuses by an appropriate grading margin [6].

4.3. Pick-up current setting selection philosophy

The pick-up current settings were chosen based on the following philosophy:

I. The pick-up current setting is set low enough to detect a fault drawing 80% of the calculated phase-to-phase fault current at the furthest downstream reach of the protection device, under minimum network conditions. For this calculation, the reach of the protective device should be set in such a way that includes coverage of the entire section of the network which, under normal network conditions is protected by the downstream protective device. This means that the protective device should be able to detect a phase-to-phase fault in its zone of protection, together with the zones of protection provided by downstream protective devices. For example, in Figure 4-2, the auto-recloser at 200 must trip for a fault at 219/132/1 (furthest point) for a condition where the auto-recloser at 219/1 fails to operate.
II. The pick-up current setting is set up such that it is between 110% and 120% of the probabilistic operating current of the thinnest/thickest conductor within the overcurrent reach.

III. The pick-up current setting is set to be at least 110% of the downstream device's pick-up current setting, and 90% of the pick-up current setting of the upstream device.
IV. For earth fault, the pick-up current setting is allocated incrementally up to but not exceeding the medium voltage busbar CB earth fault pick-up current setting.
V. A grading margin of at least 200 ms is required to maintain selectivity between the relay/ auto-recloser operating times.

4.4. Application of Eskom's MV protection philosophy

MV feeders within the Eskom environment are mostly operated radially, with resistive neutral earthing (Neutral Electro-Magnetic Coupler/neutral earthing Compensator (NEC), a Neutral Earthing Resistor, and an auxiliary Transformer (NECRT)) to limit earth fault currents to 360 A per source transformer. MV feeders are most often protected by non-directional phase over current, earth fault and sensitive earth fault (SEF) protection [6]. A typical NECRT setup is shown in Figure 4-3.

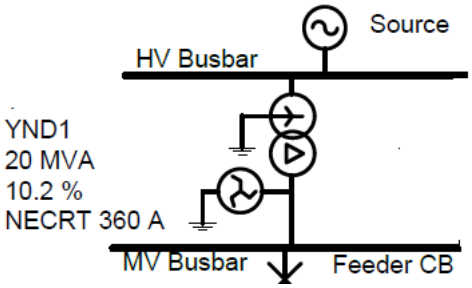


Figure 4-3: Typical NECRT set up within the Eskom environment

The aim of the NECRT as explained above is to provide a neutral point in a delta of a distribution network and limit the earth fault current to a designated level under fault or abnormal conditions [32].

The fault current on a transformer is limited to 150% of the normal current. Therefore, for the transformer in Table 4-1, the pick-up current setting of the transformer circuit breaker is calculated as follows:

$$I_F = 150\% \times \frac{S}{\sqrt{3} \times V_{LL}} = 150\% \times \frac{10\ MVA}{\sqrt{3} \times 11kV} = 787.29A \qquad (4\text{-}3)$$

Thus, the first relay/recloser on the feeder will have 787.29 A as the upstream pick-up current setting when the pick-up setting considerations in section 4.3 are applied.

Fault levels at each node for the network based on conductor length/equipment rating illustrated in Figure 4-2 are indicated in Table 4-3.

Table 4-3: Fault levels at each node

Fault levels at each node		
Node	**Three-phase fault level (A)**	**Single-phase fault level (A)**
490	132	102
423	231	118
386/45	149	116
318	306	156
289/46/9	186	144
219/132/1	182	141
219/1	437	223
200	479	245
179/120/2	213	165
179/1	529	271
101	909	469
87/11/14	552	426
87/1	1031	533
Feeder CB	7706	5379

The fault levels at each node are based on the 11 kV three-wire fox conductor system i.e. conducting area mm^2 Cu equivalent, pu R, X, B values, and conductor length. Fox conductor has a normal current of 148 A and an emergency rating of 192 A (Table D-2).

The pick-up current setting considerations in section 4.3 can be represented in the following way:

1) $80\% \times \frac{\sqrt{3}}{2} \times FL_{FP}$ (FL_{FP} fault level at the furthest point)
2) Conductor Emergency current rating $\times$ 1.1
3) $90\% \times$ upstream pick $-$ up current setting,

Where the lowest pick-up current setting is used.

Applying these considerations to Figure 4-2 and Table 4-3, the pick-up current settings for each node are shown in Table 4-4.

Table 4-4: Three-phase and earth fault pick-up current settings

Fault levels at each node		
Node	**3 phase fault pick-up current setting (A)**	**Earth fault pick-up current setting (A)**
423	124.704	20
318	138.56	35
219/1	141.54	20
200	157.27	50
179/1	176.46	30
101	196.1	65
87/1	269.28	60
Feeder CB	299.20	80

For earth fault protection, pick-up current settings within the range 20 A – 40 A allow the use of higher LV earth electrode resistances providing more sensitivity and faster isolation of MV phase to LV neutral fault conditions [6].
The earth fault protection pick-up current settings are chosen based on the following philosophy:

I. 30 A primary earth fault pick-up current setting at the furthest downstream settable protective device and upstream pick-up current settings are at least 5 A higher than downstream devices.
II. IEC normal inverse time-current curves are used. Extremely inverse curves are often not supported by every device in the MV network.
III. Fault protection is graded with downstream auto-reclosers such that it coordinates with the total clearing time of fuses on pole-top transformers for all earth fault currents exceeding 100 A primary [6].

This philosophy is illustrated in Table 4-4.
For the furthest point on the backbone, the first auto-recloser is given a TMS of 0.01 and a TMS of 0.05 is assigned for the furthest point on the T-offs/branch to start the grading calculations. A minimum grading margin of 200 ms is used between the protection devices. Table 4-5 and Table 4-6 illustrate the application of the above-mentioned points (Eskom's MV protection philosophy) to the network in Figure 4-2.

Table 4-5: Three-phase fault operating times and TMS values for the protective devices on the network

Backbone				
Node	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
423	0.01	231	124.70	0.11
318	0.03	231	138.56	0.41
318	0.03	306	138.56	0.27
200	0.09	306	157.27	0.57
200	0.09	479	157.27	0.56
101	0.11	479	196.1	0.86
101	0.11	909	196.1	0.50
CB	0.13	909	299.2	0.80
CB	0.13	7706	299.2	0.27
Trf	0.38	642	65.61	1.13
Trf	0.38	369	65.61	1.50
T-offs/branch 3				
Node	**TMS**	**Fault level (A)**	**Pick-up current Setting (A)**	**Operating time (s)**
219/1	0.05	437	141.54	0.31
200	0.09	437	157.27	0.61
200	0.09	479	157.27	0.56
T-offs/ branch 2				
Node	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
179/1	0.05	529	176.46	0.32
101	0.11	529	196.1	0.62
101	0.11	909	196.1	0.50

T-offs/ branch 1				
Node	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
87/1	0.05	1031	269.28	0.26
CB	0.13	1031	299.2	0.56
CB	0.13	7706	299.2	0.27

Table 4-6: Earth fault operating times and TMS values for the protective devices on the network

Backbone				
Node	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
423	0.05	102	20	0.21
318	0.08	102	35	0.51
318	0.08	156	35	0.36
200	0.11	156	50	0.66
200	0.11	245	50	0.47
101	0.15	245	65	0.77
101	0.15	469	65	0.52
CB	0.21	469	80	0.82
CB	0.21	5379	80	0.33
T-offs 3				
node	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
219/1	0.05	223	20	0.14
200	0.11	223	50	0.44
200	0.11	245	50	0.47
T-offs 2				
Relay/Recloser	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
179/1	0.05	271	30	0.16
101	0.15	271	65	0.46
101	0.15	469	65	0.52
T-offs 1				
Relay/Recloser	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Operating time (s)**
87/1	0.05	533	60	0.16
CB	0.21	533	80	0.46
CB	0.21	5379	80	0.33

The TMS and operating times in Table 4-5 and Table 4-6 for the different nodes are shown on the reduced network diagram in Figure 4-4 and Figure 4-5. The directional arrows represent the Time Delay (TD)/operating time of the secondary relay for a fault in the primary relay zone.

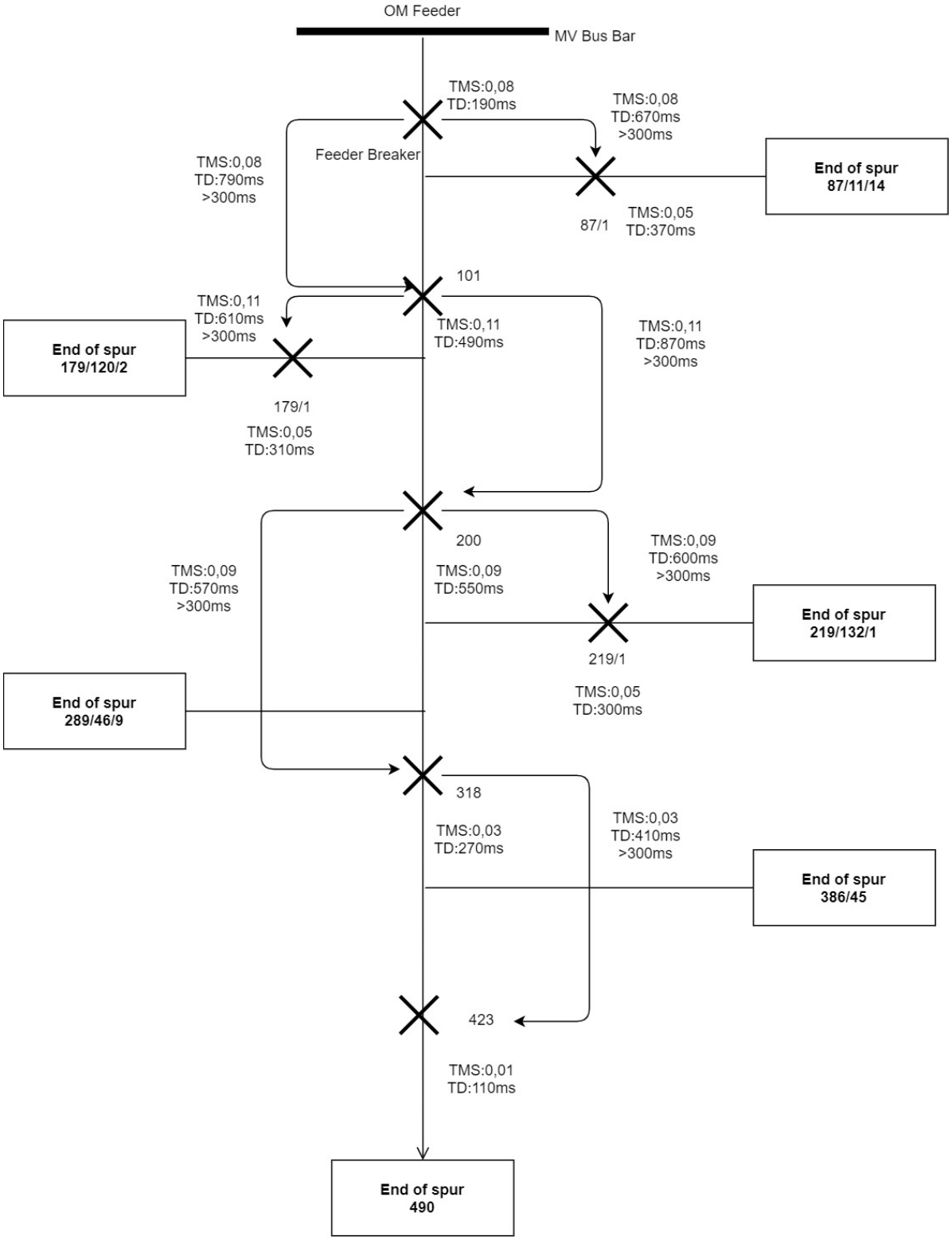


Figure 4-4: Graphical representation of the pick-up current setting selection philosophy (section 4.3) applied to the network in Figure 4-1

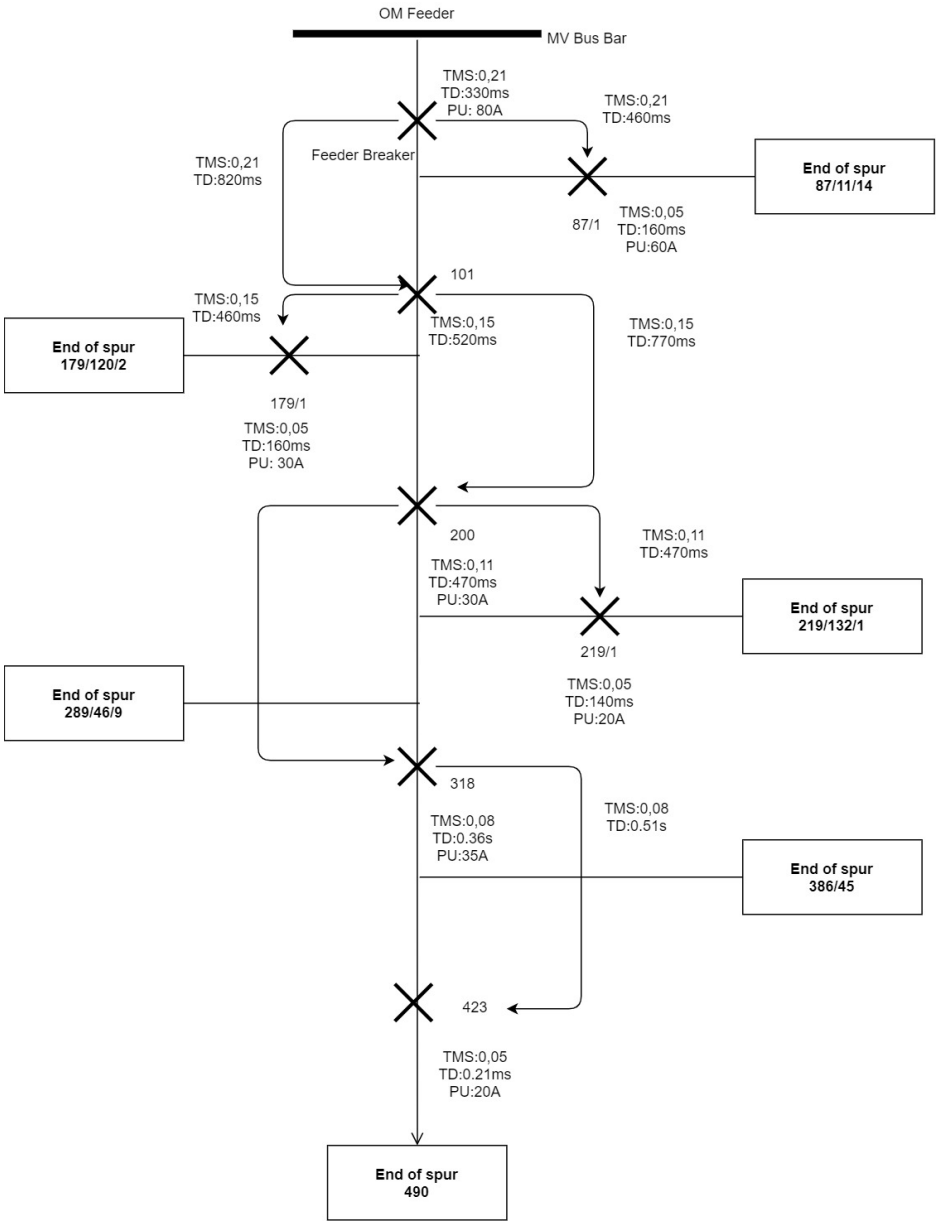


Figure 4-5: Graphical representation of Eskom MV earth fault pick-up current selection philosophy (section 4.4) applied to the network in Figure 4-1

4.5. Distribution network model

The MV protection philosophy was applied to three more distribution networks. As introduced partially in section 3.2, each distribution network falls in one of the distribution system classifications i.e. radial, ring, and parallel. As per Eskom's guidelines on data sensitization, the MV distribution network names from which the data was obtained were labelled A, B, and C. The Distribution networks are as shown in Figure 4-6 to Figure 4-8.

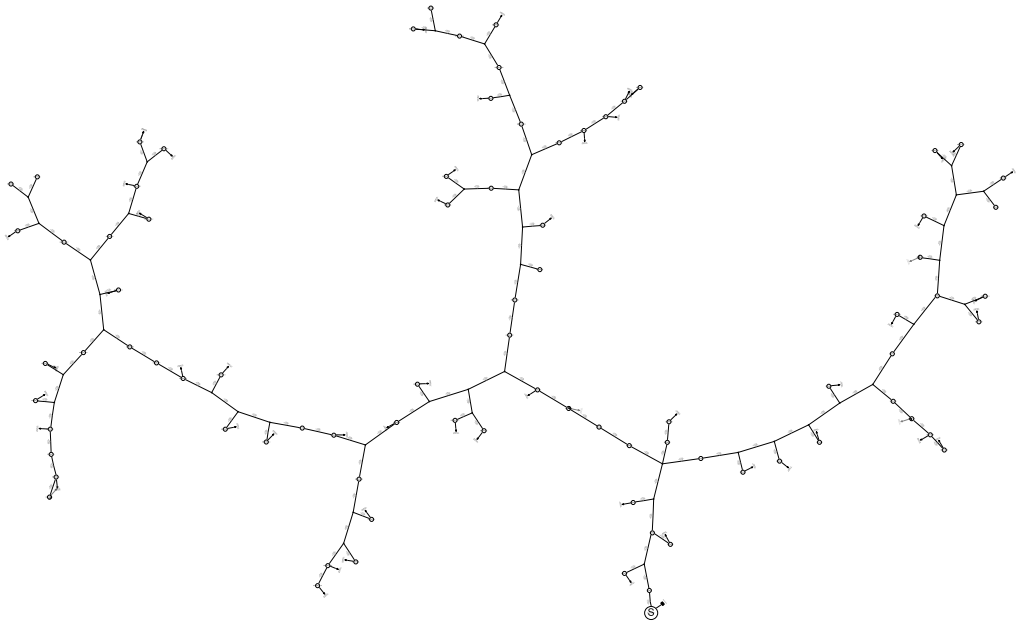


Figure 4-6: Distribution network A

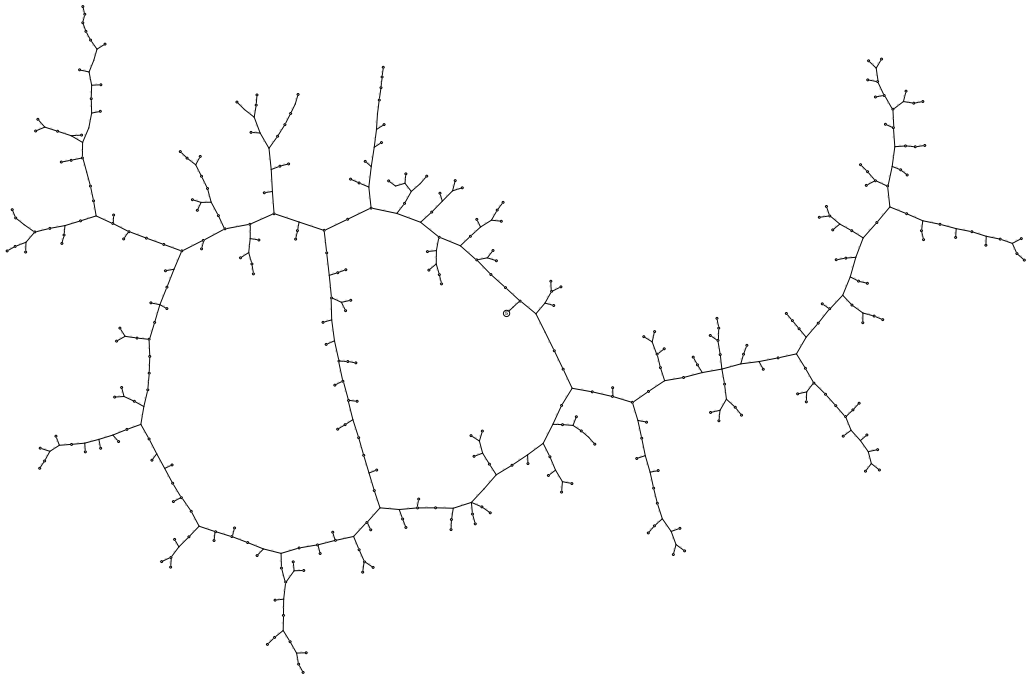


Figure 4-7: Distribution network B

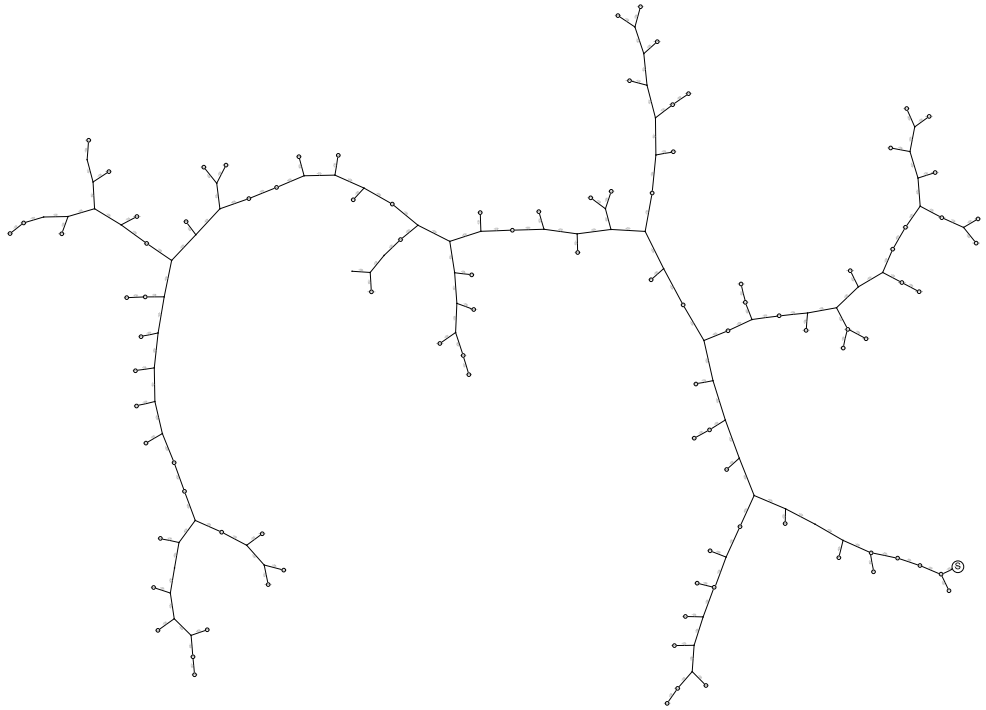


Figure 4-8: Distribution network C

4.5.1. Reduced distribution network diagrams

As mentioned in earlier sections of chapter 4, the first step in the application of the MV philosophy is the creation of a reduced network. The distribution networks in Figure 4-6, Figure 4-7, and Figure 4-8 were reduced to Figure 4-9 and Figure 4-10 and Figure 4-11.

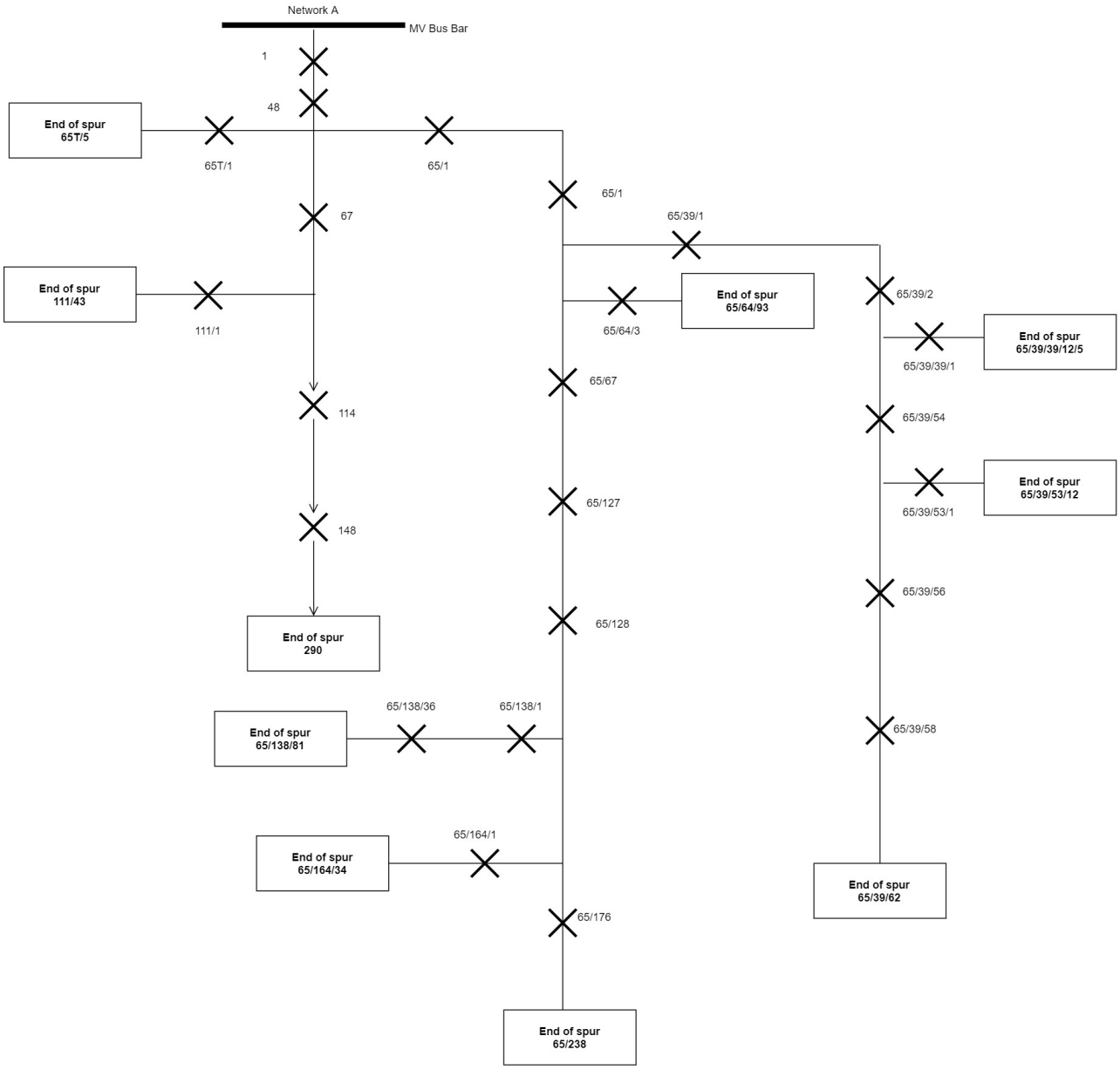


Figure 4-9: Reduced distribution network A diagram

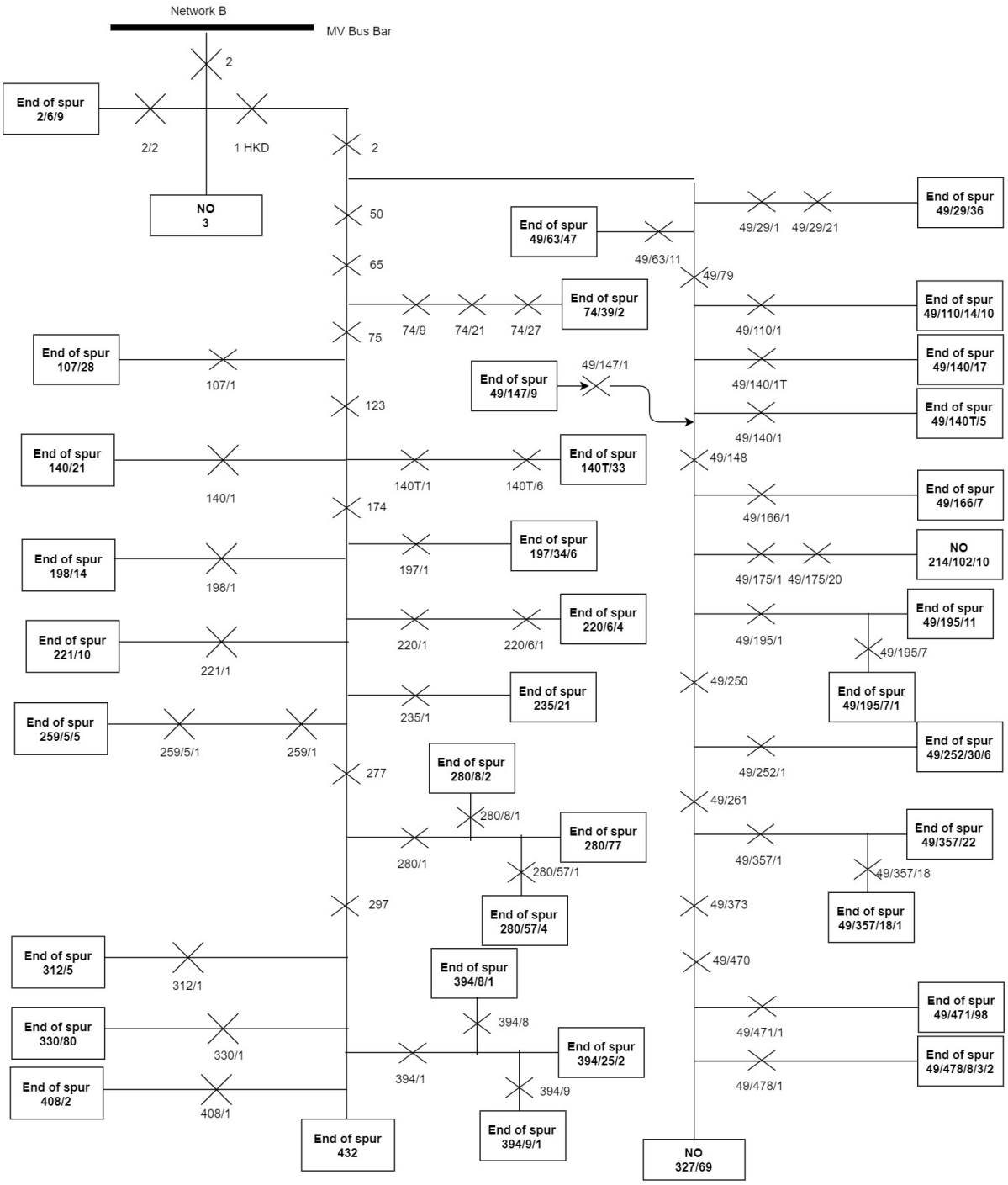


Figure 4-10: Reduced distribution network B diagram

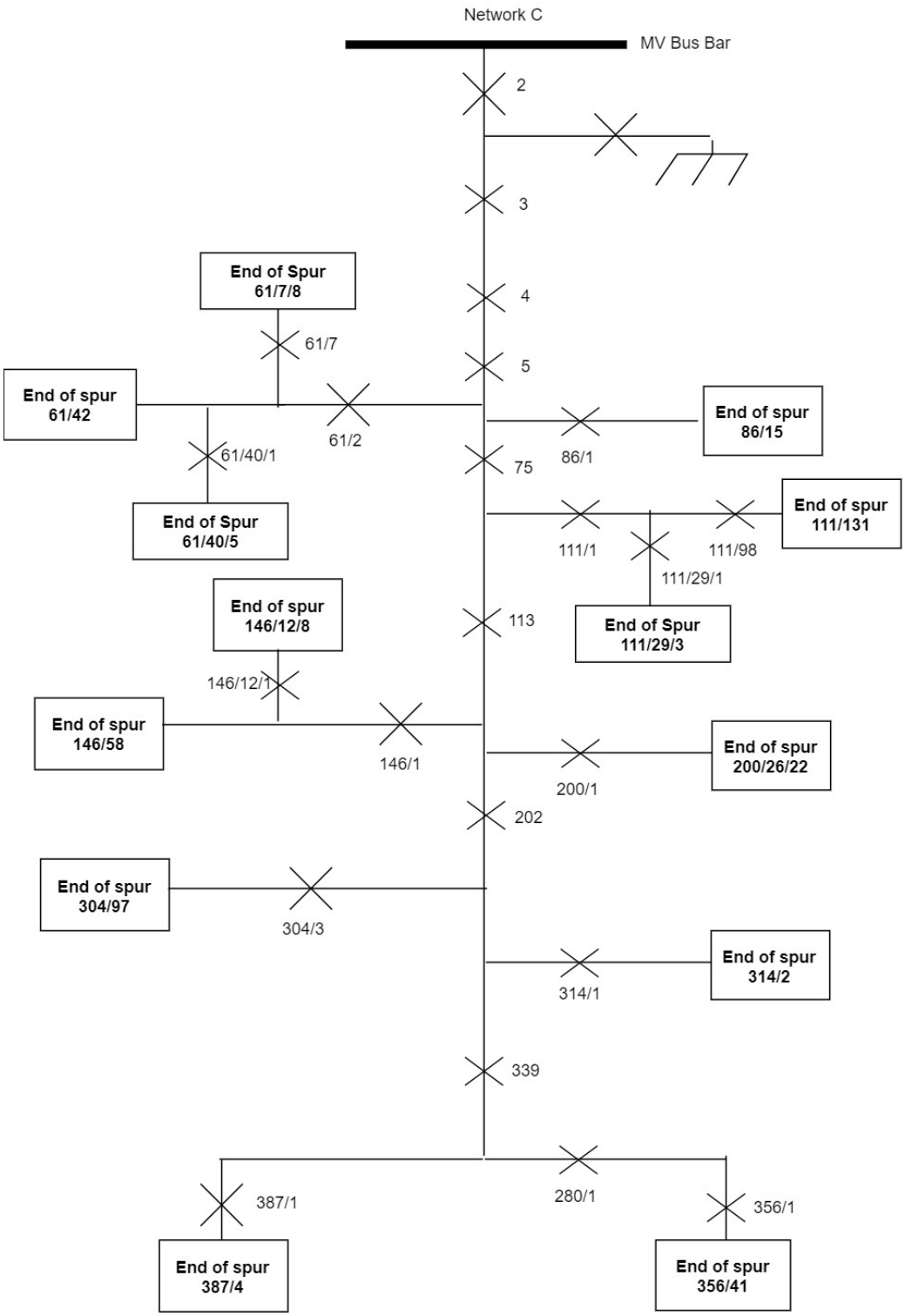


Figure 4-11: Reduced distribution network C diagram

4.5.2. Operating times and TMS values for the relays/reclosers on network A, B and C

For Figure 4-6, Figure 4-7, and Figure 4-8 the following data are known: the rating of the conductor to the first relay/recloser illustrated in Figure 4-12, Figure 4-13, and Figure 4-14. The transformer and MV circuit breaker parameters for each network are shown in Table 4-7, Table 4-8 and Table 4-9.

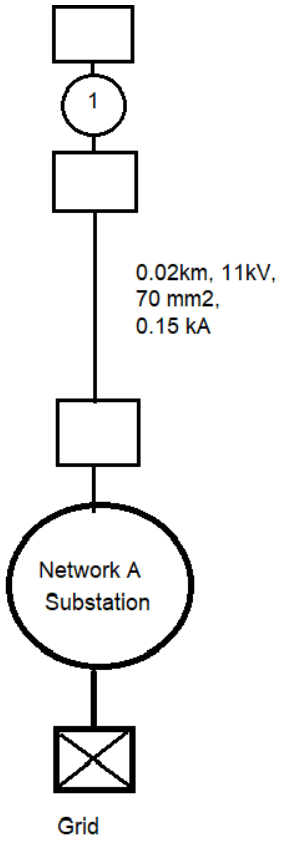


Figure 4-12: Conductor rating from network A's substation transformer

Table 4-7: Network A transformer and MV CB parameters

Transformer and MV busbar circuit breaker parameters	
Capacity	3 MVA
HV rating	44 kV
MV rating	11 kV
Overcurrent CT ratio	200:1
Earth fault CT ratio	100:1
Transformer circuit breaker operating time	1.50 s
Transformer earth fault pick-up current setting	100 A

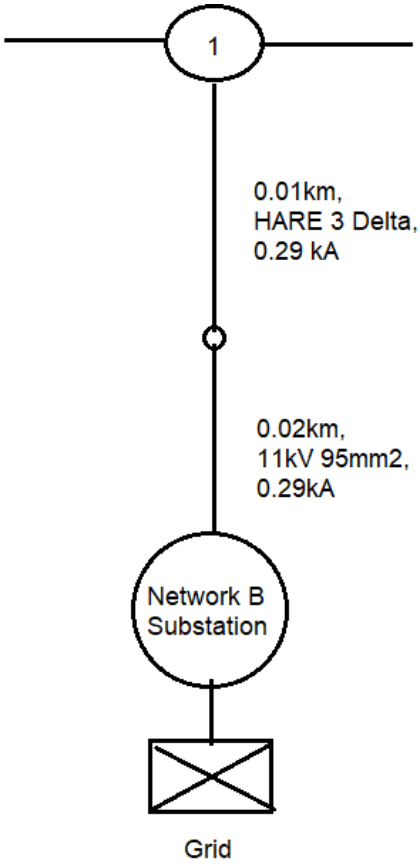


Figure 4-13: Conductor rating from network B's substation transformer

Table 4-8: Network B transformer and MV CB parameters

Transformer and MV busbar circuit breaker parameters	
Capacity	3 MVA
HV rating	44 kV
MV rating	11 kV
Overcurrent CT ratio	200:1
Earth fault CT ratio	100:1
Transformer circuit breaker operating time	1.50 s
Transformer earth fault pick-up current setting	100 A

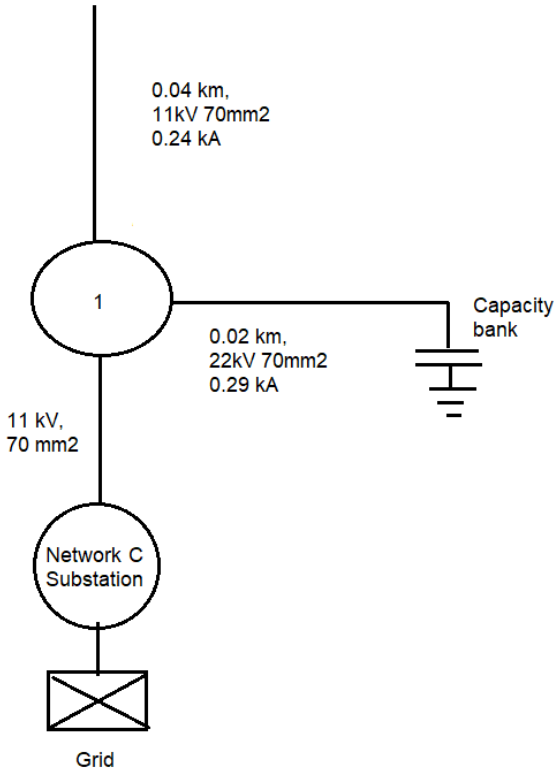


Figure 4-14: Conductor rating from network C's substation transformer

Table 4-9: Network C transformer and MV CB parameters

Transformer and MV busbar circuit breaker parameters	
Capacity	3 MVA
HV rating	44 kV
MV rating	11 kV
Overcurrent CT ratio	200:1
Earth fault CT ratio	100:1
Transformer circuit breaker operating time	1.50 s
Transformer earth fault pick-up current setting	100 A

Using the transformer and MV circuit breaker parameters, the pick-up current setting for the transformer circuit breakers were calculated as shown in equation 4-4, 4-5 and 4-6.

A $$I_F = 150\% \times \frac{S}{\sqrt{3} \times V_{LL}} = 150\% \times \frac{3\ MVA}{\sqrt{3} \times 11kV} = 236.19A \quad (4\text{-}4)$$

B $$I_F = 150\% \times \frac{S}{\sqrt{3} \times V_{LL}} = 150\% \times \frac{3\ MVA}{\sqrt{3} \times 11kV} = 236.19A \quad (4\text{-}5)$$

C $$I_F = 150\% \times \frac{S}{\sqrt{3} \times V_{LL}} = 150\% \times \frac{3\ MVA}{\sqrt{3} \times 11kV} = 236.19A \quad (4\text{-}6)$$

Thus, the first relay/recloser on the main feeder of network A, B and C have 236.19 A as the upstream pick-up current setting when the considerations in section 4.3 were applied. As indicated in Figure 4-12, Figure 4-13 and Figure 4-14, the conductor from the substation to the first relay/recloser/node has a surface area similar to that of a 0.1 hard-drawn copper (HDC) conductor with an emergency current rating of 321.8 A (Table D-2). In Figure 4-13, the conductor used from the first node to the first recloser is Hare, which has an emergency current rating of 380 A. The results from using the fault levels at each node based on the conductor rating, conductor length, equipment rating, and applying the considerations in section 4.3 to 4.4 on Figure 4-9, Figure 4-10 and Figure 4-11 are shown in Table 4-10, Table 4-11, and Table 4-12. Full data set can be found in appendix C.1.2, C-1.3, and C.1.4.

Table 4-10: Operating times and TMS values for the protective devices on network A

Relay/Recloser Position	TMS	Fault level (A)	Pick-up current setting (A)	Operating time (s)
148	0.01	519	125.52	0.05
114	0.07	519	139.47	0.35
67	0.10	630	154.97	0.60
1	0.24	1204	191.31	0.89
CB	0.37	2710	212.57	0.10
Trf	0.54	678	59.05	1.50
111/1	0.05	555	139.47	0.25
65T/1	0.05	860	154.97	0.20
65/176	0.05	264	91.51	0.33
65/127	0.12	328	112.97	0.81
65/67	0.15	330	125.52	1.10
65/1	0.26	888	154.97	1.02
48	0.17	901	172.18	1.32
65/39/58	0.06	379	82.36	0.27
65/39/56	0.12	379	91.51	0.57
65/39/54	0.17	393	101.67	0.86
65/39/2	0.21	398	112.97	1.15
65/39/1	0,26	582	125.52	1.18
65/2	0,19	587	139.47	1.47
65/64/3	0,06	473	125.52	0,31
65/164/1	0,06	277	91.51	0,38
65/128	0,09	277	101.67	0.68
65/138/36	0,08	266	82.36	0.47
65/138/1	0.06	266	91.51	0.77
65/39/39/1	0.08	435	101.67	0.38
65/39/53/1	0.08	398	91.51	0.38

Table 4-11: Operating times and TMS values for the protective devices on network B

Relay/Recloser	TMS	Fault level (A)	Pick-up current setting (A)	Operating time (s)
297	0.07	388	82.36	0.31
277	0.16	413	91.51	0.75
174	0.12	413	101.67	1.05
123	0.15	622	112.97	0.75
75	0.13	829	125.52	0.82
65	0.15	1208	139.47	0.68
50	0.24	1335	154.97	0.77
2	0.18	1833	172.18	0.97
1	0.30	3510	191.31	0.71
CB	0.42	3510	212.57	1.01
Trf	0.80	788	59.047	2.11
2/2	0.05	3510	154.97	0.11
74/21	0.05	490	101.67	0.22
74/27	0.11	337	112.96	0.70
74/9	0.14	292	125.52	1.17
107/1	0.05	764	112.97	0.18
140/6	0.05	393	91.51	0.24
140/1	0.11	393	101.67	0.54
197/1	0.05	511	91.51	0.20
198/1	0.05	507	91.51	0.20
220/6/1	0.05	370	82.36	0.23
220/1	0.11	370	91.51	0.53
235/1	0.05	397	91.51	0.24
259/5/1	0.05	383	82.36	0.22
259/1	0.11	383	91.51	0.52
280/8/1	0.05	358	74.12	0.22
280/1	0.11	358	82.36	0.52
280/57/1	0.05	284	74.12	0.26
49/373	0.09	236	91.51	0.628
49/250	0.11	387	112.97	0.94
49/148	0.12	400	125.52	0.90
49/79	0.12	590	139.47	0.82
49/29/21	0.06	803	125.52	0.22
49/29/1	0.13	803	139.47	0.52
49/1	0.21	1052	154.97	0.75
49/63/11	0.06	623	139.47	0.28
49/110/1	0.06	621	125.52	0.26
49/140/1	0.06	487	125.52	0.31
49/147/1	0.06	464	125.52	0.32
49/166/1	0.06	509	112.97	0.27
49/175/20	0.06	421	101.67	0.29
49/175/1	0.11	421	112.97	0.59
49/195/7	0.06	404	101.67	0.30

49/195/1	0.11	404	112.97	0.60
49/252/1	0.06	393	101.67	0.31
49/195/18	0.06	404	82.36	0.26
49/357/1	0.12	250	91.51	0.83
49/261	0.12	250	101.67	1.13
49/471/1	0.05	235	74.12	0.30
49/478/1	0.05	229	74.12	0.31
49/470	0.05	229	82.36	0.61

Table 4-12: Operating times and TMS values for the protective devices on network C

Relay/Recloser	TMS	Fault level (A)	Pick-up current setting (A)	Operating time (s)
4	0.26	2877	154.97	0.60
3	0.37	2877	172.18	0.90
2	0.48	2877	191.31	1.20
CB	0.57	2877	212.57	1.50
Trf	1.03	719	59.047	2.83
86/1	0.05	719	125.52	0.20
111/98	0.05	307	112.97	0.35
200/1	0.05	390	112.97	0.28
304/3	0.05	270	91.51	0.32
314/1	0.05	261	91.51	0.33
202	0.05	261	112.97	0.63
61/7/1	0.06	895	112.97	0.20
61/40/1	0.06	618	112.97	0.24
61/2	0.127	618	125.52	0.54
111/29/1	0.06	452	112.97	0.30
111/1	0.11	452	125.52	0.60
5	0.13	565	139.47	0.81
146/12/1	0.06	519	101.67	0.25
146/1	0.12	519	112.97	0.55
113	0.10	572	125.52	0.82
356/1	0.01	265	91.51	0.07
387/1	0.01	236	91.51	0.07
339	0.01	236	101.67	0.37

For the furthest point on the backbone, the last auto-recloser/relay is given a TMS of 0.01, a TMS of 0.05 is assigned for the furthest point on the T-offs/branch, a TMS of 0.06 for a T-off of a T-off to start the grading calculations. A minimum grading margin of 200 ms was used between the protection devices.

4.6. Network model used in the detailed analysis of the impact of DG on the power protection system (training data)

As discussed in section 2.8, the approach to determining protection-based DG penetration limits will be based on supervised machine learning methods. In supervised learning, the correct output from the training data set is known and expect that there is a relationship between the input and output variables. Each training sample comes with both input and output values [33].Results from the detailed analysis of the impact of PV DG on the power protection systems found in reference [22] were used to train the logistic and linear regression model in chapter 6. The training distribution network is shown in Figure 4-15.

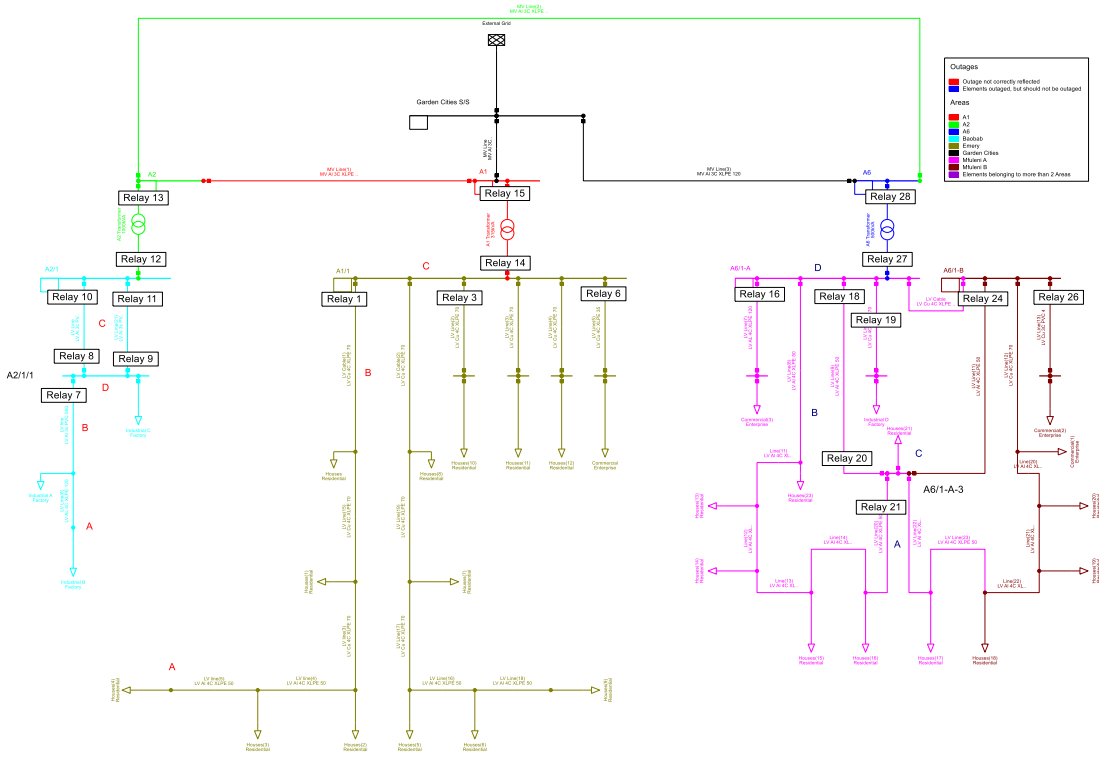


Figure 4-15: Network model for the detailed analysis of the impact of DG on the protection system (training network) [22]

For Figure 4-15, the following data is known: transformer and MV CB parameters (Table 4-13).

Table 4-13: Training network transformer and MV CB parameters

Transformer and MV busbar Circuit Breaker parameters	
Capacity	2 MVA
HV Rating	44 kV
MV Rating	11 kV
Overcurrent CT ratio	200:1
Earth Fault CT ratio	100:1
Transformer Circuit breaker operating time	1.50 s
Transformer earth fault pick-up current setting	100 A

Using the training distribution network's transformer and MV CB parameters (Table 4-13), the pick-up current setting for the transformer circuit breaker was calculated as shown in equation 4-7.

$$I_F = 150\% \times \frac{S}{\sqrt{3} \times V_{LL}} = 150\% \times \frac{2\ MVA}{\sqrt{3} \times 11kV} = 157.459A \quad (4\text{-}7)$$

Table 4-14 shows the relay operating times in Figure 4-15 during simulated fault conditions after three stages of DG penetration. Appendix E.1 presents additional details on the training data.

Table 4-14: Operating times and TMS values for the relays on the training network

Relay/ recloser	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)	Actual relay operating time (s) before	Relay operating times (s) after DG penetration of		
						35% of ADMD	65% of ADMD	75% of ADMD
6	0.11	587	147.86	0.50	6.98	10.73	14.48	21.97
6/1	0.43	587	28.85	0.30	1.14	1.75	2.37	3.59
2	0.25	749	152.10	0.86	80.11	18.51	15.05	9.78
2/1	1.03	749	33.94	0.66	61.48	14.21	11.55	7.51
1	0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81
1/1	1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48
6	0.11	587	147.86	0.50	1.74	9.31	3.54	3.63
6/1	0.43	587	28.85	0.30	0.723	1.20	1.47	1.51
2	0.25	749	152.10	0.86	6.64	7.59	10.17	13.29
2/1	1.03	749	33.94	0.66	5.10	5.82	7.81	10.20
1	0.67	1817	149.34	0.81	1.05	28.54	14.23	14.20
1/1	1.17	1817	67.56	0.61	0.82	8.22	11.06	11.04
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48
6	0.11	587	147.86	0.50	9.68	10.72	2.07	2.26
6/1	0.43	587	28.85	0.30	1.21	1.34	0.78	0.81
2	0.25	749	152.10	0.86	14.29	14.31	20.70	14.52
2/1	1.03	749	33.94	0.66	10.97	10.98	15.89	11.14
1	0.67	1817	149.34	0.81	1.01	1.06	1.10	1.08
1/1	1.17	1817	67.56	0.61	0.82	0.82	0.82	0.82
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48
6	0.11	587	147.86	0.50	0.46	0.46	0.46	0.46
6/1	0.43	587	28.85	0.30	0.32	0.32	0.32	0.32
2	0.25	749	152.10	0.86	0.85	0.85	0.85	0.64
2/1	1.03	749	33.94	0.66	0.75	0.75	0.75	0.75
1	0.67	1817	149.34	0.81	1.10	1.02	1.03	1.09
1/1	1.17	1817	67.56	0.61	0.82	0.82	0.82	0.82
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
Trf	1.10	1817	157.46	1.48	1.48	1.48	1.48	1.48

5. Recommendation systems

Recommender systems are built to look for related content from large and complex data sets [34]. Recommendation systems can be classified into three groups: content-based, collaborative filtering, and hybrid systems. Properties of the items being recommended are considered in content-based recommendation systems and the similarities between related items are used in collaborative filtering [35]. In hybrid recommender systems, two or more strategies are combined [36].

This chapter presents the application of recommendation systems on the PV and load data. The most common PV size, time of day when PV generation is at maximum, the location with the highest number of PV installations, other times in a day with a correlation to the maximum PV generation and minimum loading points, area PV installation growth rate and the most common PV sizes per postal code were recommended.

5.1. Most common PV sizes

To recommend the most common PV size, the complete data set of PV systems publicly available on the sunny portal [28], shown in Table 3-4 was used. The parent data frame was reduced to two columns: city and PV size (power). A *describe()* function (Figure 5-1) was used to determine the total number of data points, unique PV sizes and frequently repeated PV size (mode) within the data set.

```
#Data split
popular_PVsize =  pv_data[['City', 'Power']]
popular_PVsize.head()
```

	City	Power (kW)
0	Montana	0.0
1	Tshokwane Picnic Site Kruger National Park	0.0
2	Nkuhlu Picnic Site Kruger National Park	0.0
3	Edleen	0.0
4	Cape Town	0.0

```
popular_PVsize['Power'].astype(str).describe()
```

```
count     351
unique    227
top       5.3
freq       26
Name: Power, dtype: object
```

Figure 5-1: Output of a describe() function from the reduced parent data set

From the entire data set, the most common PV size was found to be 5.3 kW. To understand the distribution of the power generated by a 5.3 kW PV system, a *describe()* function was used on the actual power generation data recorded every 15 minutes (Table 3-1). The maximum output and 75th percentile power output were identified to be 3.96 kW and 2.77 kW, respectively (Table 5-1 and Figure 5-2).

Table 5-1: 5.3 kW PV generation data description

Count	Mean	Standard deviation	Power (kWp)				
			minimum	25%	50%	75%	Maximum
86	1.18	1.49	0	0	0.05	2.77	3.96

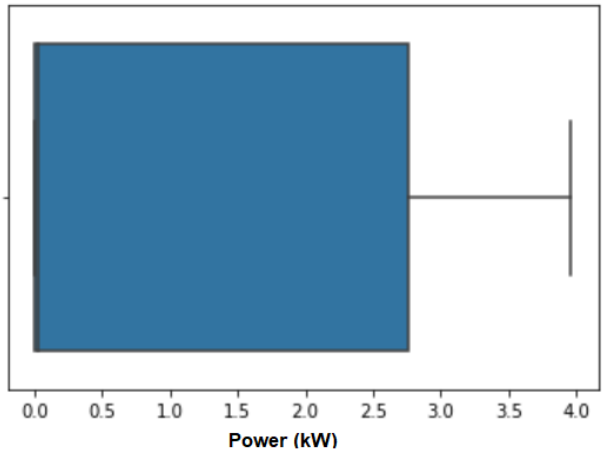


Figure 5-2: 5.3 kW PV generation data spread

5.2. Time of day when PV generation is at maximum

To recommend the time of day where a 5.3 kW PV generation peaked (3.96 kW), as shown in Table 5-1 and Figure 5-2, a *sort values()* function was used on the reduced parent data set (Figure 5-1). The time of day when the PV was at maximum generation was determined as 11:45 AM.

5.3. The place with the highest PV systems

Using recommendation by count on the PV: postal code and size data, the areas, by postal code with the highest number of PV systems was determined as shown in Figure 5-3 and Figure 5-4. Table 3-3 and Figure 3-4 in chapter 3 were used to translate the postal code data to places. The first two places were found to be Centurion (0157, Gauteng) followed by Somerset West area (7130, Western Cape). Both these areas consist of affluent neighbourhoods.

```
zip_count=pd.DataFrame(pv_data.groupby('Zip Code')['Power'].count())
zip_count.sort_values('Power', ascending=False).head()
```

Zip Code	Power (kW)
157	11
7130	9
7806	8
7600	8
81	6

Figure 5-3: Postal codes with most PVs

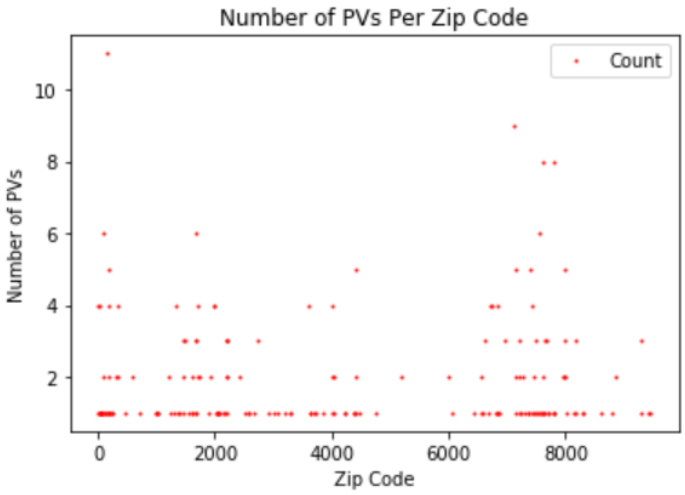


Figure 5-4: Number of PVs per postal code

5.4. Time of day where maximum generation and minimum loading occurs

By merging the two data frames (5,3 kW PV generation: Table 3-1 and load data: Table 3-2) on the same time axis and sorting the data, the time of day where the 5.3 kW PV generation was at maximum and the distribution network was under minimum loading conditions was found to be 11:30 AM -12:00 PM as indicated in Figure 5-5.

```
summary = pd.merge(loading, pv_maxGen, on='Time')
summary=summary[['Load(KVA)','StartkW']]
summary.sort_values('StartkW', ascending=False).head()
```

Time	Load (kVA)	Solar Power (kW)
11:30 AM	511.509896	3.913333
12:00 PM	512.259176	3.856667
11:00 AM	514.310195	3.626667
10:30 AM	520.212243	3.610000
12:30 PM	506.563976	3.470000

Figure 5-5: Load and PV generation data on the same time axis

5.5. Other times in a day with a correlation to the maximum PV generation and minimum loading points

To find other times in a day where minimum loading and maximum PV generation is observed, similarities based on Pearson's correlation coefficient (Pearson's R) were evaluated. The minimum loading time, determined as 11:30 AM in section 5.4 and the maximum PV generation time, determined as 11:45 AM in section 5.2 were used as pivot points. Figure 5-6 and Figure 5-7 present other times with similar minimum loading and maximum PV generation behaviour.

```
similar_to_loadmin = load_data_crosstab.corrwith(load_min)

corr_loadmin = pd.DataFrame(similar_to_loadmin, columns=['PearsonR'])
corr_loadmin.dropna(inplace=True)
corr_loadmin.head()
```

Time	Pearson's R
01:00 AM	0.432128
01:00 PM	0.956746
01:30 AM	0.448820
01:30 PM	0.901800
02:00 AM	0.434116

Figure 5-6: Loading similar to 11:30 AM

```
similar_to_Peak = pv_starts_crosstab.corrwith(peak_production_1)

corr_Peak = pd.DataFrame(similar_to_Peak, columns=['PearsonR'])
corr_Peak.dropna(inplace=True)
corr_Peak.head()
```

Time	Pearson's R
10:00 AM	1.0
10:15 AM	1.0
10:30 AM	1.0
10:45 AM	-1.0
11:00 AM	1.0

Figure 5-7: Generation similar to 11:45 AM

The time of day with a high correlation to minimum loading was 1:00 PM (Figure 5-6), 10:00 AM, 10:15 AM, 10:30 AM, and 11:00 AM for maximum PV generation (Figure 5-7).

5.6. Most common PV sizes per postal code

To find the most common PV size in each location (postal code) from the complete data set of PV systems publicly available on the sunny portal [28], the parent data frame was split to PV size, postal code, and date of installation. To reduce the number of groups from the date of installation entry, only the year was used (transition from Table 5-2 to Figure 5-8).

Table 5-2: PV systems sampled data [28]

PV System Name	Country	Zip Code	City	Power (kWp)	Date of installation
Weenen	South Africa	9460	Wesselsbron	25.75	12/04/2015
H9	South Africa	9430	Virginia	5.30	07/03/2018
G. Scholtz 47 Koedoeweg, Woodl	South Africa	9301	Bloemfontein	4.95	22/01/2019
Mr JPC Antunes	South Africa	9301	Bloemfontein	5.60	10/01/2018

```
pv_data =  pv_data[['Zip Code', 'Power','date' ]]
pv_data['date'] = pd.DataFrame(pv_data['date'].str[6:])
pv_data =  pv_data[['Zip Code', 'Power','date' ]]
pv_data.head()
```

	Zip Code	Power	date
0	182.0	0.0	2017
1	1350.0	0.0	2018
2	1350.0	0.0	2018
3	1709.0	0.0	2018
4	7405.0	0.0	2014

Figure 5-8: PV system reduced data set

Grouping the data by PV size, postal code, counting postal code occurrences, and sorting the count by descending order, the number of each PV size per postal code was determined as shown in Figure 5-9 (first 5 rows).

```
zip_count=pd.DataFrame(pv_data.groupby(['Power', 'Zip Code'])['Zip Code'].count())
zip_count = zip_count.rename(columns = {"Power": "Power",
                                        "Zip Code":"Zip Code",
                                        "Zip Code": "Zip Code Count"})
zip_count.sort_values('Zip Code Count', ascending=False).head()
```

Power (kW)	Zip Code	Zip Code Count
5.30	27.0	3
15.90	1.0	2
6.75	7646.0	2
10.71	348.0	2
5.00	2001.0	2

Figure 5-9: Number of each PV size per postal code

The most common PV sizes per postal/zip code were found to be less than 10 kW. Full data set in Figure 5-9 was graphically represented in Figure 5-11 to Figure 5-12 (y-axis limited to 40 kW).

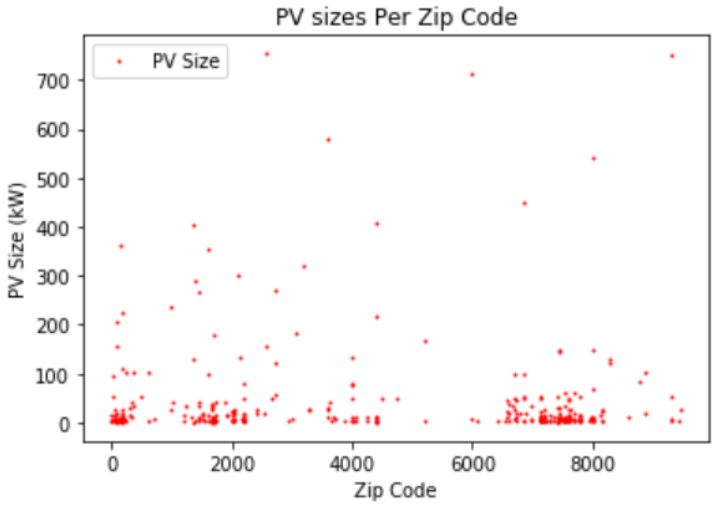


Figure 5-10: PV sizes per postal code

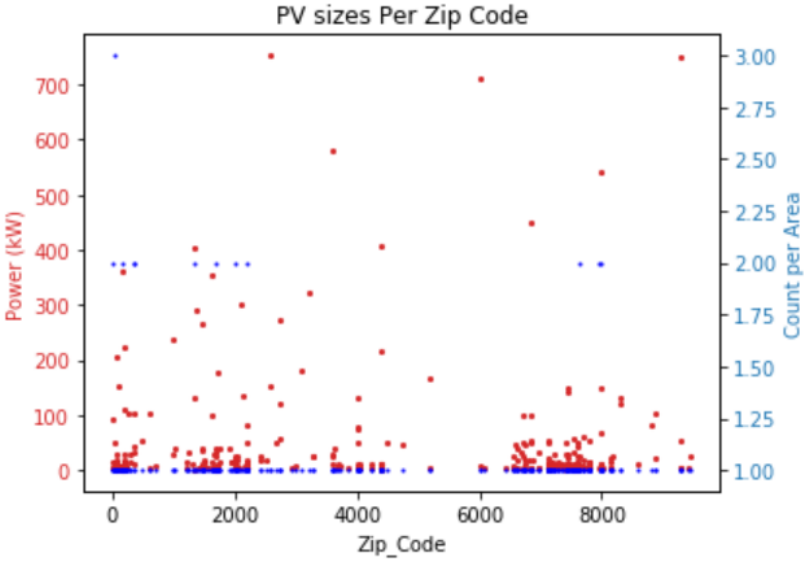


Figure 5-11: PVs sizes and count per postal/zip code (area)

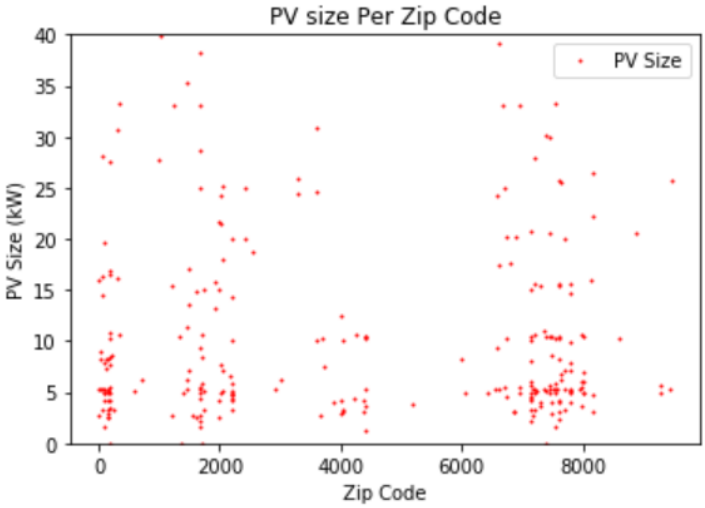


Figure 5-12: PV sizes (40 kW max) per postal code

5.7. PV size yearly growth rate

To determine the number of PV size installation per year, the newly constructed data frame in section 5.6 was grouped by date and PV size. By sorting the PV size count by descending order, the number of PV size installation per year was determined as indicated in Figure 5-13 (first 5 rows).

```
power_count=pd.DataFrame(pv_data.groupby(['date','Power'])['Power'].count())
power_count = power_count.rename(columns = {"date": "Date",
                                            "Power": "Power",
                                            "Power": "Power Count"})
power_count.sort_values('Power Count', ascending=False).head()
```

Date	Solar Power (kW)	Solar Power Count
2018	5.3	6
2015	5.3	6
	5.0	5
2016	5.3	3
	5.0	3

Figure 5-13: The number of PV size installation per year

Full data set in Figure 5-13 was graphically represented in Figure 5-14 and Figure 5-15 (y-axis limited to 40 kW).

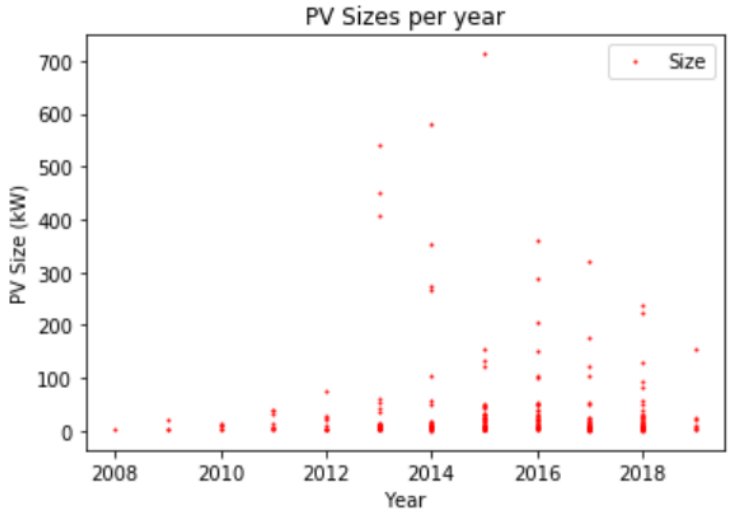


Figure 5-14: PV Size installations per year

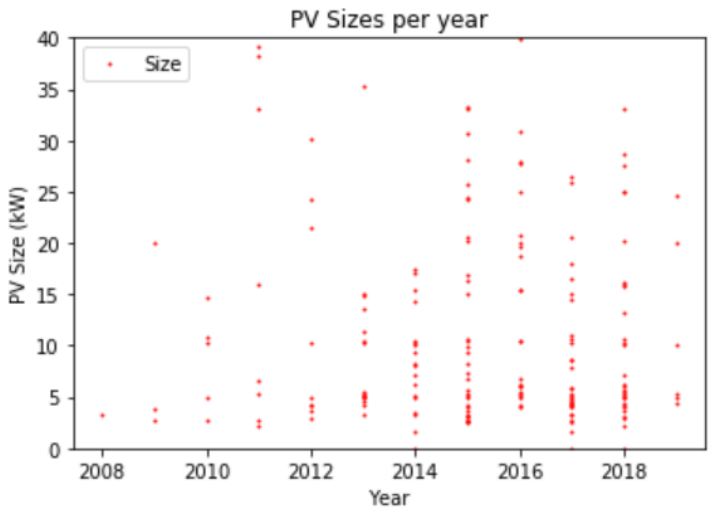


Figure 5-15: PV Size (40 kW max) installations per year

Figure 5-15 indicate growth in the installations of PV systems with less than 10 kW every year, a PV installation growth rate of 1.5 times the number of 5.3 kW installations in the previous year.

5.8. The growth rate of PV installation by area

To determine the growth rate of area PV installations, the newly constructed data frame in section 5.6 was grouped by date, zip code, and the date count sorted by descending order. The number of area PV installation per year was determined as indicated in Figure 5-16 (first 5 rows).

```
date_count=pd.DataFrame(pv_data.groupby(['date','Zip Code'])['date'].count())
date_count = date_count.rename(columns = {"date": "Date",
                                          "date": "date_count"})
date_count.sort_values('date_count', ascending=False).head()
```

Date	Zip Code	Date Count
2018	157.0	5
2012	2001.0	4
2013	4400.0	4
2015	348.0	4
	7806.0	4

Figure 5-16: Number of area PV installation per year

The full data set in Figure 5-16 was graphically represented in Figure 5-18 and Figure 5-17.

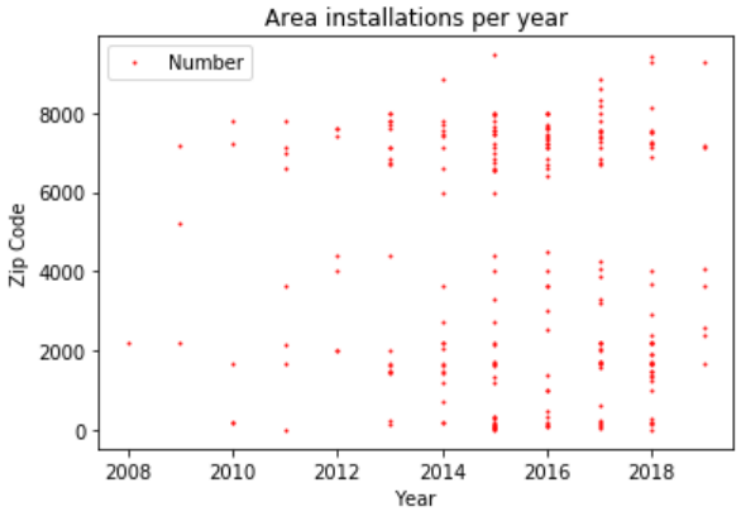


Figure 5-17: Area PV installation per year

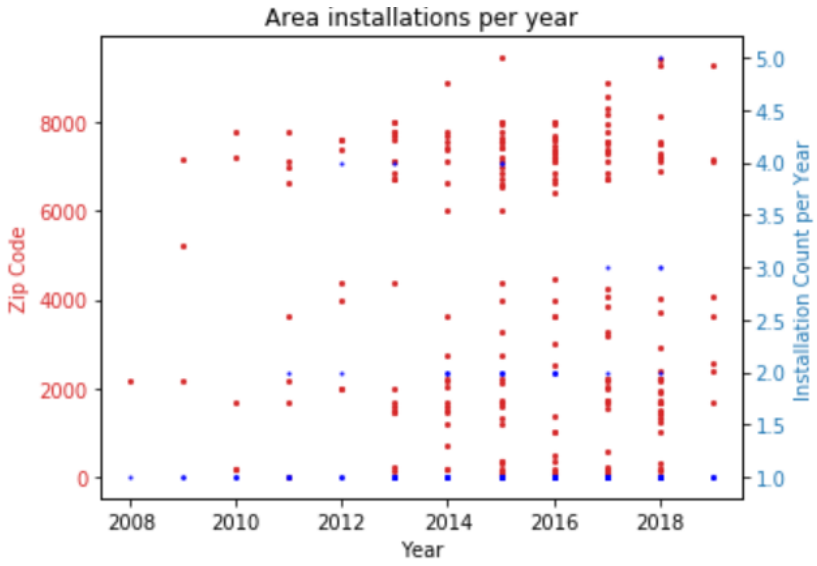


Figure 5-18: Area PV installation and count per year

Figure 5-17 indicates growth in two area clusters (postal code range 0 to 2000 and 6000+).

5.9. Chapter discussion

The network DG planning dilemma introduced in section 1.1 is:

Protective systems are designed such that lower short circuit levels are experienced by downstream devices compared to the short circuit levels that are upstream (unidirectional power flow). The original overcurrent, earth fault, and sensitive earth fault protection coordination designs were not performed with DG source integration considerations. Therefore, for certain levels of DG penetration (feeder, size, and type-specific), inadequacy in existing protection coordination and alteration of the system short circuit current characteristics result in loss of protection coordination [3] [4] *The other problem with distributed generators is the vast difference in the technology, location, size, connection sequence, and protection scheme requirements which result in future DG network planning inaccuracies.*

The results in this chapter- areas with a high number of PVs (section 5.3 and Figure 5-4), the maximum generation output of a 5.3 kW PV (section 5.2), peak generation times, minimum loading times (Figure 5-5), similar to peak and minimum times (section 5.5), popular PV size (Figure 5-1), PV size per postal code(Figure 5-11), PV size installation per year (Figure 5-14), area PV installation per year (section 5.8)- solve the second part of the network DG planning dilemma, by answering the when, where, and likely PV size that will be connected to the network. To achieve a better resolution of where PVs are currently installed, a heat map was created as shown in Figure 5-19.

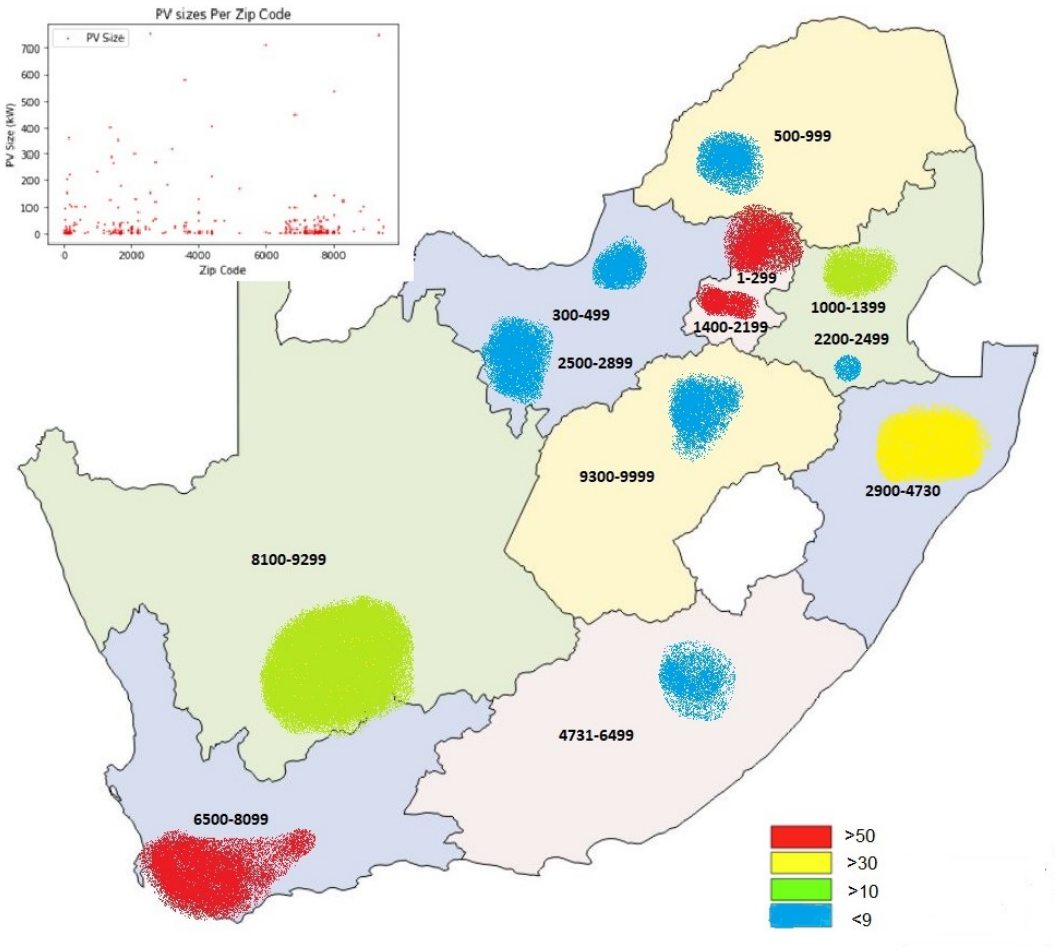


Figure 5-19: PV size clusters and postal code range per province

By using both the actual PV installation heat map (Figure 5-19) and the global horizontal irradiation map (Figure 3-5) in DG penetration planning, a better resolution of where PV installations will occur is achieved.

6. Model training: value estimation and classification systems

This chapter presents the trainings of models that can predict relay operating times at certain DG penetration levels, from the current network parameters and which class of protection miscoordination is likely to occur. Thus, solving the first part of the network DG planning dilemma.

Results from the detailed analysis of the impact of DG on the power systems protection in [37] were used as training data for both the supervised machine learning regression and classification model. The regression model was used to calculate relay/recloser actual operating times before DG penetration and operating times after three stages of DG penetration. The classification model determines which relay/recloser protection feature results in which coordination problem after three stages of DG penetration.

6.1. Model classes and success criteria

In this section, the DG protection hosting capacity calculation considerations for each protection miscoordination group was defined. As discussed in section 4.2, relay/recloser operating time used in protection coordination is calculated using equation 6-8.

$$\text{Normal inverse} \qquad t = 0{,}14 \times \frac{TMS}{((\frac{I_{Fault}}{I_{PU}})^{0.02} - 1)} \qquad (6\text{-}8)$$

Where,

TMS: time multiplier setting
I_{Fault}: measured fault current magnitude.
I_{PU}: Pick-up current setting
t: operating time

In section 2.3, the impact of DGs on protection philosophies was introduced. Loss of protection coordination was cast into three groups: loss of coordination, loss of grading and de-sensitization. The success criteria for the models will be the accuracy in predicting the relay operating times at different DG penetration levels and the accuracy in classifying the protection miscoordination based on the loss of protection coordination groups.

6.2. Linear regression as a relay operating time predictor

The aim is to train a model that can predict the relay operating time at certain DG penetration levels from the current network protection parameters. Four combinations of the network protection parameters (features) were used as inputs i.e., all features as inputs, fitting (adding polynomials), eliminating features, adding polynomials and eliminating certain features.

6.2.1. All features as inputs

When predicting the relay operating times using all the features as inputs, the data in Table 4-14 was split into two: inputs (five features) and outputs (four relay operating times) as shown in Table 6-1. The model was used to predict the relay operating times before and after DG penetration for each penetration scenario.

Table 6-1: Inputs (recloser number, TMS, fault level, pickup current and calculated operating time) and outputs (actual relay operating times before and after DG penetration) for the all features as inputs model (sample)

Features (inputs)					Output			
Relay/ recloser	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of		
						35% of ADMD	65% of ADMD	75% of ADMD
6	0.11	587	147.86	0.50	6.98	10.73	14.48	21.97
6/1	0.43	587	28.85	0.30	1.14	1.75	2.37	3.59
2	0.25	749	152.10	0.86	80.11	18.51	15.05	9.78
2/1	1.03	749	33.94	0.66	61.48	14.21	11.55	7.51
1	0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81
1/1	1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48

The hypothesis function (introduced in appendix B1.1a) is defined as shown in equation 6-9.

$$h_\Theta(x) = \Theta_0 + \Theta_1 x \qquad (6\text{-}9)$$

Where,

$h_\Theta(x)$ is the relay operating time,
x is the input (five features), and
Θ_0, Θ_1 are parameters gradient descent is trying to optimise.

To evaluate the model's performance and ensure that the gradient descent is working, the graph of the cost function vs the number of iterations should decrease with each iteration. Table 6-2 shows a summary of gradient descent on the *all features as inputs* model's cost function for the four scenarios (before and after DG penetration).

Table 6-2: A summary of gradient descent on the all features as inputs model's cost function

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Gradient descent on the model's cost function decreased and saturated between a cost of:			
115 and 120 with a learning rate of 0.1 (Figure 6-1).	**15 and 20** with a learning rate of 0.1 (Figure 6-2).	**10 and 15** with a learning rate of 0.1 (Figure 6-3).	**5 and 10** with a learning rate of 0.1 (Figure 6-4).

Gradient descent on the *all features as inputs* model's cost function for the other learning rates (0.03, 0.01 and 0.0037) decreased and saturated at a cost higher than that of the 0.1 learning rate for all the DG penetration scenarios.

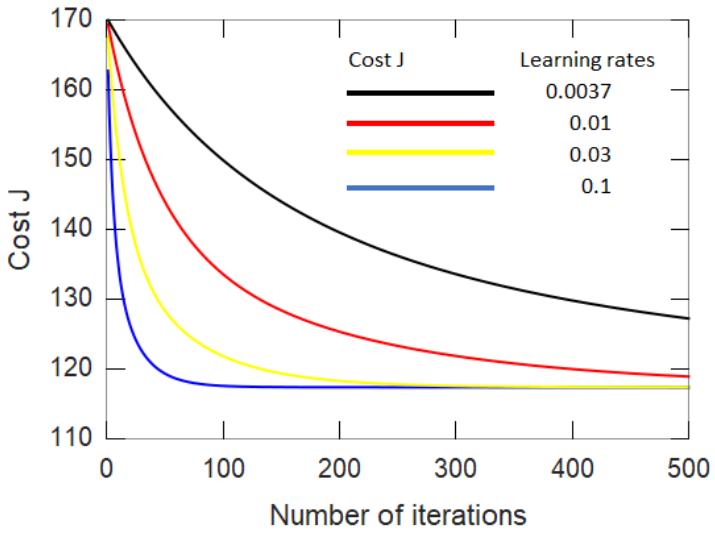


Figure 6-1: Convergence of the all features as inputs model's gradient descent for four different learning rates

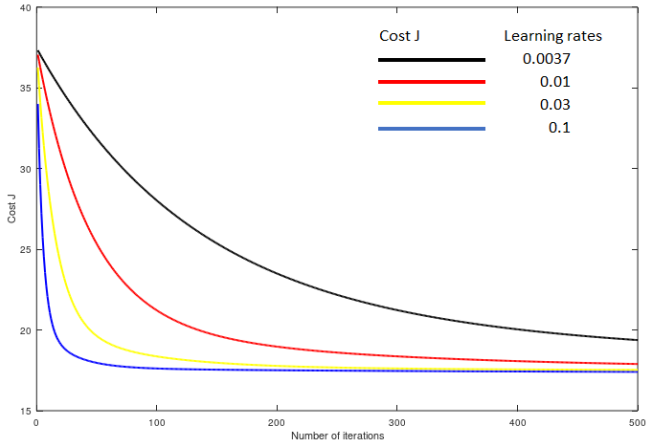


Figure 6-2: Convergence of the all features as inputs model's gradient descent after 35% of DG penetration for four different learning rates

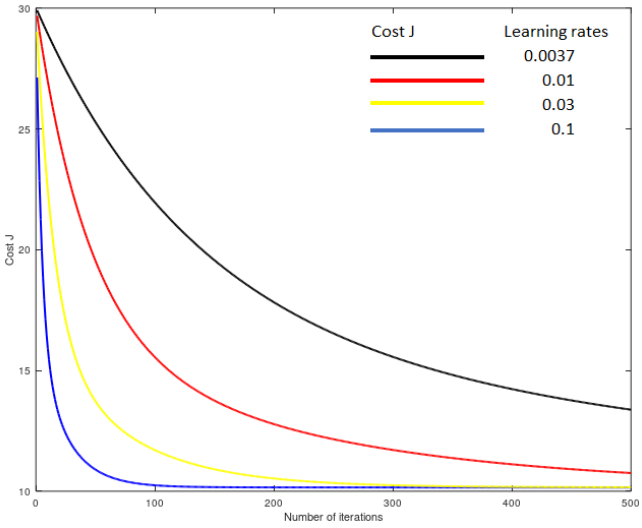


Figure 6-3: Convergence of the all features as inputs model's gradient descent after 65% of DG penetration for four different learning rates

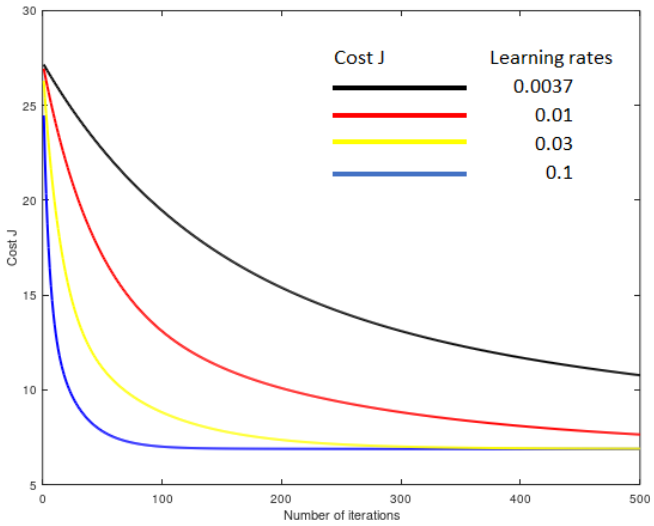


Figure 6-4: Convergence of the all features as inputs model's gradient descent after 75% of DG penetration for four different learning rates

6.2.2. Fitting (polynomials as added inputs)

When predicting the relay operating times using all the features with second-degree polynomials as additional inputs, the data in Table 4-14 was modified and split into two: inputs (nine features) and outputs (four relay operating times) as shown in Table 6-3.

Table 6-3: Inputs (second-degree polynomials-TMS, fault level, pickup current and calculated operating time) and outputs for the second-degree polynomials as additional inputs model (sample)

Features (inputs)					**Output**			
Features in Table 6-1	***Calculated operating time*²**	***TMS*²**	***Fault Level*²**	***Pick up current*²**	**Actual relay operating time (s) before DG penetration**	**Relay operating times (s) after DG penetration of**		
						35% of ADMD	**65% of ADMD**	**75% of ADMD**
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81
	0.37	1.37	3301489	4564.75	4.87	5.61	4.90	4.65
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48

The model was used to predict the relay operating times. Table 6-4 shows a summary of gradient descent on the *polynomials as added inputs* model's cost function for the four scenarios (before and after DG penetration).

Table 6-4: A summary of gradient descent on the polynomials as added inputs model's cost function

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Gradient descent on the model's cost function decreased and saturated between a cost of:			
110 and 120 with a learning rate of 0.1 (Figure 6-5).	**15 and 20** with a learning rate of 0.1 (Figure 6-6).	**10 and 15** with a learning rate of 0.1 (Figure 6-7).	**5 and 10** with a learning rate of 0.1 (Figure 6-8).

Gradient descent on the *polynomials as added inputs* model's cost function for the other learning rates (0.03, 0.01 and 0.0037) decreased and saturated at a cost higher than that of the 0.1 learning rate for all the DG penetration scenarios.

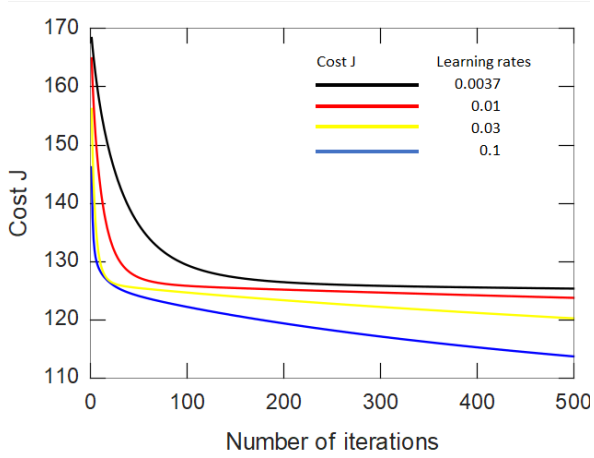


Figure 6-5: Convergence of the second-order polynomials as inputs model's gradient descent for four different learning rates

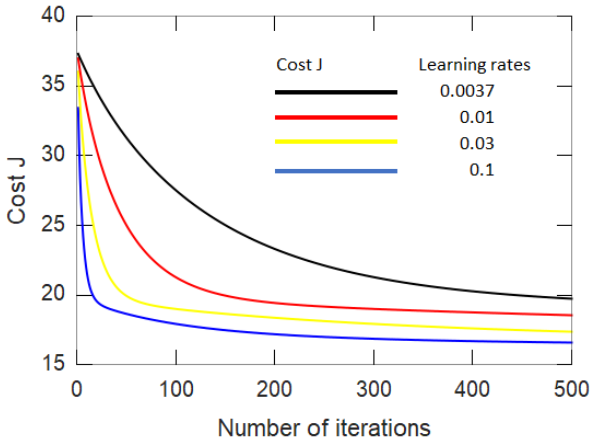


Figure 6-6: Convergence of the second-order polynomials as inputs model's gradient after 35% of DG penetration for four different learning rates

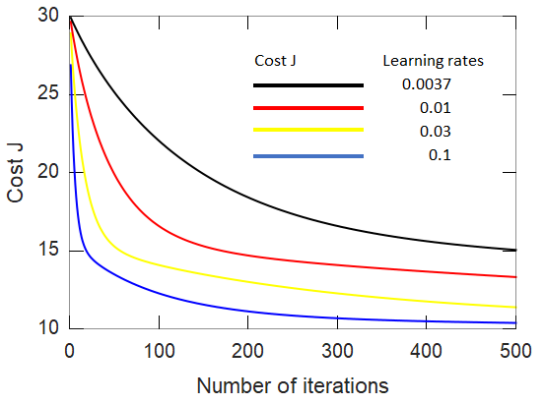


Figure 6-7: Convergence of the second-order polynomials as inputs model's gradient after 65% of DG penetration for four different learning rates

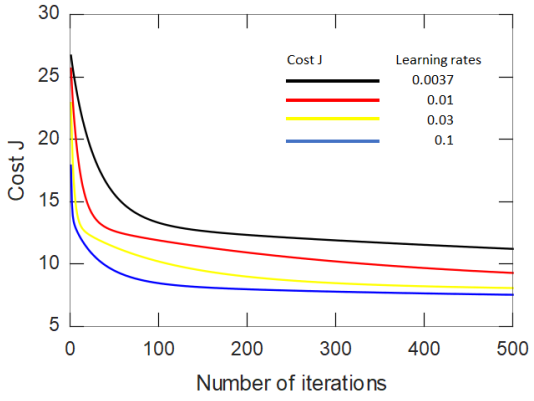


Figure 6-8: Convergence of the second-order polynomials as inputs model's gradient after 75% of DG penetration for four different learning rates

6.2.3. Eliminating features

When predicting the relay operating times using some of the features as inputs (relay/recloser positions eliminated), the data in Table 4-14 was split into two: inputs (four features) and outputs (four relay operating times) as shown in Table 6-5.

Table 6-5: Inputs (relay/recloser number eliminated) and outputs for the eliminated features as inputs model (sample)

Features (inputs)				**Output**			
TMS	**Fault level (A)**	**Pick-up current setting (A)**	**Calculated operating time (s)**	**Actual relay operating time (s) before DG penetration**	**Relay operating times (s) after DG penetration of**		
					35% of ADMD	**65% of ADMD**	**75% of ADMD**
0.11	587	147.86	0.50	6.98	10.73	14.48	21.97
0.43	587	28.85	0.30	1.14	1.75	2.37	3.59
0.25	749	152.10	0.86	80.11	18.51	15.05	9.78
1.03	749	33.94	0.66	61.48	14.21	11.55	7.51
0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81
1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65
0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18
1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48

The model was used to predict the relay operating times. Table 6-6 shows a summary of gradient descent on the *feature elimination* model's cost function for the four scenarios (before and after DG penetration).

Table 6-6: A summary of gradient descent on the feature elimination model's cost function

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Gradient descent on the model's cost function decreased and saturated between a cost of:			
135 and 140 with a learning rate of 0.1 (Figure 6-9).	**20** with a learning rate of 0.1 (Figure 6-10).	**16 and 18** with a learning rate of 0.1 (Figure 6-11).	**14 and 16** with a learning rate of 0.1 (Figure 6-12).

Gradient descent on the *feature elimination* model's cost function for the other learning rates (0.03, 0.01 and 0.0037) decreased and saturated at a cost higher than that of the 0.1 learning rate for all the DG penetration scenarios.

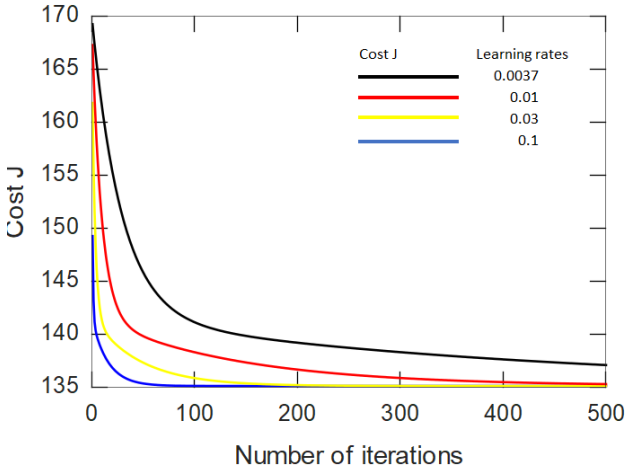


Figure 6-9: Convergence of the feature elimination model's gradient descent for four different learning rates

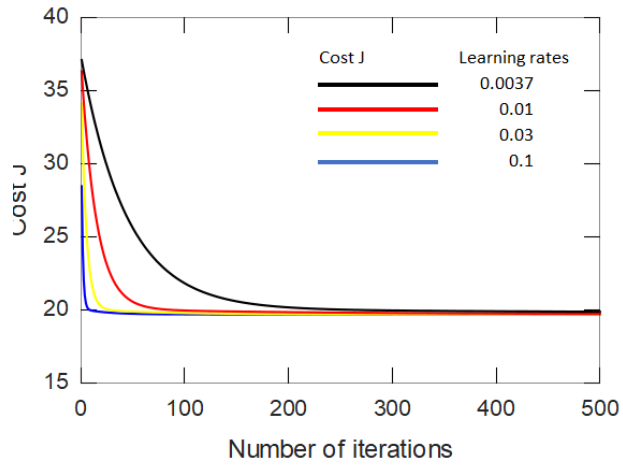


Figure 6-10: Convergence of the feature elimination model's gradient descent after 35% of DG penetration for four different learning rates

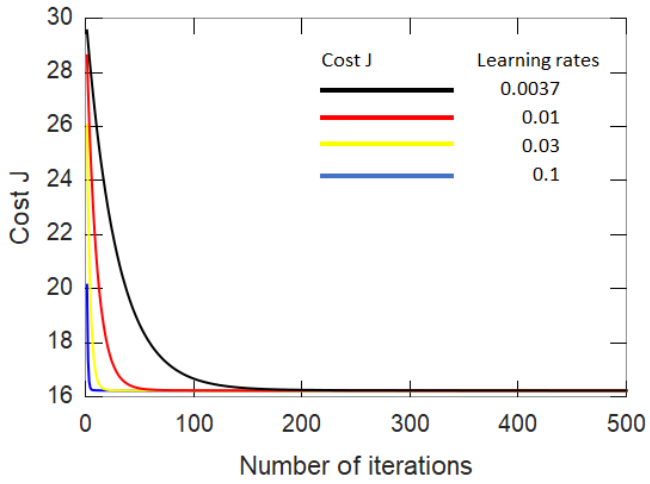


Figure 6-11: Convergence of the feature elimination model's gradient descent after 65% of DG penetration for four different learning rates

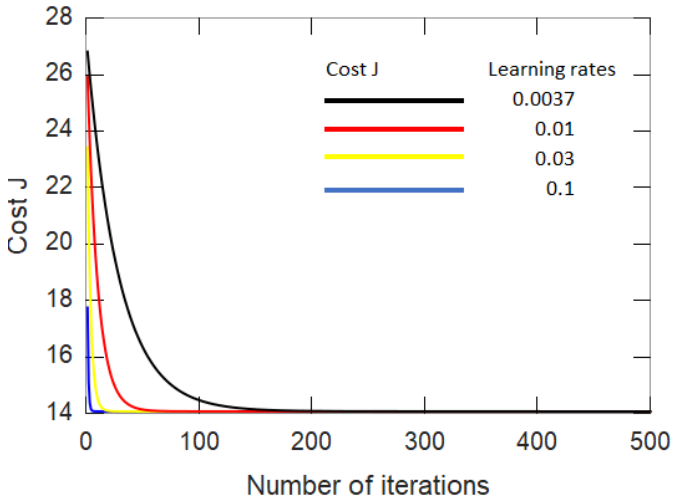


Figure 6-12: Convergence of the feature elimination model's gradient descent after 75% of DG penetration for four different learning rates

6.2.4. Fitting and eliminating features

When predicting the relay operating times using some of the features (relay/recloser positions eliminated) with second-degree polynomials as additional inputs, the data in Table 4-14 was modified and split into two: inputs (eight features) and outputs (four relay operating times) as shown in Table 6-7.

Table 6-7: Inputs (added second-degree polynomials and relay number eliminated) and outputs for the fitting and feature elimination model (sample)

Features (inputs)					Output			
Features in Table 6-5	***Calculated operating time²***	***TMS²***	***Fault Level²***	***Pick up current²***	**Actual relay operating time (s) before DG penetration**	**Relay operating times (s) after DG penetration of**		
						35% of ADMD	**65% of ADMD**	**75% of ADMD**
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81
	0.37	1.37	3301489	4564.75	4.87	5.61	4.90	4.65
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48

The model was used to predict the relay operating times. Table 6-8 shows a summary of gradient descent on the *fitted and eliminated feature* model's cost function for the four scenarios (before and after DG penetration).

Table 6-8: A summary of gradient descent on the fitted and eliminated feature model's cost function

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Gradient descent on the model's cost function decreased and saturated between a cost of:			
110 and 120 with a learning rate of 0.1 (Figure 6-13).	**15 and 20** with a learning rate of 0.1 (Figure 6-14).	**10 and 15** with a learning rate of 0.1 (Figure 6-15).	**10 and 15** with a learning rate of 0.1 (Figure 6-16).

Gradient descent on the *fitted and eliminated feature* model's cost function for the other learning rates (0.03, 0.01 and 0.0037) decreased and saturated at a cost higher than that of the 0.1 learning rate for all the DG penetration scenarios.

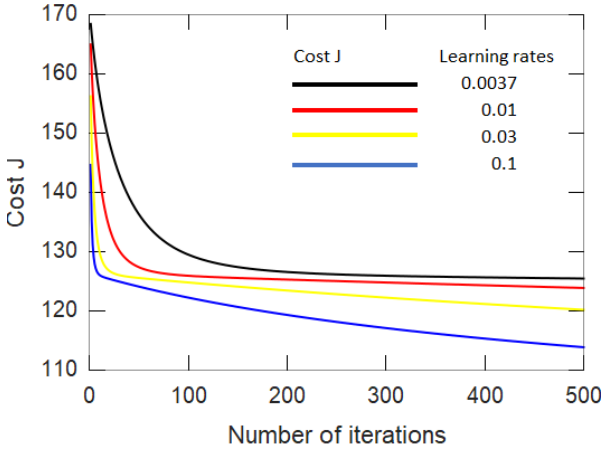


Figure 6-13: Convergence of the feature elimination and second-order polynomials as inputs model's gradient descent before DG penetration for four different learning rates

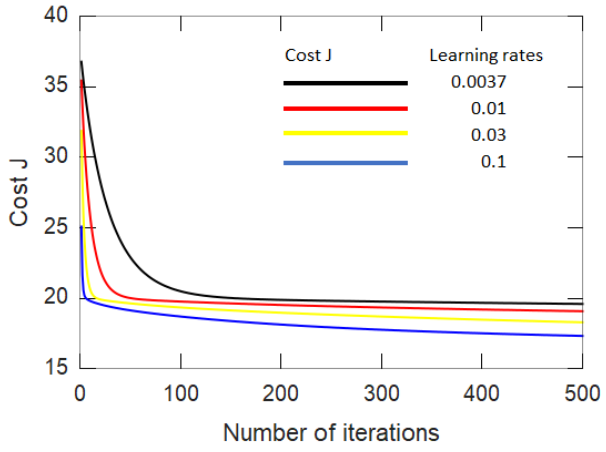


Figure 6-14: Convergence of the feature elimination and second-order polynomials as inputs model's gradient descent after 35% of DG penetration for four different learning rates

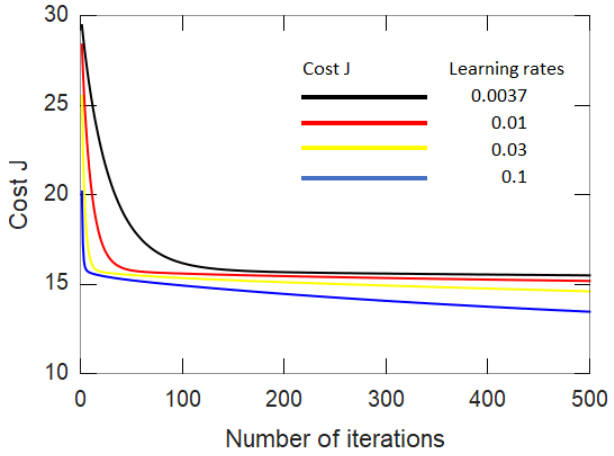


Figure 6-15: Convergence of the feature elimination and second-order polynomials as inputs model's gradient descent after 65% of DG penetration for four different learning rates

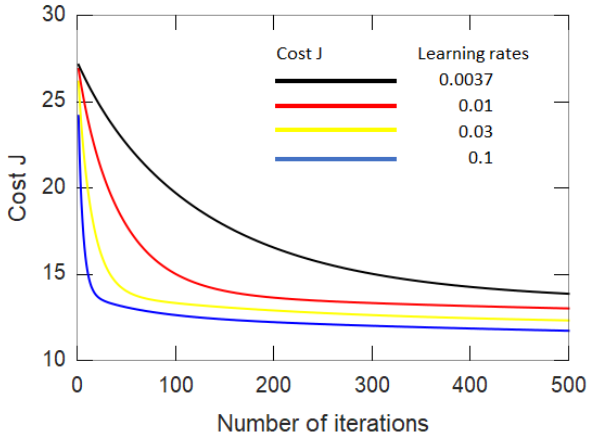


Figure 6-16: Convergence of the feature elimination and second-order polynomials as inputs model's gradient descent after 75% of DG penetration for four different learning rates

6.3. Linear and polynomial regression debugging: learning curves

To understand the shortcomings of regression models and to select the best model, learning curves are used. As detailed in appendix B.1.2c), the relationship between the cross-validation and training error functions indicate whether the model is suffering from a high bias or high variance problem. The ideal state of desired performance for a regression model is an error equal to zero as the number of training examples are increased. Not all models can achieve the ideal state of desired performance, therefore, an error of less than 0.3±0.1 was selected as the desired performance for all the regression models in this section. To aid in the selection of the best performing model, learning curves were plotted for each model i.e. all features, fitting (adding polynomials), eliminating features, adding polynomials and eliminating other features.

6.3.1. All features as inputs

As introduced in section 6.2.1, five features (network protection parameters) were inputs to the hypothesis function with one output: which makes visualisation of the inputs in relation to the output difficult. Figure 6-17 shows a 2D graph of one of the features (normalised pick-up current settings for multiple relays) in relation to the output (actual relay operating time before DG penetration). As evident in Figure 6-17, a normal linear equation doesn't fit this data.

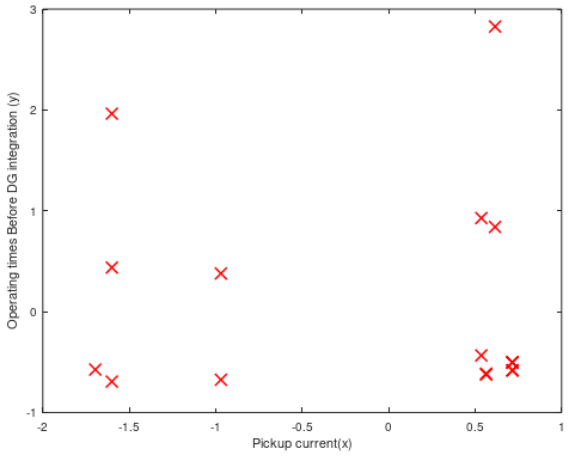


Figure 6-17: Normalised pick-up current settings vs operating time before DG penetration

As introduced in section 6.3, the relationship between the cross-validation and training error functions indicate whether the model is suffering from a high bias or high variance problem. A high variance problem is a situation where the cross-validation error is high and training error is low (introduced in detail in appendix B.1.2c). Appendix B.1.2a) also introduced the regularization parameter as a magnitude that reduces the contribution of high-order polynomials to achieve a "best fit". Table 6-9 shows a summary of regression and regularization learning curves for the *all feature as inputs* model on four scenarios (before and after DG penetration).

Table 6-9: A summary of regression and regularization learning curves for the all feature as inputs model

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Linear regression learning curve			
Indicated a high variance problem after 7 training examples. However, after 17 training examples, the cross-validation and training errors approached the desired performance of less than 0.3±0.1 as illustrated in Figure 6-18 and Table F-1 in appendix F.	Indicated a high variance problem (Figure F-18 in appendix F).	Indicated a high variance problem (Figure F-22 in appendix F).	Indicated a high variance problem initially. However, after 11 training examples, the cross-validation and training errors approached the desired performance as illustrated in Figure 6-21 and Table F-10 in appendix F.
8th-degree polynomial regression learning curve			
Indicated a high variance problem (Figure 6-19)	Indicated a high variance (Figure F-19 in appendix F).	Indicated a high variance (Figure F-23 in appendix F).	The cross-validation and training errors approached the desired performance of less than 0.3±0.1 (Figure 6-22 and Table F-11 in appendix F).
Regularization parameter learning curve			
Indicated an optimal value of 40, where the training and cross-validation error are equal (Figure 6-20).	Indicated an optimal value of 90 (Figure F-20 in appendix F).	Indicated an optimal value beyond 100 (Figure F-24 in appendix F)	Indicated an optimal value beyond 100 (Figure 6-23).

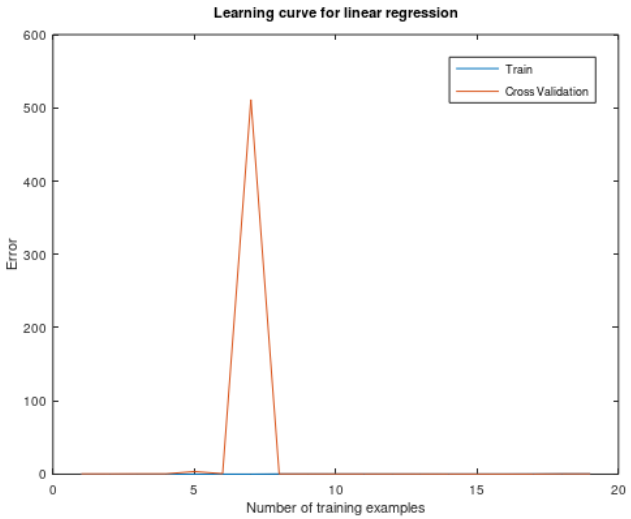


Figure 6-18: All features as inputs model linear regression learning curves (relay operating time before DG penetration)

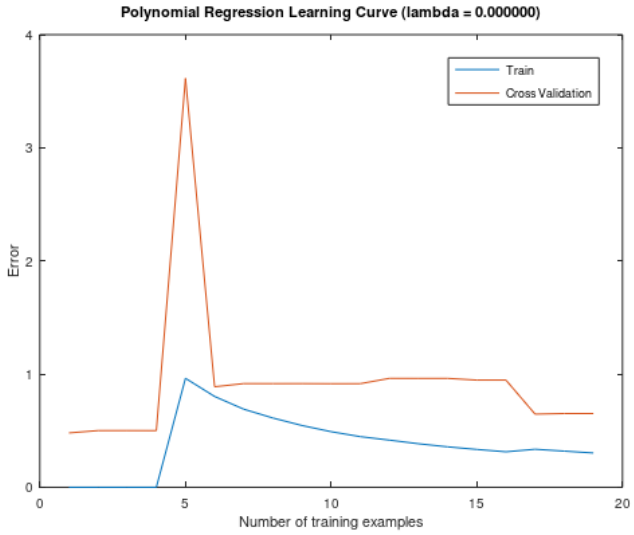


Figure 6-19: All-feature model polynomial regression learning curves (relay operating time before DG penetration)

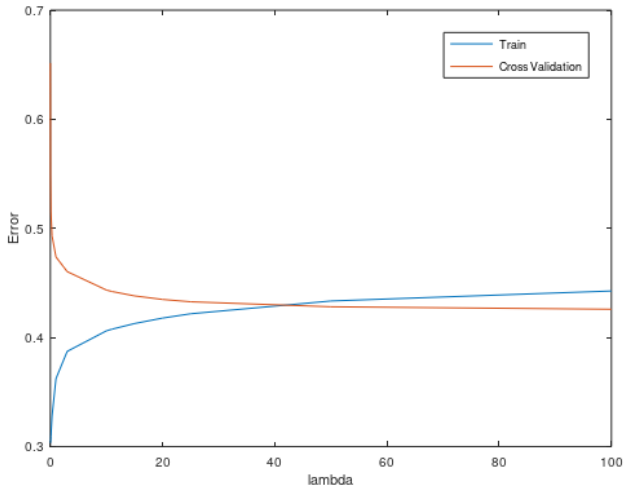


Figure 6-20: All-feature model regularization learning curves (relay operating time before DG penetration)

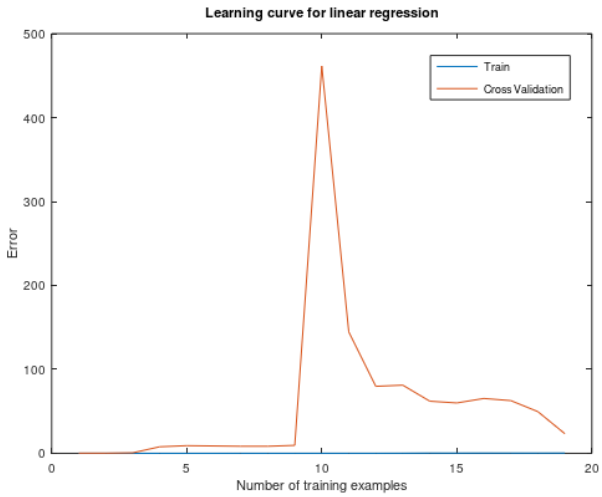


Figure 6-21: All features as inputs model linear regression learning curves (relay operating time after 75% of DG penetration)

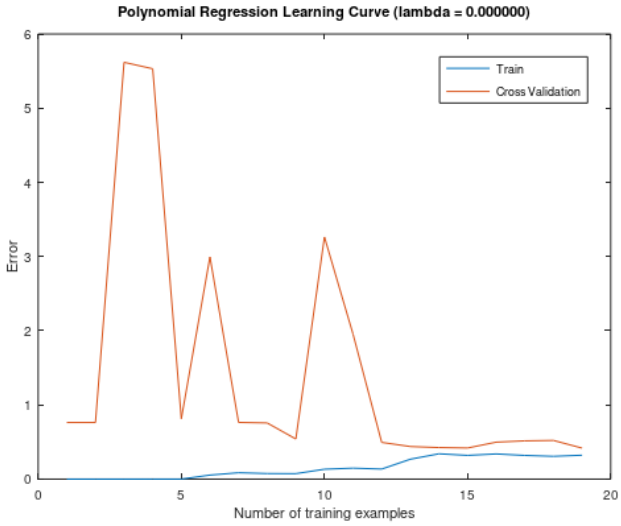


Figure 6-22: All-feature model polynomial regression learning curves (relay operating time after 75% DG penetration)

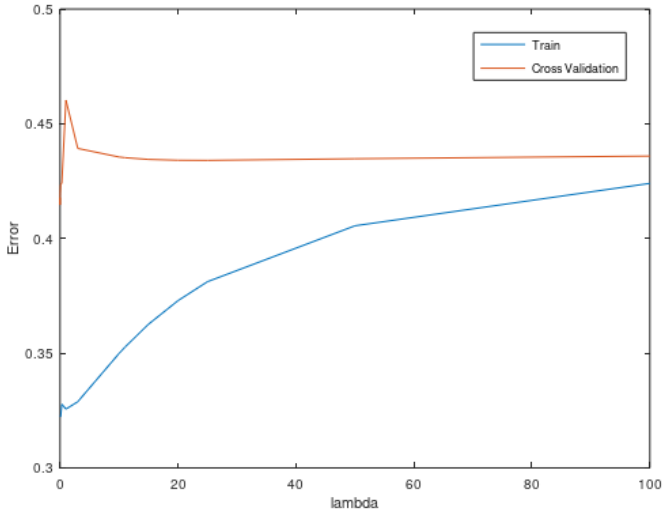


Figure 6-23: All-feature model regularization learning curves (relay operating time after 75% DG penetration)

6.3.2. Fitting (polynomials as added inputs)

As introduced in section 6.2.2, four additional features were added to the original five features as inputs to the hypothesis function. Table 6-10 shows a summary of regression and regularization learning curves for the *polynomials as added inputs* model on four scenarios (before and after DG penetration).

Table 6-10: A summary of regression and regularization learning curves for the polynomials as added inputs model

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Linear regression learning curve			
Indicated a high variance problem (Figure F-26 in appendix F).	Indicated a high variance problem initially. However, after 17 training examples, the cross-validation and training errors approached the desired performance of less than 0.3±0.1 as illustrated in Figure 6-24 and Table F-16 in appendix F).	Indicated a high variance problem initially. However, after 10 training examples, the cross-validation and training error approached the desired performance of less than 0.3±0.1 as illustrated in Figure F-31 and Table F-19 in appendix F).	Indicated a high variance problem (Figure F-34 in appendix F).
8th-degree polynomial regression learning curve			
Indicated a high variance problem (Figure F-27 in appendix F).	The cross-validation and training errors approached the desired performance of less than 0.3±0.1(Figure 6-25 and Table F-17 in appendix F).	Indicated a high variance (Figure F-32 in appendix F).	The cross-validation and training errors approached the desired performance of less than 0.3±0.1 (Figure F-35 and Table F-23 in appendix F).
Regularization parameter learning curve			
Indicated an optimal value of 0.003, where the training and cross-validation error are equal (Figure F-28 and Table F-15 in appendix F).	Indicated an optimal value of 0.3 (Figure 6-26 and Table F-18 in appendix F).	Indicated an optimal value of 15 (Figure F-33 in appendix F)	Indicated an optimal value beyond 50 (Figure F-36 in appendix F)

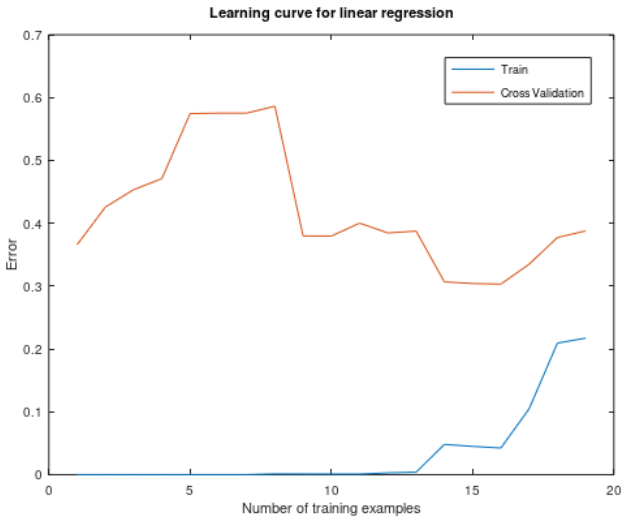


Figure 6-24: Fitted model linear regression learning curves (relay operating time after 35% DG penetration)

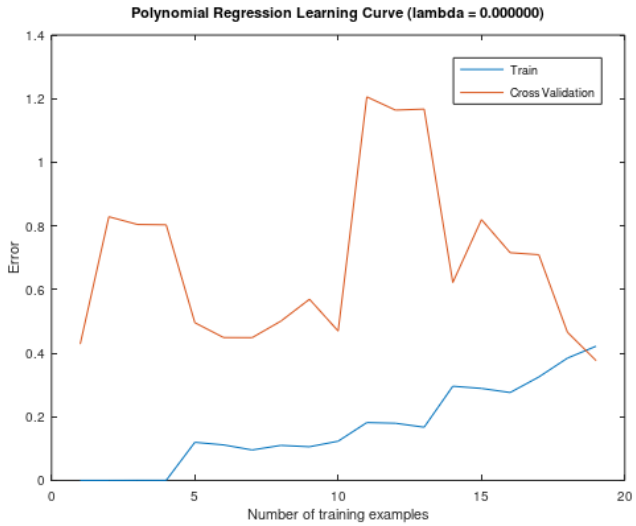


Figure 6-25: Fitted model polynomial regression learning curves (relay operating time after 35% DG penetration)

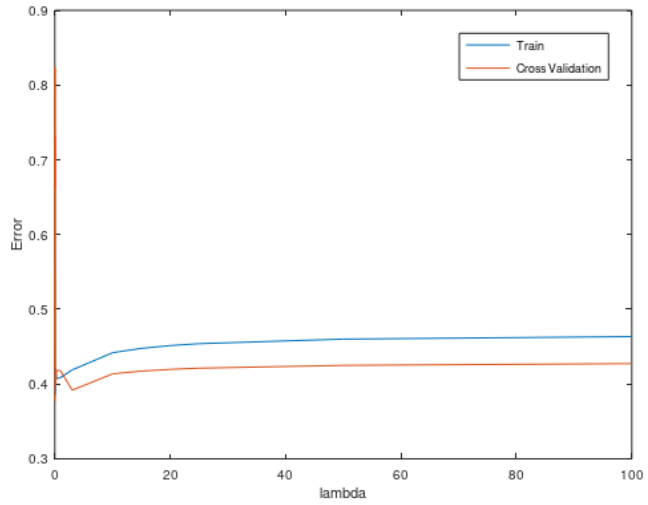


Figure 6-26: Fitted model regularization learning curves (relay operating time after 35% DG penetration)

6.3.3. Eliminating features

As introduced in section 6.2.3, relay positions were eliminated leaving four features used as inputs to the hypothesis function. Table 6-11 shows a summary of regression and regularization learning curves for the *eliminated features* model on four scenarios (before and after DG penetration).

Table 6-11: A summary of regression and regularization learning curves for the eliminated features model

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Linear regression learning curve			
Indicated a high variance problem (Figure F-37 in appendix F).	Indicated a high variance problem initially. However, after 15 training examples, the cross-validation and training error approached the desired performance (Figure 6-27)	Indicated a high variance problem (Figure F-40 in appendix F).	Indicated a high variance problem (Figure F-42 in appendix F).
8th-degree polynomial regression learning curve			
Indicated a high variance problem (Figure F-38 in appendix F).	Indicated a high variance problem (Figure F-29 in appendix F).	The cross-validation and training errors approached the desired performance of less than 0.3±0.1 (Figure 6-28).	Indicated a high variance problem (Figure F-43 in appendix F).
Regularization parameter learning curve			
Indicated an optimal value of 25, where the training and cross-validation error are equal (Figure F-39 in appendix F).	Indicated an optimal value of 50 (Figure F-30 and Table F-30 in appendix F).	Indicated an optimal value beyond 100 (Figure F-41 in appendix F)	Indicated an optimal value of 3 (Figure 6-29 and Table F-36 in appendix F)

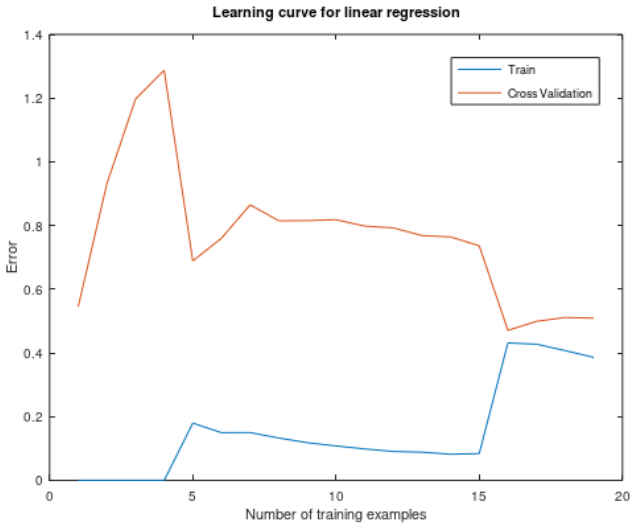


Figure 6-27: Feature elimination model linear regression learning curves (relay operating time after 35% DG penetration)

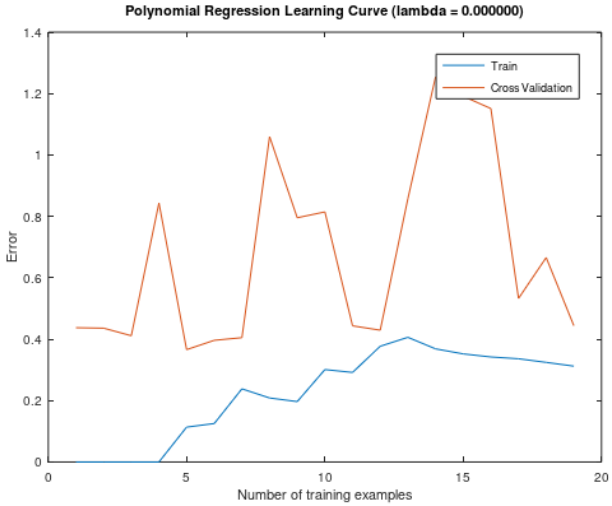


Figure 6-28: Feature elimination model polynomial regression learning curves (relay operating time after 65% DG penetration)

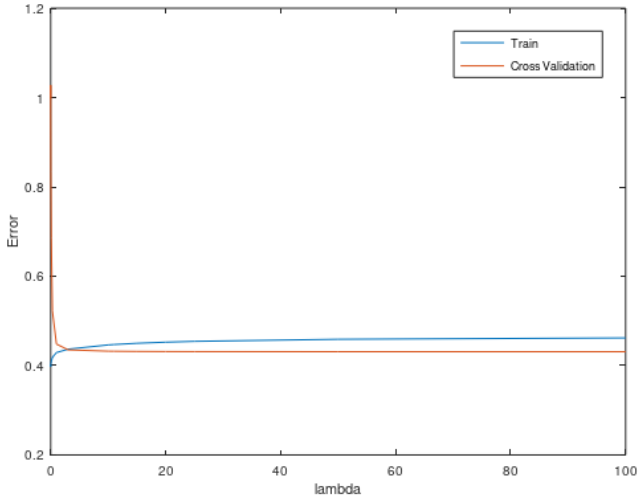


Figure 6-29: Feature elimination model regularization learning curves (relay operating time after 75% DG penetration)

6.3.4. Fitting and eliminating features

As introduced in section 6.2.4, relay positions were eliminated with four second-degree polynomials as additional inputs to the remaining four features were used in the hypothesis function. Table 6-12 shows a summary of regression and regularization learning curves for the *fitted and eliminated features* model on four scenarios (before and after DG penetration).

Table 6-12: A summary of regression and regularization learning curves for the fitted and eliminated features model

Predicting relay operating times:			
Before DG penetration	**After DG penetration of 35% of the ADMD**	**After DG penetration of 65% of the ADMD**	**After DG penetration of 75% of the ADMD**
Linear regression learning curve			
Indicated a high variance problem (Figure F-44 in appendix F).	Indicated a high variance problem initially. However, after 9, training examples, the cross-validation and training error approached the desired performance of less than 0.3±0.1 (Figure F-47 and Table F-40 in appendix F).	Indicated a high variance problem initially. However, after 12 training examples, the cross-validation and training error approached the desired performance of less than 0.3±0.1 (Figure F-50 and Table F-43 in appendix F).	Indicated a high variance problem initially. However, after training examples, the cross-validation and training error approached the desired performance of less than 0.3±0.1 (Figure 6-30 and Table F-46 in appendix F).
8^{th}-degree polynomial regression learning curve			
Indicated a high variance problem (Figure F-45 in appendix F).	Indicated a high variance problem (Figure F-48 in appendix F).	Indicated a high variance problem (Figure F-51 in appendix F).	Indicated a high variance problem (Figure 6-31).
Regularization parameter learning curve			
Indicated an optimal value beyond 100 (Figure F-46 in appendix F).	Indicated an optimal value beyond 100 (Figure F-49 in appendix F).	Indicated an optimal value beyond 100 (Figure F-52 in appendix F).	Indicated an optimal value beyond 100 (Figure 6-32).

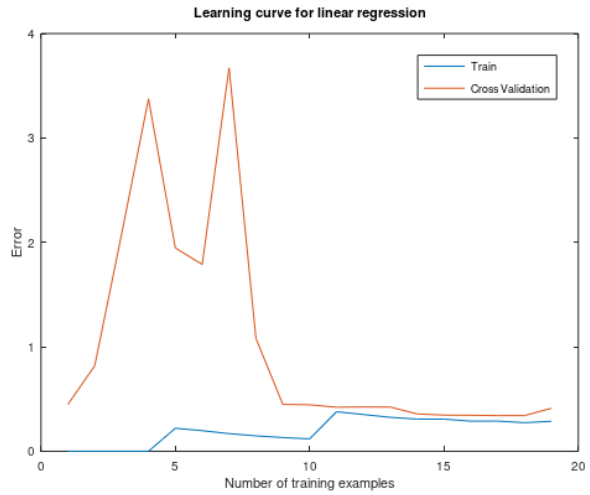


Figure 6-30: Feature elimination and fitted model linear regression learning curves (relay operating time after 75% DG penetration)

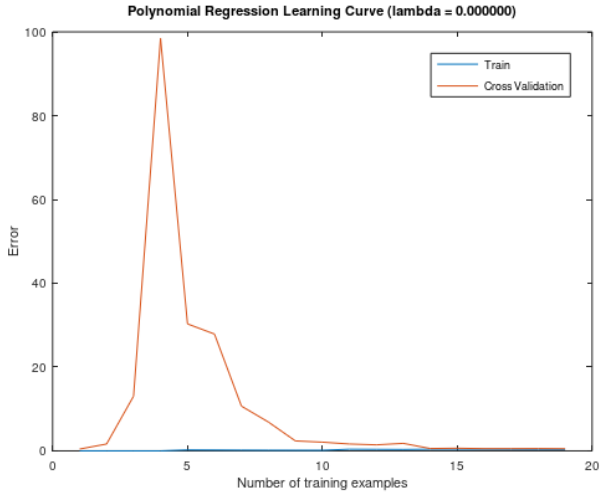


Figure 6-31: Feature elimination and fitted model polynomial regression learning curves (relay operating time after 75% DG penetration)

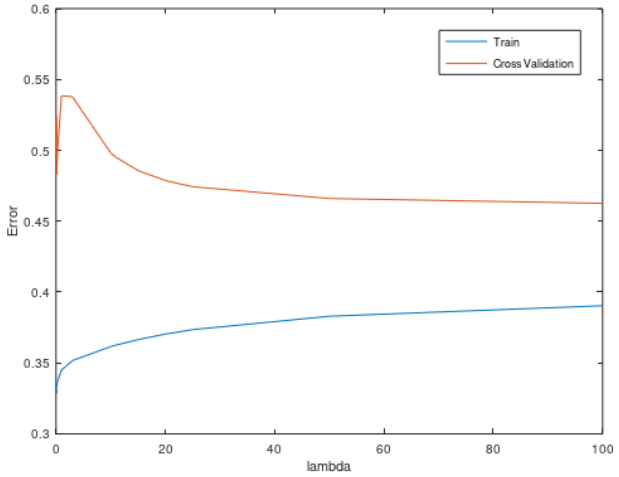


Figure 6-32: Feature elimination and fitted model regularization learning curves (relay operating time after 75% DG penetration)

6.4. Logistic regression as a protection miscoordination classifier

The aim was to train a model that can determine from the current network parameters and penetration levels which class of protection miscoordination is likely to occur.

6.4.1. Loss of coordination

In classifying loss of coordination, it is safer to predict a false positive, therefore, the decision boundary was set to true (output y equal to 1) when the sigmoid function (h(x)) was greater than or equal to 0.7 and false (output y equal to 0) when the sigmoid function was less than 0.7. As described in appendix B1.2a), the regularization parameter (lambda) determines the magnitude in which the theta parameter (Θ) is magnified to reduce the contribution of high-order polynomials. Given the number of high-order polynomials in the loss of coordination models, the regularization parameter was set to1 and 10. The accuracy of the model was evaluated under the following conditions: all features as inputs, fitting (polynomials as additional inputs), eliminating features, fitting and eliminating features (Table 6-13).

Table 6-13: A summary of the training and validation accuracies for the loss of coordination models

When classifying loss of coordination using			
All the features as inputs	**All the features with second-degree polynomials as additional inputs**	**Some of the features as inputs (relay positions eliminated)**	**Some of the features (relay positions eliminated) with second-degree polynomials as added inputs**
The data in Table 4-14, results from the detailed analysis on the impact of PV DG on the power protection systems [22] [37], and protection coordination issue triggers (appendix E.1.iv) were combined and split into:			
Nine features as inputs and a binary output (loss of coordination) as shown in Table 6-14.	Thirteen features as and a binary output as shown in Table 6-15.	Eight features as inputs and a binary output as shown in Table 6-16.	Twelve features as inputs and a binary output as shown in Table 6-17.
Evaluated with the regularization parameter set to 10.			
Training accuracy of **95.65%** and validation accuracy of **88.89%** (Figure F-53 in appendix F).	Training accuracy of **91.30%** and validation accuracy of **100%** (Figure F-55 in appendix F).	Training accuracy of **95.65%** and validation accuracy of **88.89%** (Figure F-54 in appendix F).	Training accuracy of **100%** and a cross-validation accuracy of **77.78%** (Figure F-56 in appendix F).

Table 6-14: Inputs (recloser number, TMS, fault level, pickup current, calculated operating time before and after DG penetration) and an output (loss of coordination binary output) for the all features as inputs model (sample)

Features (inputs)									**Output**
Network parameters					**Actual relay operating time (s) before DG penetration**	**Relay operating times (s) after DG penetration of**			**Loss of coordination**
Relay/ recloser	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Calculated operating time (s)**		**35% of ADMD**	**65% of ADMD**	**75% of ADMD**	
6	0.11	587	147.86	0.50	6.98	0.46	0.46	0.46	0
6/1	0.43	587	28.85	0.30	1.14	0.32	0.32	0.32	0
2	0.25	749	152.10	0.86	80.11	0.85	0.85	0.64	1
2/1	1.03	749	33.94	0.66	61.48	0.75	0.75	0.75	1
1	0.67	1817	149.34	0.81	10.54	1.02	1.03	1.09	0
1/1	1.17	1817	67.56	0.61	4.87	0.82	0.82	0.82	0
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-15: Inputs (second-degree polynomials-TMS, fault level, pickup current and calculated operating time) and an output (loss of coordination binary output) for the second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-14	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of coordination
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	0.46	0.46	0.46	0
	0.09	0.19	344569	832.05	1.14	0.32	0.32	0.32	0
	0.74	0.06	561001	23134.16	80.11	0.85	0.85	0.64	1
	0.44	1.06	561001	1151.94	61.48	0.75	0.75	0.75	1
	0.66	0.45	3301489	22303.17	10.54	1.02	1.03	1.09	0
	0.37	1.37	3301489	4564.75	4.87	0.82	0.82	0.82	0
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

Table 6-16: Inputs (recloser number eliminated) and an output (loss of coordination binary output) for the feature eliminated model (sample)

Features (inputs)								Output
Network parameters				Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of coordination
TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)		35% of ADMD	65% of ADMD	75% of ADMD	
0.11	587	147.86	0.50	6.98	0.46	0.46	0.46	0
0.43	587	28.85	0.30	1.14	0.32	0.32	0.32	0
0.25	749	152.10	0.86	80.11	0.85	0.85	0.64	1
1.03	749	33.94	0.66	61.48	0.75	0.75	0.75	1
0.67	1817	149.34	0.81	10.54	1.02	1.03	1.09	0
1.17	1817	67.56	0.61	4.87	0.82	0.82	0.82	0
0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-17: Inputs (second-degree polynomials and relay position eliminated) and an output (loss of coordination binary output) for the relay positions eliminated with second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-16	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of coordination
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	0.46	0.46	0.46	0
	0.09	0.19	344569	832.05	1.14	0.32	0.32	0.32	0
	0.74	0.06	561001	23134.16	80.11	0.85	0.85	0.64	1
	0.44	1.06	561001	1151.94	61.48	0.75	0.75	0.75	1
	0.66	0.45	3301489	22303.17	10.54	1.02	1.03	1.09	0
	0.37	1.37	3301489	4564.75	4.87	0.82	0.82	0.82	0
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

6.4.2. Loss of grading

In classifying loss of grading, the decision boundary was set to true (output y equal to 1) when the sigmoid function (h(x)) was greater than or equal to 0.7 and false (output y equal to 0) when the sigmoid function was less than 0.7. The accuracy of the model was evaluated under the following conditions: all features as inputs, fitting (polynomials as added inputs), eliminating features, fitting and eliminating features (Table 6-18).

Table 6-18: A summary of the training and validation accuracies for the loss of grading models

When classifying loss of grading using			
All the features as inputs	**All the features with second-degree polynomials as additional inputs**	**Some of the features as inputs (relay positions eliminated)**	**Some of the features (relay positions eliminated) with second-degree polynomials as added inputs**
The data in Table 4-14, results from the detailed analysis on the impact of PV DG on the power protection systems [22] [37], and protection coordination issue triggers (appendix E.1.iv) were combined and split into:			
Nine features as inputs and a binary output (loss of grading) as shown in Table 6-19.	Thirteen features as and a binary output as shown in Table 6-20.	Eight features as inputs and a binary output as shown in Table 6-21.	Twelve features as inputs and a binary output as shown in Table 6-22.
Evaluated with the regularization parameter set to 10			
Training accuracy of **78.26%** and validation accuracy of **22.22%** (Figure F-57 in appendix F).	Training accuracy of **69.56%** and validation accuracy of **33.33%** (Figure F-59 in appendix F).	Training accuracy of **73.91%** and validation accuracy of **55.56%** (Figure F-58 in appendix F).	Training accuracy of **78.26%** and validation accuracy of **22.22%** (Figure F-60 in appendix F).

Table 6-19: Inputs (recloser number, TMS, fault level, pickup current calculated, operating time before and after DG penetration) and an output (loss of grading binary output) for the all features as inputs model (sample)

Features (inputs)									Output
Network parameters					Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of grading
Relay/ recloser	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)		35% of ADMD	65% of ADMD	75% of ADMD	
6	0.11	587	147.86	0.50	6.98	10.73	14.48	21.97	1
6/1	0.43	587	28.85	0.30	1.14	1.75	2.37	3.59	1
2	0.25	749	152.10	0.86	80.11	18.51	15.05	9.78	1
2/1	1.03	749	33.94	0.66	61.48	14.21	11.55	7.51	1
1	0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81	1
1/1	1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65	1
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-20: Inputs (second-degree polynomials-TMS, fault level, pickup current and calculated operating time) and an output (loss of grading binary output) for the second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-19	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of grading
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97	1
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59	1
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78	1
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51	1
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81	1
	0.37	1.37	3301489	4564.75	4.87	5.62	4.90	4.65	1
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

Table 6-21: Inputs (recloser number eliminated) and an output (loss of grading binary output) for the feature eliminated model (sample)

Features (inputs)								Output
Network parameters				Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of grading
TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)		35% of ADMD	65% of ADMD	75% of ADMD	
0.11	587	147.86	0.50	6.98	10.73	14.48	21.97	1
0.43	587	28.85	0.30	1.14	1.75	2.37	3.59	1
0.25	749	152.10	0.86	80.11	18.51	15.05	9.78	1
1.03	749	33.94	0.66	61.48	14.21	11.55	7.51	1
0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81	1
1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65	1
0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-22: Inputs (second-degree polynomials and relay position eliminated) and an output (loss of grading binary output) for the relay positions eliminated with second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-21	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of grading
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97	1
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59	1
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78	1
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51	1
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81	1
	0.37	1.37	3301489	4564.75	4.87	5.62	4.90	4.65	1
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

6.4.3. De-sensitization

In classifying de-sensitization, the decision boundary was also set to true (output y equal to 1) when the sigmoid function (h(x)) was greater than or equal to 0.7 and false (output y equal to 0) when the sigmoid function was less than 0.7. The accuracy of the model was evaluated under the following conditions: all features as inputs, fitting (polynomials as added inputs), eliminating features, fitting and eliminating features (Table 6-23).

Table 6-23: A summary of the training and validation accuracies for the de-sensitization models

When classifying de-sensitization using			
All the features as inputs	**All the features with second-degree polynomials as additional inputs**	**Some of the features as inputs (relay positions eliminated)**	**Some of the features (relay positions eliminated) with second-degree polynomials as added inputs**
The data in Table 4-14, results from the detailed analysis on the impact of PV DG on the power protection systems [22] [37], and protection coordination issue triggers (appendix E.1.iv) were combined and split into:			
Nine features as inputs and a binary output (de-sensitization) as shown in Table *6-24*.	Thirteen features as and a binary output as shown in Table 6-25.	Eight features as inputs and a binary output as shown in Table 6-21.	Twelve features as inputs and a binary output as shown in Table 6-27.
Evaluated with the regularization parameter set to 10			
Training accuracy of **91.30%** and validation accuracy of **33.33%** (Figure F-61 in appendix F).	Training accuracy of **82.61%** and validation accuracy of **66.67%** (Figure F-63 in appendix F).	Training accuracy of **86.96%** and validation accuracy of **44.44%** (Figure F-62 in appendix F).	Training accuracy of **73.91%** and validation accuracy of **66.67%** (Figure F-64 in appendix F).

Table 6-24: Inputs (recloser number, TMS, fault level, pickup current calculated, operating time before and after DG penetration) and an output (de-sensitization binary output) for the all features as inputs model (sample)

Features (inputs)									**Output**
Network parameters					**Actual relay operating time (s) before DG penetration**	**Relay operating times (s) after DG penetration of**			**De-sensitization**
Relay/ recloser	**TMS**	**Fault level (A)**	**Pick-up current setting (A)**	**Calculated operating time (s)**		**35% of ADMD**	**65% of ADMD**	**75% of ADMD**	
6	0.11	587	147.86	0.50	6.98	10.73	14.48	21.97	1
6/1	0.43	587	28.85	0.30	1.14	1.75	2.37	3.59	1
2	0.25	749	152.10	0.86	80.11	18.51	15.05	9.78	0
2/1	1.03	749	33.94	0.66	61.48	14.21	11.55	7.51	0
1	0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81	0
1/1	1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65	0
CB	0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
Trf	1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-25: Inputs (second-degree polynomials-TMS, fault level, pickup current and calculated operating time) and an output (de-sensitization binary output) for the second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-24	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			De-sensitization
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97	1
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59	1
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78	0
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51	0
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81	0
	0.37	1.37	3301489	4564.75	4.87	5.62	4.90	4.65	0
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

Table 6-26: Inputs (recloser number eliminated) and an output (de-sensitization binary output) for the feature eliminated model (sample)

Features (inputs)								Output
Network parameters				Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			De-sensitization
TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)		35% of ADMD	65% of ADMD	75% of ADMD	
0.11	587	147.86	0.50	6.98	10.73	14.48	21.97	1
0.43	587	28.85	0.30	1.14	1.75	2.37	3.59	1
0.25	749	152.10	0.86	80.11	18.51	15.05	9.78	0
1.03	749	33.94	0.66	61.48	14.21	11.55	7.51	0
0.67	1817	149.34	0.81	10.54	13.43	10.64	9.81	0
1.17	1817	67.56	0.61	4.87	5.62	4.90	4.65	0
0.92	1817	157.46	1.18	1.18	1.18	1.18	1.18	0
1.15	1817	157.46	1.48	1.48	1.48	1.48	1.48	0

Table 6-27: Inputs (second-degree polynomials and relay position eliminated) and an output (de-sensitization binary output) for the relay positions eliminated with second-degree polynomials as additional inputs model (sample)

Features (inputs)									Output
Network parameters in Table 6-26	$Calculated\ operating\ time^2$	TMS^2	$Fault\ Level^2$	$Pick\ up\ current^2$	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			De-sensitization
						35% of ADMD	65% of ADMD	75% of ADMD	
	0.25	0.01	344569	21862.27	6.98	10.73	14.48	21.97	1
	0.09	0.19	344569	832.05	1.14	1.75	2.37	3.59	1
	0.74	0.06	561001	23134.16	80.11	18.51	15.05	9.78	0
	0.44	1.06	561001	1151.94	61.48	14.21	11.55	7.51	0
	0.66	0.45	3301489	22303.17	10.54	13.43	10.64	9.81	0
	0.37	1.37	3301489	4564.75	4.87	5.62	4.90	4.65	0
	1.38	0.84	3301489	24793.34	1.18	1.18	1.18	1.18	0
	2.18	1.33	3301489	24793.34	1.48	1.48	1.48	1.48	0

6.5. Chapter Discussion

This chapter presented the trainings of models that can predict relay operating times at certain DG penetration levels, from the current network parameters and the determination of which class of protection miscoordination is likely to occur. The success criteria were based on each model's cost function minimization value within the first 100 iterations, the cross-validation and training error approaching desired performance of less than 0.3±0.1 on the learning curves.

When predicting relay operating times before DG penetration from **the model that uses all the features as inputs** (section 6.2.1), the cost function was minimized to a value below **120** $seconds^2$ (Figure 6-1). Learning curves for this model are shown in Figure 6-18 and Table F-1 in appendix F.

When predicting relay operating times at 35% of PV DG penetration from **the model of fitting** (**adding polynomials** (section 6.2.2)), the cost function was minimized to a value below **20** $seconds^2$ (Figure 6-6). Learning curves for this model are shown in Figure 6-24 and Table F-16 in appendix F.

When predicting relay operating times at 65% of PV DG penetration from **the model of fitting** (**adding polynomials** (section 6.2.2)), the cost function was minimized to a value below **15** $seconds^2$ (Figure 6-7). Learning curves for this model are shown in Figure F-31 and Table F-19 in appendix F.

When predicting operating time at 75% of PV DG penetration from **the model that uses all the features** (section 6.2.1), the cost function was minimized to a value below **10** $seconds^2$ (Figure 6-4). From the learning curves, the cross-validation and training error indicated a high variance problem (Figure 6-21). However, from the **fitted polynomials and feature elimination** model learning curves, the cross-validation and training error approached the desired performance as illustrated in Figure 6-30 and Table F-46 in appendix F.

In section 1.2 b), the objective for the classifier model was defined as **accuracy and precision**: when a positive value is predicted, how often is the prediction correct. Because false positives (recloser/relay not experiencing protection miscoordination and classified as true) are more acceptable than false negatives (recloser/relay experiencing protection miscoordination and classified as false). Hence false negatives must be as low as possible. The performance metrics (Figure B-11) for each model (loss of coordination, loss of grading and de-sensitization) are shown in Figure 6-33 to Figure 6-38.

When classifying loss of coordination, the model with second-degree polynomials added (section 6.4.1 had a training accuracy of **91.30%** and validation accuracy of **100%.** Thus, out of twenty-three instances where loss of coordination was experienced in the training data set, the model was able to predict twenty-one instances correctly. The confusion matrices for the loss of coordination classifier are shown on the bottom right of Figure 6-33 and Figure 6-34.

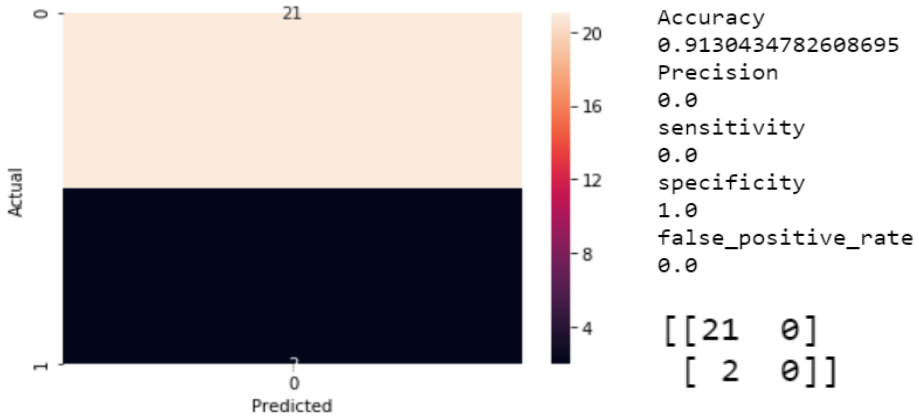


Figure 6-33: Loss of coordination classifier (fitted model) training set confusion matrix and performance metrics

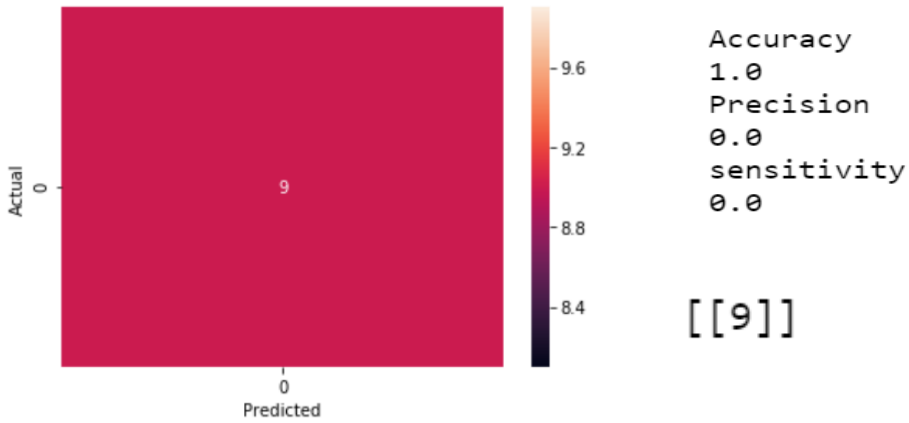


Figure 6-34: Loss of coordination classifier (fitted model) cross-validation data set confusion matrix and performance metrics

When classifying the loss of grading, the model with relay position eliminated (section 6.4.2 had a training accuracy of **73.91%** and validation accuracy of **55.56%.** Thus, out of seven instances where the loss of grading was experienced in the cross-validation data set, the model was able to predict three instances correctly. The confusion matrices for the loss of grading classifier are shown in Figure 6-35 and Figure 6-36.

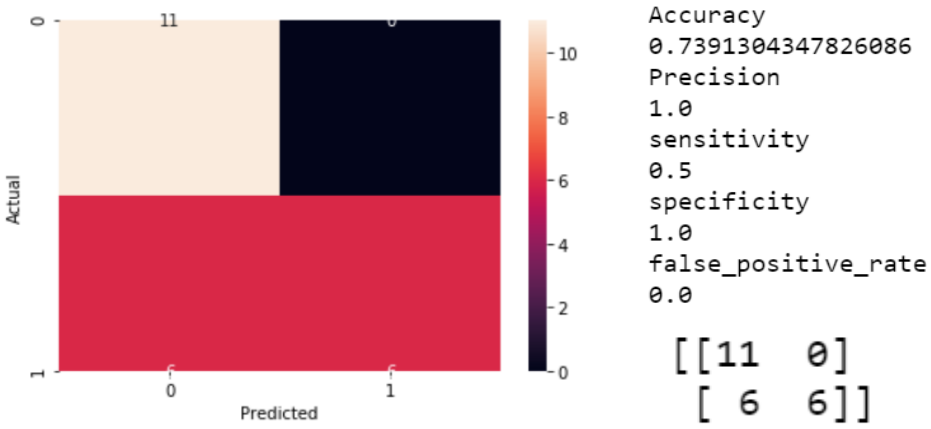


Figure 6-35: Loss of grading classifier (eliminated feature model) training data set confusion matrix and performance metrics

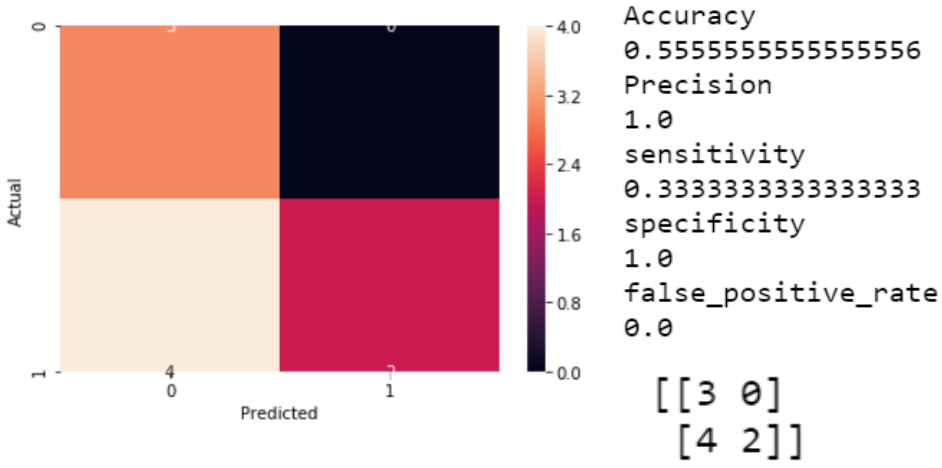


Figure 6-36: Loss of grading classifier (eliminated feature model) cross-validation set confusion matrix and performance metrics

When classifying de-sensitization, the model with second-degree polynomials as added inputs (section 6.4.3 had a training accuracy of **82.61%** and validation accuracy of **66.67%.** Thus, out of three instances where de-sensitization was not experienced in the training data set, the model was able to predict two instances correctly. The confusion matrices for the de-sensitization classifier are shown in Figure 6-37and Figure 6-38.

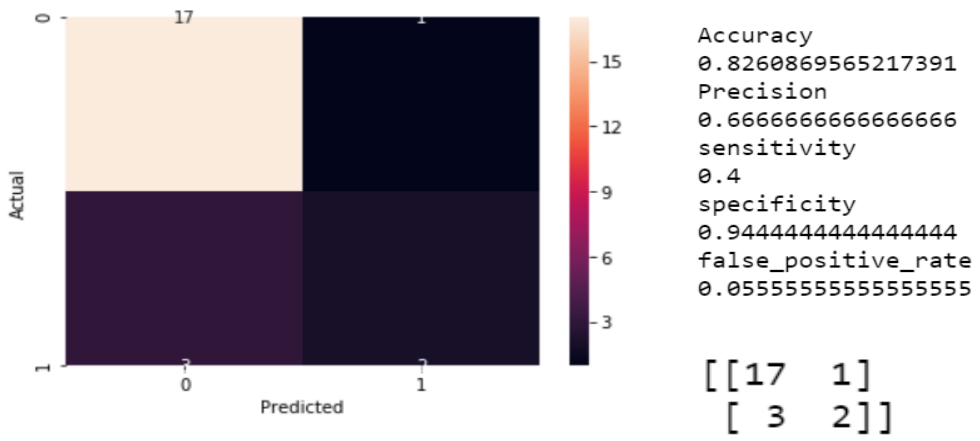


Figure 6-37: De-sensitization classifier (fitted model) training data set confusion matrix and performance metrics

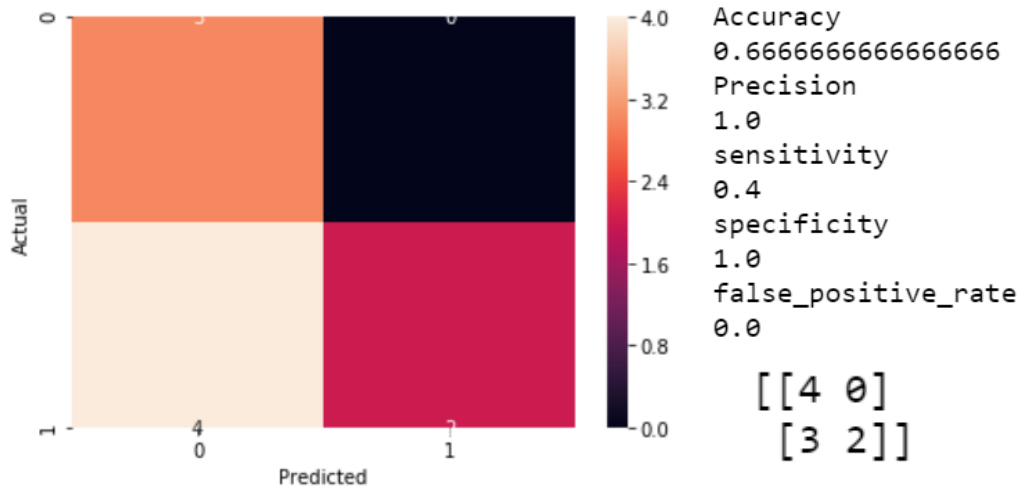


Figure 6-38: De-sensitization classifier (fitted model) cross-validation data set confusion matrix and performance metrics

7. Model application: results and analysis

This chapter presents the identification of the type of feeders that can handle a high penetration of DGs and the type of feeders that cannot. This was done by using the best regression model (section 6.5) to predict the operating times at three levels of DG penetration and the best classifier model to classify protection coordination issues that will occur on four different MV distribution networks (section 4.5). The results from the best regression model were used as inputs to the most accurate logistic regression model to classify which set of network protection features and DG penetration level results in which protection coordination problem. Table 7-1 shows the values of the predicted relay operating times and protection miscoordination groups after three stages of DG penetration.

*Table 7-1: Predicted relay operating times and protection miscoordination groups after three stages of DG penetration**

Relay	TMS	Fault level (A)	Pick up current	Calculated operating time (s)	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of coordination	De-sensitization	Loss grading
						35% of ADMD	65% of ADMD	75% of ADMD			
Trf	0.54	677.50	59.05	1.50	13.68	52.77	70.53	31.72	0	0	1
Trf	0.80	787.50	59.05	2.10	12.19	84.96	149.06	50.92	0	1	1
49/470	0.05	229.00	82.36	0.61	18.13	18.13	18.13	18.13	0	0	1
Trf	1.04	719.25	59.05	2.83	13.76	104.32	232.83	65.35	0	1	1

*filtered to show network protection features where protection miscoordination was predicted.

From Table 7-1, Table 4-5, Table 4-10, Table 4-11, and Table 4-12 the network protection features for which protection miscoordination was predicted and not predicted are grouped in Table 7-2, Table 7-3 and Table 7-4 respectively.

Table 7-2: Network number and protection features where the loss of grading was predicted

Network	Relay/ Recloser position	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)
Network B	Trf	0.80	787.50	59.05	2.10
Network C	Trf	1.04	719.25	59.05	2.83
Network A	Trf	0.54	677.5	59.05	1.5
Network B	49/470	0.05	229	82.36	0.61

Table 7-3: Network number and protection features where the loss of grading was not predicted

Network	Relay/ Recloser position	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)
Illustration network (Figure 4-1)	Trf	0.38	368.5	65.61	1.5

In Table 7-2 and Table 7-3, the network protection features indicate that loss of grading occurs on transformer circuit breakers where **0.5 ≤ TMS ≤ 1.1** and **650≤ Fault level ≤800A.** It is noted

that relay 49/470 is a point of interest as shown in Table 7-7 (relay operating time indicated protection miscoordination before DG penetration), thus a false positive. In Table 7-4 and Table G-1, the network protection features indicate that de-sensitization occurs where **0.6 ≤ TMS ≤ 1.1** and **710≤ Fault level ≤ 800A.** Network A's transformer circuit breaker TMS and fault level were within the de-sensitization range, however, the model did not predict de-sensitization for the network protection features (false negative). The loss of protection coordination groups introduced in section 2.3 are categorized based on the operating times of secondary relays in relation to primary relays. Relay operating times are calculated using the TMS and fault level values (section 4.2, equation 4-1). TMS is proportional to the relay operating time, therefore, it is chosen such that the relay operating time coordination with downstream devices is optimized. Also detailed in section 2.3 is the added fault current introduced in the distribution network by DG. Thus, TMS and fault level are accepted as the combination of protection features that determine the protection-based DG penetration limit.

Table 7-4: Network number and protection features that resulted in de-sensitization

Network	Relay/ Recloser position	TMS	Fault level (A)	Pick-up Current setting (A)	Calculated operating time (s)
Network B	Trf	0.80	787.5	59.05	2.10
Network C	Trf	1.04	719.25	59.05	2.83

As seen in Table 7-1 and Table G-1 (in appendix G), loss of coordination was not predicted for the MV distribution networks (network A, B, C and Illustration network Figure 4-1) using the horizontal approach (linear and logistic regression models). Based on the relay operating times after DG penetration, protection miscoordination (loss of grading and de-sensitization) occurs after a DG penetration of 35% of ADMD. Using the findings in section 5.9, the TMS and fault level range where protection miscoordination occurs, a PV DG Medium Voltage Opportunity Network (MVON) guide was defined. The MVON guide is for the DG penetration level, location and network protection features that determine which protection miscoordination group (section 2.3) will occur as shown in Table 7-5 and Table 7-6.

Table 7-5: PV DG Medium voltage opportunity guide for penetration levels

PV DG Penetration level \ Postal Range	1-299	1400-2199	6500-8099	2900-4730	8100-9299
35% of ADMD					x
65% of ADMD				x	
75% of ADMD	x	x	x		

Table 7-6: PV DG Medium Voltage opportunity network guide for network protection feature ranges for protection miscoordination groups

Protection miscoordination issue / Network Protection features	TMS	Fault level (A)	Relay/ Recloser position	Pick-up current Setting (A)	Calculated operating time (s)
Loss of grading	x	x			
De-sensitization	x	x	x		

A vertical analysis was done: the predicted operating time of each relay was compared to the operating time of the downstream and upstream relay. Table 7-7 shows the points of interest from the vertical analysis and the predicted operating times.

Table 7-7: Vertical analysis points of interest and outliers on predicted relay operating times, protection miscoordination groups before and after three stages of DG penetration

Relay	TMS	Fault level (A)	Pick-up current setting (A)	Calculated relay operating time (s) before DG penetration	Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss of coordination	De-sensitization	Loss grading
						35% of ADMD	65% of ADMD	75% of ADMD			
423	0.01	231.00	124.70	0.11	18.45	18.45	18.45	7.36	0	0	0
318	0.03	231.00	138.56	0.41	5.58	5.58	5.58	0.15	0	0	0
101	0.11	479.00	196.10	0.86	10.46	10.46	10.46	10.46	0	0	0
CB	0.13	909.00	299.20	0.80	6.14	6.14	6.14	6.14	0	0	0
219/1	0.05	437.00	141.54	0.31	6.23	6.23	6.23	5.95	0	0	0
87/1	0.05	1031.00	269.28	0.26	4.02	4.02	4.02	7.58	0	0	0
148	0.01	519.00	125.52	0.05	7.85	7.85	7.85	9.42	0	0	0
114	0.07	519.00	139.47	0.35	9.42	9.42	9.42	6.61	0	0	0
67	0.10	630.00	154.97	0.60	9.84	9.84	9.84	2.77	0	0	0
1	0.24	1204.00	191.31	0.89	3.36	9.90	4.97	10.90	0	0	0
CB	0.37	2710.00	212.57	1.00	0.99	30.50	31.95	25.17	0	0	0
111/1	0.05	555.00	139.47	0.25	8.86	8.86	8.86	7.79	0	0	0
65/176	0.05	264.00	91.51	0.33	16.16	16.16	16.16	5.31	0	0	0
65/127	0.12	328.00	112.97	0.81	15.55	15.55	15.55	15.55	0	0	0
65/67	0.15	330.00	125.52	1.10	15.81	15.81	15.81	15.81	0	0	0
65/1	0.26	888.00	154.97	1.02	8.98	12.96	8.50	9.40	0	0	0
48	0.17	901.00	172.18	1.31	7.55	7.55	7.55	7.55	0	0	0
65/39/58	0.06	379.00	82.36	0.27	14.43	14.43	14.43	8.33	0	0	0
65/39/56	0.12	379.00	91.51	0.57	14.67	14.67	14.67	5.79	0	0	0
65/39/54	0.17	393.00	101.67	0.86	14.69	14.69	14.69	2.91	0	0	0
65/39/2	0.21	398.00	112.97	1.15	14.87	4.76	14.87	14.87	0	0	0

65/39/1	0.26	582.00	125.52	1.18	11.72	12.79	8.26	4.74	0	0	0
65/2	0.19	587.00	139.47	1.47	15.05	0.29	15.05	15.05	0	0	0
65/39/39/1	0.08	435.00	101.67	0.38	13.40	13.40	13.40	7.29	0	0	0
65/39/53/1	0.08	398.00	91.51	0.38	14.13	14.13	14.13	7.43	0	0	0
297	0.07	388.00	82.36	0.31	4.72	4.72	4.72	8.25	0	0	0
174	0.12	413.00	101.67	1.05	10.52	10.52	10.52	10.52	0	0	0
123	0.15	622.00	112.97	0.74	8.48	8.48	8.48	4.57	0	0	0
75	0.13	829.00	125.52	0.82	7.45	7.45	7.45	7.45	0	0	0
65	0.16	1208.00	139.47	0.68	1.49	1.49	1.49	7.14	0	0	0
50	0.24	1335.00	154.97	0.77	0.60	10.45	5.63	15.28	0	0	0
2	0.18	1833.00	172.18	0.97	0.95	1.13	0.95	0.94	0	0	0
1	0.30	3510.00	191.31	0.71	0.50	19.59	16.97	25.25	0	0	0
2/2	0.05	3510.00	154.97	0.11	0.01	0.01	0.01	12.32	0	0	0
74/21	0.05	490.00	101.67	0.22	11.98	11.98	11.98	8.78	0	0	0
74/9	0.14	292.00	125.52	1.17	19.28	19.28	19.28	19.28	0	0	0
140/6	0.05	393.00	91.51	0.24	10.92	10.92	10.92	8.21	0	0	0
140/1	0.11	393.00	101.67	0.54	11.15	11.15	11.15	5.34	0	0	0
220/6/1	0.05	370.00	82.36	0.23	8.09	8.09	8.09	8.44	0	0	0
259/5/1	0.05	383.00	82.36	0.22	6.26	6.26	6.26	8.61	0	0	0
280/8/1	0.05	358.00	74.12	0.22	5.90	5.90	5.90	8.78	0	0	0
49/373	0.09	236.00	91.51	0.63	17.75	17.75	17.75	0.00	0	0	0
49/250	0.11	387.00	112.97	0.94	15.68	15.68	15.68	15.68	0	0	0
49/148	0.12	400.00	125.52	0.90	15.14	15.14	15.14	15.14	0	0	0
49/79	0.12	590.00	139.47	0.82	11.85	11.85	11.85	11.85	0	0	0
49/29/21	0.06	803.00	125.52	0.22	7.87	7.87	7.87	9.90	0	0	0
49/29/1	0.13	803.00	139.47	0.52	7.96	7.96	7.96	9.28	0	0	0
49/1	0.21	1052.00	154.97	0.75	6.15	5.22	6.15	11.65	0	0	0
49/261	0.12	250.00	101.67	1.13	18.30	18.30	18.30	18.30	0	0	0
49/470	0.05	229.00	82.36	0.61	18.13	18.13	18.13	18.13	0	0	1
4	0.26	2877.00	154.97	0.60	0.36	11.82	7.35	22.69	0	0	0
86/1	0.05	719.00	125.52	0.20	11.15	11.15	11.15	9.49	0	0	0
111/98	0.05	307.00	112.97	0.35	13.26	13.26	13.26	4.68	0	0	0
202	0.05	261.00	112.97	0.63	10.80	10.80	10.80	10.80	0	0	0
61/7/1	0.06	895.00	112.97	0.20	0.04	0.04	0.04	10.60	0	0	0
61/40/1	0.06	618.00	112.97	0.24	10.41	10.41	10.41	9.23	0	0	0
61/2	0.13	618.00	125.52	0.54	13.09	13.09	13.09	7.71	0	0	0
111/29/1	0.06	452.00	112.97	0.30	10.94	10.94	10.94	7.44	0	0	0
111/1	0.11	452.00	125.52	0.60	11.15	11.15	11.15	4.08	0	0	0
5	0.13	565.00	139.47	0.81	13.98	13.98	13.98	13.98	0	0	0
146/12/1	0.06	519.00	101.67	0.25	8.59	8.59	8.59	8.89	0	0	0
113	0.10	572.00	125.52	0.82	9.82	9.82	9.82	9.82	0	0	0
356/1	0.01	265.00	91.51	0.07	3.66	3.66	3.66	8.88	0	0	0
387/1	0.01	236.00	91.51	0.07	2.83	2.83	2.83	8.62	0	0	0

In Table 7-7, points of interest (red highlighted cells) indicate protection miscoordination before DG penetration when compared to operating times of secondary relays (*twenty-two points in*

total). Nineteen green highlighted cells with a red text show outliers: model predicted a negative relay operating time (shown in Table G-2 in appendix G), the calculated relay operating time before DG penetration was used as a replacement of the negative value. Orange highlighted cells indicate a loss of coordination (*twenty-eight points in total*). A vertical analysis was done on the data in Table 7-7 to identify network protection features that occasioned in the loss of grading and de-sensitization after three stages of DG penetration. With the vertical approach, as illustrated in Table G-2 in appendix G, the downstream relay operating time is subtracted from the upstream relay operation time and the difference in the ratio of the upstream to downstream relay operating time is used. Twenty eight loss of coordination points (2 points at a DG penetration of 35%, 1 point at a DG penetration of 65%, 25 points at a DG penetration of 75% of ADMD), 42 de-sensitization points (12 points at a DG penetration of 35%, 12 points at a DG penetration of 65%, 18 points at a DG penetration of 75% of ADMD) and 12 loss of grading points (1 at a DG penetration of 35%, 1 point at a DG penetration of 65% and 10 points at a DG penetration of 75% of ADMD) were observed from the vertical approach (Figure 7-1).

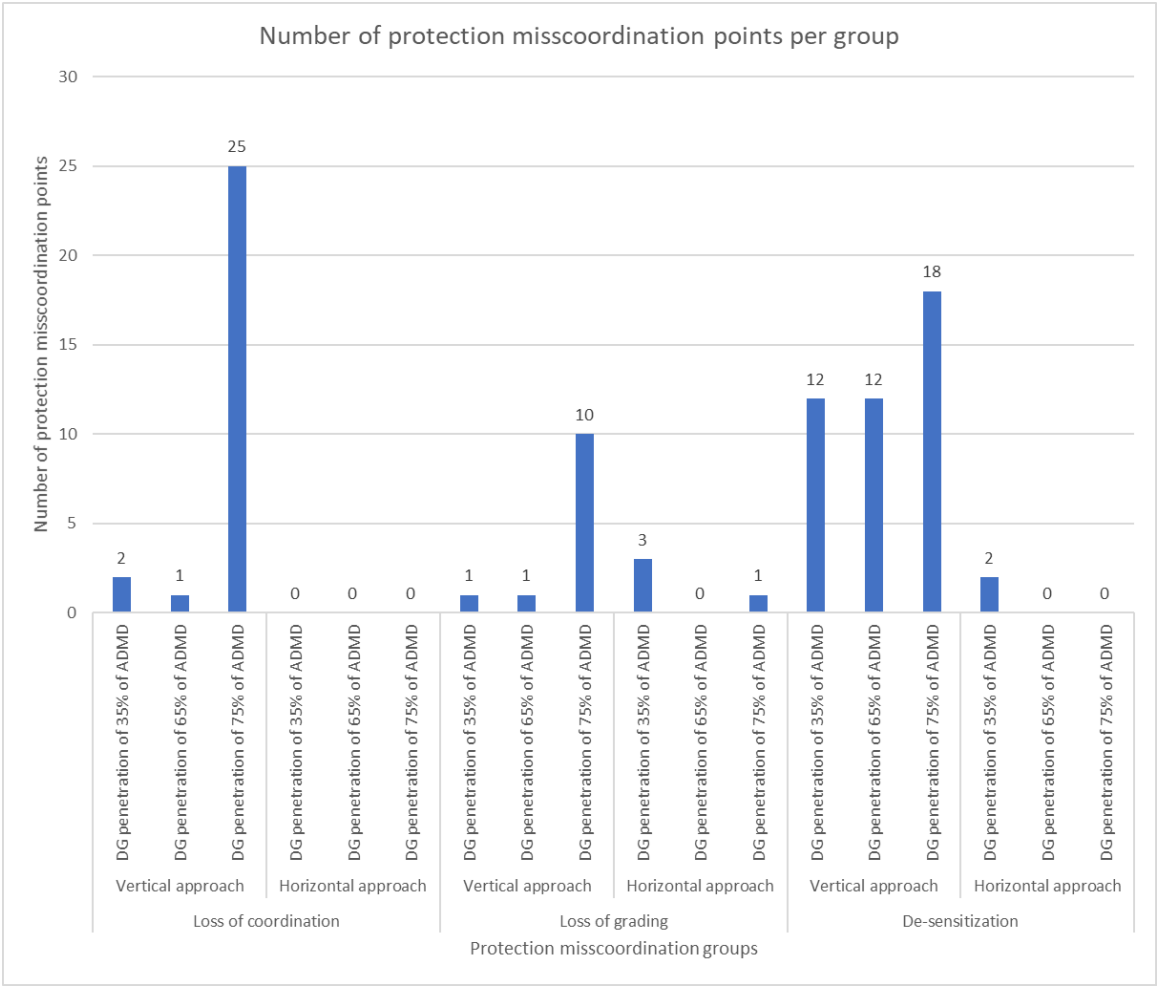


Figure 7-1: Number of protection miscoordination points per protection miscoordination group

As illustrated in Figure 7-1, more protection miscoordination points were found using the vertical approach compared to the horizontal approach (linear and logistic regression models). The difference implies that the horizontal approach cannot be used as a standalone method for the classification of protection miscoordination groups.

8. Conclusions

This research project aimed to investigate the effect of network protection parameters on relay operating times before and after DG penetration, develop a machine learning model to determine protection miscoordination before and after DG penetration on MV distribution networks, and develop a protection-based DG penetration limit medium voltage opportunity network guide.

Based on the recommendation systems (section 5.9), model performance (section 6.5), and model application results (chapter 7), the following conclusions have been drawn:

8.1. Summary of answers to the research questions

8.1.1. What is the impact of DG on feeders protected by Eskom's MV protection philosophy?

The impact of DGs on the network can be grouped into three groups: voltage, loading, and protection. Each group has a subcategory. Voltage group subcategories: over and under-voltage. Loading group subcategories: thermal charging and discharging. Protection group subcategories: additional fault current, sympathetic relay tripping, relay reach reduction, reverse power flow, and unintentional islanding.

In this research, the impact on protection was investigated. All the subcategories for protection were grouped into three groups with a specific criterion for each. From the three types of protection miscoordination groups (defined in section 6.1), only loss of grading and de-sensitization were experienced by the four MV distribution networks protected by Eskom's MV protection philosophy.

8.1.2. What are the factors that affect DG planning and connection approvals?

Planning for DG network connection requires fault level contribution (its contribution to fault current over time) of the DG and all information that will enable the planner to model the DG. This information is not readily available for small-scale DGs that can be connected to MV networks, and the impact assessments studies take time to execute for each DG using load flow programs e.g., DIgSILENT. Also, the capacity, location, type, and date when the customer installations are connected are unplanned and random from the network planner's perspective (**network DG planning dilemma**). Because this information must be supplied by the owner of the DG before the network planners can perform feasibility studies, the connection approval process is affected. This limits the growth of DG in South Africa.

8.1.3. What are the tools/concepts used to establish a network's DG hosting capacity?

Tools used to determine a network's DG hosting capacity consists of the following limiting factors: thermal limits, power quality/voltage, protection coordination, and reliability/safety. In this research, two tools were discussed: The Herman Beta algorithm modified for voltage calculation in active LV feeders using the negative load approach for DG models to analyse voltage rise constraints [21], and the Electric Power Research Institute (EPRI) system-wide DG assessment. The modified Herman beta algorithm only uses one of the limiting factors. Although the EPRI tool can use three of the four limiting factors, the entire network still needs to be modelled in DIgSILENT which becomes time-consuming for complex, large and real networks, thus,

worsening the network DG planning dilemma. This brings about a delay in the approval of new connection requests, thus slowing down the growth of distributed generation.

8.1.4. What combination of protection features determine which protection coordination problem?

Three protection coordination problems were defined. Loss of coordination was described as an increase in short circuit level which makes secondary relays have shorter operation times than primary relays. Loss of grading was described as a situation when there is an operating time difference between primary and secondary relays greater/less than 300ms. Lastly, de-sensitization was described as the reduction in the fault current level drawn from the mini substation i.e. slow operating time. Therefore, to answer the question in the subheading: TMS and fault levels were found to be the combination of protection features that determine loss of grading while TMS, fault level, and relay location determine de-sensitization.

8.1.5. What are the protection-based DG penetration limits for MV networks protected by Eskom's MV protection philosophy?

Protection based DG penetration limits for MV networks protected by Eskom's protection philosophy were found to be less than 35% of the After Diversity Maximum Demand (ADMD). Amongst all the four MV distribution networks, only de-sensitization and loss of grading were experienced using the horizontal approach. From the network protection features, TMS and fault levels had a direct impact on the penetration limit. Loss of grading was only experienced on the transformer breakers with a TMS between 0.5 and 1.1 and a fault level between 650 A and 800 A. Transformer breakers also experienced de-sensitization when the TMS was greater than 0.6 and less than 1.1 and the fault level was greater than 710 A and less than 800 A. A transformer breaker with a TMS and a fault level outside the above-mentioned range ($0.5 \leq TMS \leq 1.1$ and $650 \leq$ Fault level ≤ 800 A in Table 7-3) had a DG penetration limit of 75% of the ADMD. Therefore, for an area projected to have a high DG penetration, only feeders with a TMS outside the range of 0.5 and 1.1 and a fault level less than 650 A or greater than 800 A need to be incorporated into the design.

8.1.6. Where, when, and how much PV is/will be connected.

The results from the recommendation systems alone already provided a partial solution to the DG planning dilemma through insights to the randomness of the number, size, and location of DGs to be connected. The most likely PV that will be connected by consumers was found to be 5.3 kW panels (Figure 5-1). The growth rate in the number of PV installation was found to be 1.5 per year. Two areas with the highest number of PV installations were found to be Centurion, Gauteng (postal code of 0157) followed Somerset area, Western Cape (7130). Both these areas consist of affluent neighbourhoods. The time of day where maximum generation and minimum loading occurs was found to be 11:30 am to 12:00 pm, other times of the day similar to these times were 10:00, 10:15, 10:30, 11:00 am for maximum PV generation, and 1:30 pm for minimum loading.

8.2. Model Performance

The cost function on the model for calculating relay operating times before DG penetration using all the network features was minimized to a value below 120 $seconds^2$ within the first 100 iterations and the cross-validation and training error approached the desired performance of 0.3±0.1. The cost function on the model for calculating relay operating time after 35% of DG penetration using all the network features with added polynomials was minimized to a value below 20 $seconds^2$ within the first 100 iterations and the cross-validation and training error approached the desired performance of 0.3±0.1. The cost function on the model for calculating

relay operating time after 65% of DG penetration using all the network features with added polynomials was minimized to a value below 15 $seconds^2$ within the first 100 iterations, and the cross-validation and training error approached the desired performance of 0.3±0.1.

The cost function on the model for calculating relay operating time after 75% of DG penetration using all the network features was minimized to a value below 10 $seconds^2$ within 100 iterations (Figure 6-4), however, the cross-validation and training error indicated a high variance problem (cross-validation error was high and training error was low). The cost function on the model for calculating the relay operating time with fitted polynomials and features eliminated was minimized to a value below 15 $seconds^2$ within the first 100 iterations and the cross-validation and training error approached the desired performance of less than 0.3±0.1.

A training accuracy of 91.30% and validation accuracy of 100% was achieved when classifying loss of coordination for the model with second-degree polynomials added. This model had two false negatives for the training set and zero false negatives for the cross-validation set. A training accuracy of 73.91% and validation accuracy of 55.56% was achieved when classifying loss of grading for the model with relay position eliminated. This model had 6 false negatives for the training set and 4 false negatives for the cross-validation set. A training accuracy of 82.61% and validation accuracy of 66.67% was achieved when classifying de-sensitization for the model with second-degree polynomials. This model had 3 false negatives for the training set and 3 false negatives for the cross-validation set. In DG network planning, protection engineers can use the model to predict loss of coordination with 100% accuracy, loss of grading with 55.56% accuracy and de-sensitization with 66.67% accuracy for DG penetration up to 75% of the ADMD.

When the linear and logistic regression models (horizontal network protection features) were applied to four MV distribution networks, loss of coordination was not predicted. The loss of the grading model had one false positive, and the de-sensitization model had one false negative. From the vertical analysis, comparing the operating times of the upstream, downstream relays/reclosers, 22 relay operating times were predicted with protection miscoordination before DG penetration. Also, 28 points indicated loss of coordination (2 points at a DG penetration of 35%, 1 point at a DG penetration of 65% and 25 points at a DG penetration of 75% of the ADMD), 42 points indicated de-sensitization (12 points at a DG penetration of 35%, 12 points at a DG penetration of 65% and 18 points at a DG penetration of 75%of the ADMD) and 12 points indicated loss of grading (1 point at a DG penetration of 35%, 1 point at a DG penetration of 65% and 10 points at a DG penetration of 75% of ADMD (Figure 7-1)).

Thus, the horizontal approach cannot be used as a standalone method of classifying protection miscoordination in MV distribution networks and must be used in conjunction with the vertical approach.

8.3. The validity of the hypothesis

The hypothesis, “Machine learning techniques: value estimation and recommendation systems can be modified, trained and applied as a tool to determine/analyse protection-based penetration limits on distribution networks protected using Eskom’s MV protection philosophy”, was addressed in this study and is demonstrated to be valid. The research demonstrated that recommendation systems can be used to solve the where, when, and how much DG will be connected part of the DG network planning dilemma. The study also demonstrated a successful way of modelling the network parameters as features for machine learning. Machine learning value estimation techniques were used as an engine for determining DG protection-based penetration limits. Regression was used in predicting relay operating times before and after

different levels of DG penetration. Logistic regression was used in the classification of protection miscoordination. These techniques together with the vertical analysis on the network protection features formed an important section for this research project- network protection features that determine the protection-based DG penetration limits on MV feeders. These network protection features were found to be the TMS and fault level.

9. Recommendations

Based on the above conclusions, the following recommendations are made.

9.1. Medium voltage opportunity network guide

Protection coordination reinforcements (against loss of coordination and desensitization) are required for DG connection where the MV transformer circuit breaker TMS is between **0.5 and 1.1** and where the network fault level is between **650 and 800A.**

Distribution networks in affluent neighbourhoods similar to those in the Western Cape (Somerset West area) and Gauteng (Centurion area) need to be reinforced to accommodate maximum DG penetration up to the limit of 75% of the ADMD. DG network planning engineers need to incorporate a 5.3 kW PV system installation growth rate of 1.5 per year in their planning scenarios. New feeder protection coordination designs for projected DG penetration growth areas should be outside the range: $0.5 \leq TMS \leq 1.1$ and $650 \leq$ Fault level $\leq 800A$.

9.2. Protection-based DG penetration limits

The protection-based DG penetration limit for networks protected by Eskom's MV protection philosophy was found to be 35% of the ADMD. DG connections up to 35% of the feeder ADMD can be allowed without the need for protection coordination analysis and redesign.

9.3. Improve model accuracy

To improve the accuracy of the models, relay operating times from detailed analysis on the impact of PV DG on power protection systems: single-phase to ground/neutral fault response data should be collected to use in training the models to solve the high variance problem is recommended.

9.4. Model modification

Add upstream and downstream network protection features as inputs in the classification model to solve high bias problems.

References

[1] S. Yomna and I. Sherwat Elwan, "Is the sharing economy a valid option for scientific research technologies in emerging economies?," *IEEE Engineering Management Review,* vol. 46, no. 3, pp. 87-93, 2018.

[2] P. I. Santos e Abreu and A. G. Martins, "Assessment of the behavior of protection systems in radial networks with distributed generation," in *2016 51st International Universities Power Engineering Conference (UPEC)*, Coimbra, 2016.

[3] K. A. Wheeler, S. O. Faried and M. Elsamahy, "Fault impedance effects on distributed generation influences in overcurrent protection," *2017 IEEE Manchester PowerTech,* pp. 1-6, 2017.

[4] T. K. Abdel-Galil, A. E. B. Abu-Elanien, E. F. El-Saadany, A. Girgis, Y. A.-R. I. Mohamed, M. M. A. Salama and H. H. M. Zeineldin, "Protection Coordination Planning with distributed generation," NRCA, Montreal, 2007.

[5] M. Slabbert and S. Van Zyl, "Validation of Medium Voltage Distribution Feeder Protection Settings," in *Eskom Protection Workshop*, Midrand, 2015.

[6] Eskom PTM&C, "Protection Settings Philosophy for Medium Voltage Distribution Networks," Eskom, Johannesburg, 2015.

[7] Eskom PTM&C, "Standard for the Interconnection of Embedded Generation," Eskom, Sunninghill, 2013.

[8] E. Purwar, D. N. Vishwakarma and S. P. Singh, "A New Adaptive Inverse-time Protection Scheme for Modern Distribution Systems with Distributed generation," in *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington, DC, 2017.

[9] S. A. M. Javadian, M. R. Haghifam and M. Massaeli, "Risk analysis of protection system's operation in distribution networks with DG," in *2011 International Conference on Power Engineering, Energy and Electrical Drives*, Malaga, Spain, 2011.

[10] B. Hussain, S. M. Sharkh, S. Hussain and M. A. Abusara, "An Adaptive Relaying Scheme for Fuse Saving in Distribution Networks With Distributed Generation," *IEEE Transactions on Power Delivery,* vol. 28, no. 2, pp. 669-677, 2013.

[11] Electrica4u, "Protection of Lines or Feeder," [Online]. Available: https://www.electrical4u.com/protection-of-lines-or-feeder/. [Accessed 17 August 2017].

[12] H. Yazdanpanahi, Y. W. Li and W. Xu, "A New Control Strategy to Mitigate the Impact of Inverter-Based DGs on Protection System," *IEEE Transactions on Smart Grid,* vol. 3, no. 3, pp. 1427-1436, 2012.

[13] H. Margossian, G. Deconinck and J. Sachau, "Distribution network protection considering grid code requirements for distributed generation," *IET Generation, Transmission & Distribution,* vol. 9, no. 12, pp. 1377-1381, 2015.

[14] S. S. Alkaabi, H. H. Zeineldin and V. Khadkikar, "Adaptive planning approach for customers DG installation in smart distribution networks," *IET Renewable Power Generation,* vol. 12, no. 1, pp. 81-89, 2018.

[15] M. Theku, B. Dlamini, C. Ndlhovu and T. Ngcobo, "Strategic Plan for Integrating Future Renewable Energy Generators in South Africa," in *CIGRE Electricity Supply to Africa and Developing Economies: Challenges and Opportunities Conference*, Cape Town, 2017.

[16] J. Smith, M. Rylander, M. Bello, D. Gusain and D. Montenegro, "Distribution Planning with DER: System-Wide Assessment," Electric Power Research Institute, 2015.

[17] T. Aziz and N. Ketjoy, "PV Penetration Limits in Low Voltage Networks and Voltage Variations," *IEEE Access,* vol. 5, pp. 16784-16792, 2017.

[18] K. Kemper, U. Minnaar and A. Rix, "Analysis of Utility Scale Wind and Solar Plant Performance in South Africa Relative to Daily Electricity Demand," in *CIGRE Electricity Supply to Africa and Developing Economies: Challenges and Opportunities Conference*, Cape Town, 2017.

[19] S. Chaitusaney and A. Yokoyama, "Prevention of Reliability Degradation from Recloser–Fuse Miscoordination Due To Distributed Generation," *IEEE Transaction on Power Delivery,* vol. 23, no. 4, pp. 2545-2554, 2008.

[20] S. Chaitusaney and A. Yokoyama, "An Appropriate Distributed Generation Sizing Considering Recloser-Fuse Coordination," in *IEEE/PES Transmission and Distribution Conference & Exhibition*, Dalian, 2005.

[21] E. Namanya and C. Gaunt, "Voltage Calculation on Low Voltage Feeders with Distributed Generation," Masters Dissertation, University of Cape Town, Cape Town, 2014.

[22] E. Nxumalo, K. Awodele and A. B. Sebitosi, "Protection Based Distributed Generation Penetration Limits For the South African NRS 097-2-3 Standard," in *26th Southern African Universities Power Engineering Conference*, Johannesburg, 2018.

[23] M. Bello, M. Rylander, J. Smith, D. Gusain, S. Malapermal and K. Dedekind, "Innovative Method to Determine Distributed Energy Resources Hosting Capacity," in *CIGRE Electricity Supply to Africa and Developing Economies: Challenges and Opportunities Conference*, Cape Town, 2017.

[24] A. Fatama, A. Haque and M. Ali Khan, "A Multi Feature Based Islanding Classification Technique for Distributed Generation Systems," in *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, 2019.

[25] H. Monowar, M. Saad, D. Malihe, O. Lanre and S. Shahaboddin, "Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems," *Journal of Cleaner Production,* vol. 167, pp. 395-405, 2017.

[26] O. A. Alimi, K. Ouahada and A. M. Abu-Mahfouz, "A Review of Machine Learning Approaches to Power System Security and Stability," *IEEE Access,* vol. 8, pp. 113512-113531, 2020.

[27] A. Ng, "Machine Learning | Coursera," Coursera, 15 August 2011. [Online]. Available: https://www.coursera.org/learn/machine-learning/home/welcome. [Accessed 21 April 2020].

[28] SMA Solar Technology, "SMA Sunny portal," SMA Solar Technology AG, 2019. [Online]. Available: https://www.sunnyportal.com/Templates/PublicPagesPlantList.aspx. [Accessed 02 November 2019].

[29] geonames.org, "GeoNames," GeoNames, [Online]. Available: https://www.geonames.org/postalcode-search.html?q=&country=ZA. [Accessed 03 05 2020].

[30] d-maps.com, "d-maps.com >Africa >South Africa >South Africa >outline, provinces, color (white)," d-maps.com, [Online]. Available: https://d-maps.com/carte.php?num_car=34179&lang=en. [Accessed 03 05 2020].

[31] CRSES, "Centre For Renewable and Sustainable Energy Studies," University of Stellenbosch, 2014. [Online]. Available: http://www.crses.sun.ac.za/research-publications-resources. [Accessed 10 November 2017].

[32] ACTOM Pty Ltd, "Distribution Transformers - Actom," 16 November 2016. [Online]. Available: https://www.actom.co.za/wp-content/uploads/2016/10/Distribution-Transformers-Brochure.pdf. [Accessed 19 September 2020].

[33] Ö. Başţanlar and Ö. Mustafa, "Introduction to Machine Learning," *miRNomics: MicroRNA Biology and Computational Analysis. Methods in Molecular Biology (Methods and Protocols),* vol. 1107, pp. 105-128, 2014.

[34] X. Xu, J. Zhou, Y. Liu, Z. Xu and X. Zhao, "Taxi-RS: Taxi-Hunting Recommendation System Based on Taxi GPS Data," *IEEE Transactions on Intelligent Transportation Systems,* vol. 16, no. 4, pp. 1716-1727, 2015.

[35] S. Shaikh, S. Rathi and P. Janrao, "Recommendation System in E-Commerce Websites: A Graph Based Approached," in *2017 IEEE 7th International Advance Computing Conference (IACC)*, Hyderabad, 2017.

[36] K. AL Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riff and A. Sabr, "Hybrid recommender system for tourism based on big data and AI: A conceptual framework," *Big Data Mining and Analytics,,* vol. 4, no. 1, pp. 47-55, 2021.

[37] E. Nxumalo and K. Awodele, "Impact of Distributed Generation (DG) on the Protection System of a Network using DIgSILENT Power Factory," BSc. Research Report, University of Cape Town, Cape Town, 2017.

[38] C. Liu, "5 Concepts You Should Know About Gradient Descent and Cost Function," Fintech Industry, May 2020. [Online]. Available: https://www.kdnuggets.com/2020/05/5-concepts-gradient-descent-cost-function.html. [Accessed 10 August 2020].

[39] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters,* vol. 27, pp. 861-874, 2006.

[40] R. Ng, "Evaluating a Classification Model," Ritchie Ng, [Online]. Available: https://www.ritchieng.com/machine-learning-evaluate-classification-model/. [Accessed 06 05 2020].

[41] Alstom Firm, "Overcurrent Protection for Phase and Earth faults," in *Network Protection & Automation Guide*, Saint-Ouen, Alstom Grid, 2011, pp. 131-136.

[42] South African National Standard, "NRS 097-2-3: Grid interconnection of Embedded Generation," SABS Standards Division, Pretoria, 2014.

[43] South African National Standard, "NRS 097-2-1: Grid interconnection of Embedded Generation," SABS Standards Division, Pretoria, 2010.

[44] Republic of South Africa: Department of Energy, "Integrated Resource Plan For Electricity (IRP) 2010-2030," 21 November 2013. [Online]. Available: http://www.doe-irp.co.za/content/irp2010_updatea.pdf. [Accessed 14 October 2017].

[45] Eskom, "Facts & Figures," January 2015. [Online]. Available: http://www.eskom.co.za/AboutElectricity/FactsFigures/Documents

/TD_0003TransmissionDistributionElectricityRev8.pdf. [Accessed 15 October 2017].

[46] Y. Paithankar and S. Bhide, "Overcurrent Protection of Transmission lines," in *Fundamentals of Power System Protection*, New Delhi, Prentice-Hall of India Private Limited, 2003, pp. 38-41.

[47] RSA Grid Code Secretariat, "Grid Connection Code for Renewable Power plants connected to the electricity transmission system or the distribution system in South Africa," Eskom Transmission Division, Johannesburg, 2014.

[48] South African National Standard, "NRS 034-1: Electricity Distribution-Guidelines for the Provision of Electricity Distribution Networks in Residential Areas," SABS Standards Division, Pretoria, 2007.

[49] D. J. Glover, M. S. Sarma and T. J. Overbye, Power System Analysis & Design 5th edition, Illinois: Cengage Learning, 2012.

[50] H. Alireza, G. A. Vassilios, Z. Hadi and H. Maryam, "Prevention of overcurrent relays miscoordination in distribution system due to high penetration of distributed generation," in *2013 International Conference on Renewable Energy Research and Applications (ICRERA)*, Spain, 2013.

[51] K. Awodele, "Overcurrent Relays EEE4089F 2017," Lecture Notes, University of Cape Town, Cape Town, 2017.

A. Appendix A - Publications

From the research, the following paper was published.

1. “A Review on the Use of Recommendation and Value Estimation Systems to Determine Protection Based Distributed Generation Penetration Limits,” in International SAUPEC/RobMech/PRASA Conference, Cape Town, 2020.

B. Appendix B – Machine Learning

This appendix holds machine learning techniques and approaches for improving their performances.

1. Machine learning

Machine learning provides an opportunity to automatically build computational models of complex relationships (real-world problems that cannot be modelled directly as a closed-form input-output relationship) through the processing of data available and maximising a problem-dependent performance criterion [28]. Arthur Samuel [27], described machine learning as the field of study that gives computers the ability to learn without being explicitly programmed. Tom Mitchel [27] defined a learning problem as follows: "*A computer program is said to learn from* ***experience E*** *with respect to some* ***task T*** *and some performance* ***measure P****, if its performance on T, as measured by P, improves with experience E*". A top-level machine learning diagram is shown in Figure B-1.

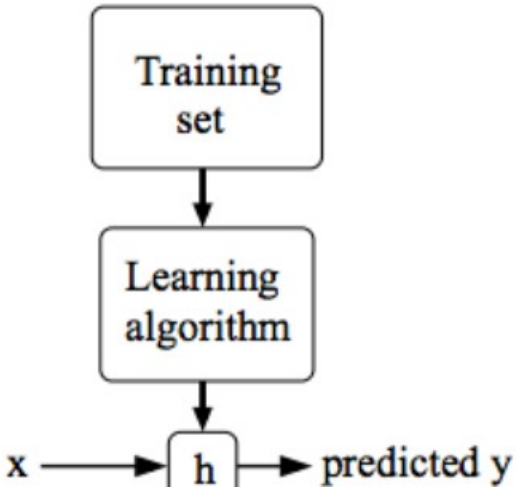


Figure B-1: Top-level flow diagram of machine learning (where h is the hypothesis function) [27]

1.1. Machine learning categories

Classical machine learning can be grouped into two main categories:

- Supervised learning
- Unsupervised learning

Other groups include Reinforcement learning and recommender systems. Figure B-2 shows some of the other machine learning categories.

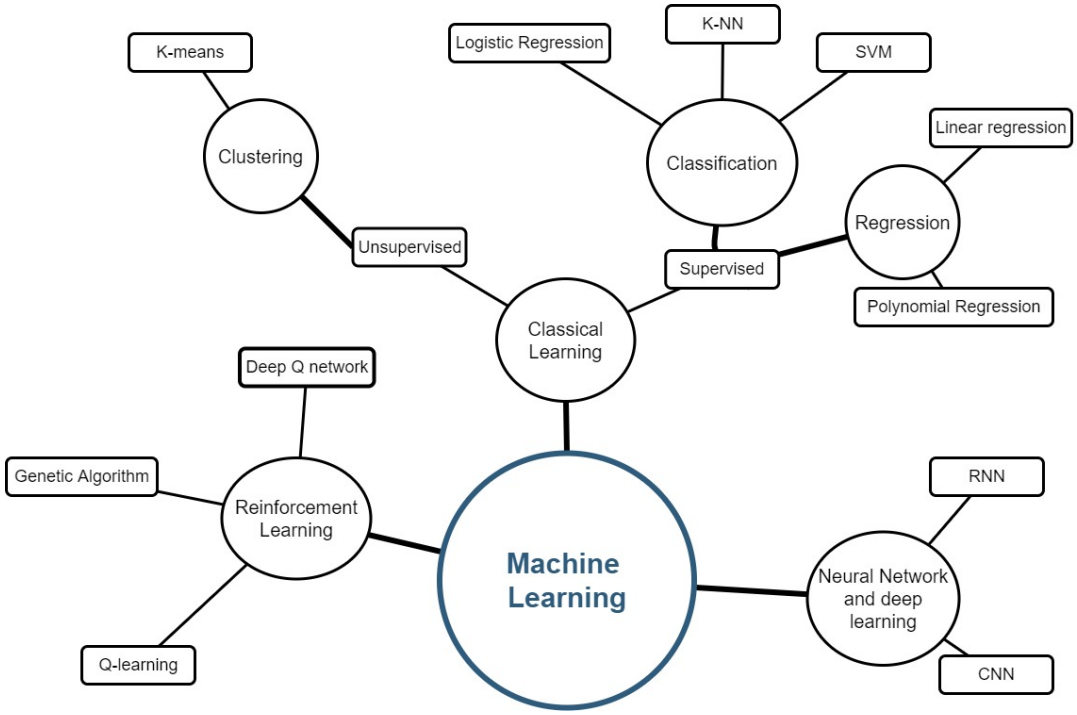


Figure B-2: Machine learning categories

In supervised learning, the correct output from the training data set is known and expect that there is a relationship between the input and output variables. Each training sample comes with both input and output values [28]. Supervised learning can also be broken down into two groups:

- Regression and
- Classification (logistic regression).

For regression problems, the goal is to predict the output of a continuous function. For classification problems, the aim is to group discrete categories into a discrete group. For unsupervised learning, a structure is derived from data sets whereby the effect of the variables is unknown: no feedback based on predicted results [27].

a) Linear regression

As introduced in section 1.1, in linear regression problems, the goal is to predict the output of a continuous function. This is done by using a function (linear or polynomials) to determine an output. The objective is to get the best possible line that fits the data. The criterion for the best line is one whereby the average squared vertical distances of the scattered points from the line is minimal. The ideal state is when the best fit line passes through all the data points [27]. As in the example below, the goal is to minimise the squared error function.

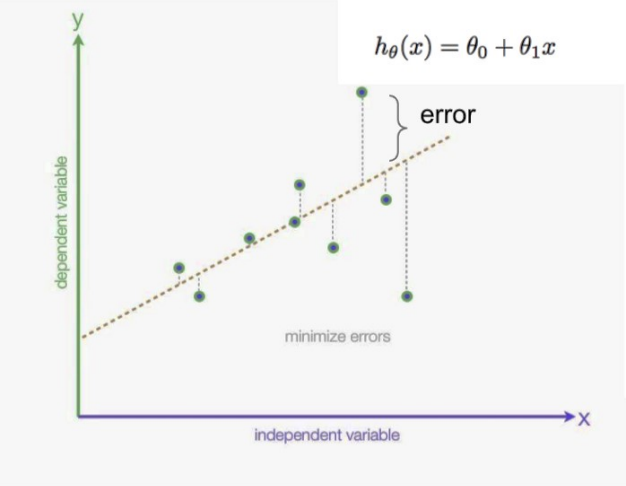


Figure B-3: Linear regression example [38]

Hypothesis: $$h_\Theta(x) = \Theta_0 + \Theta_1 x \quad \text{(B-10)}$$

Parameters: $$\Theta_0, \Theta_1 \quad \text{(B-11)}$$

Cost function: *"One half mean squared Error"* $$J(\Theta_0\Theta_1) = \frac{1}{2m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})^2 \quad \text{(B-12)}$$

Goal: $$\min_{\Theta_0\Theta_1} J(\Theta_0, \Theta_1) \quad \text{(B-13)}$$

$$\frac{\partial}{\partial\Theta_0}J(\Theta_0\Theta_1) = \frac{1}{m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)}) \quad \text{(B-14)}$$

Derivatives

$$\frac{\partial}{\partial\Theta_1}J(\Theta_0\Theta_1) = \frac{1}{m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)}) * x^{(i)} \quad \text{(B-15)}$$

As indicated in Figure B-3, the cost function measures the accuracy of the hypothesis function. The cost function is an 'average' of the difference of all the results of the hypothesis function from the actual values (one half of the mean squared error) [27]. For this research, gradient descent -

which is an efficient optimization algorithm that attempts to find a local or global minimum of a function - will be used.

Gradient descent iterates and finds optimal values of the parameters that correspond to the minimum values of the cost function using calculus. It is important to note that for this to work we need to know the slope and direction to move the coefficient values to get a lower cost on the next iteration [38]. As shown in Figure B-4 and Figure B-5, the final point is the global minimum. This is done by taking steps down the cost function in the direction of the steepest descent. The size of each step is determined by α, the learning rate [27]. The effects of the learning rate are shown in Figure B-6.

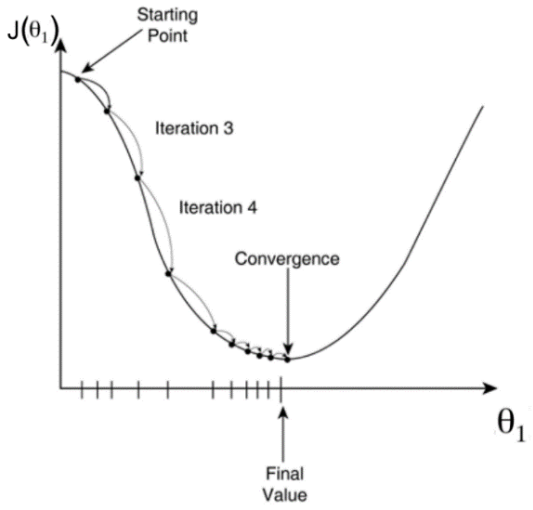


Figure B-4: θ1 gradually converges towards a minimum value [38]

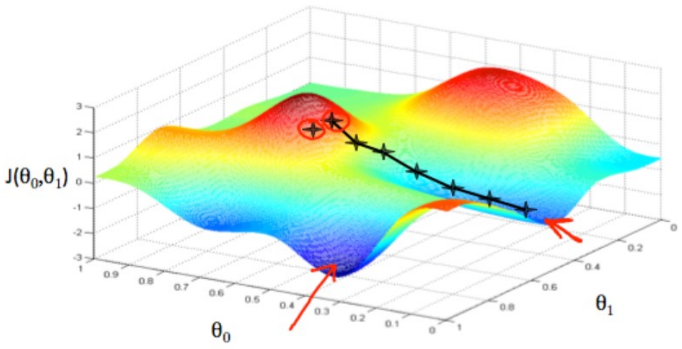


Figure B-5: 3D visualisation of the cost function and gradient descent [27]

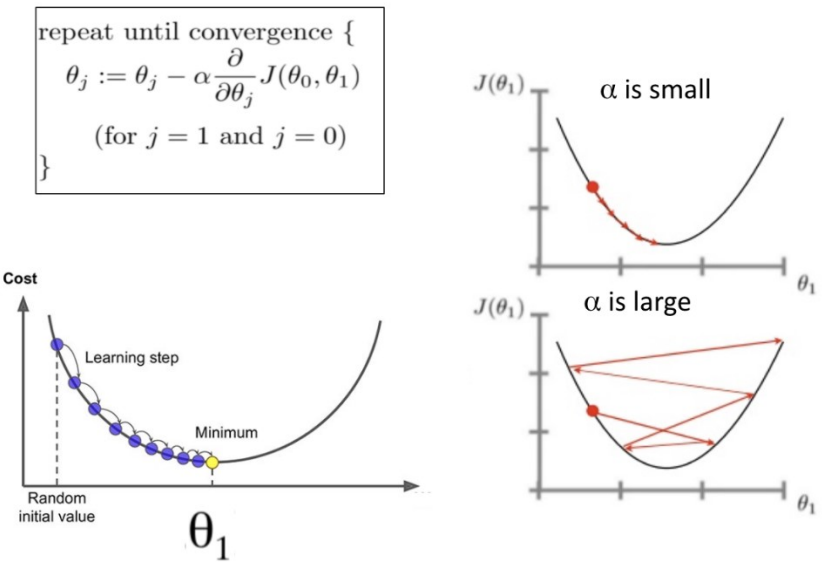


Figure B-6: Effects of the learning rate on gradient descent [38]

To make sure that the gradient descent is working, the graph of the cost function vs the number of iterations should decrease by less than 10^{-3} in one iteration [27].

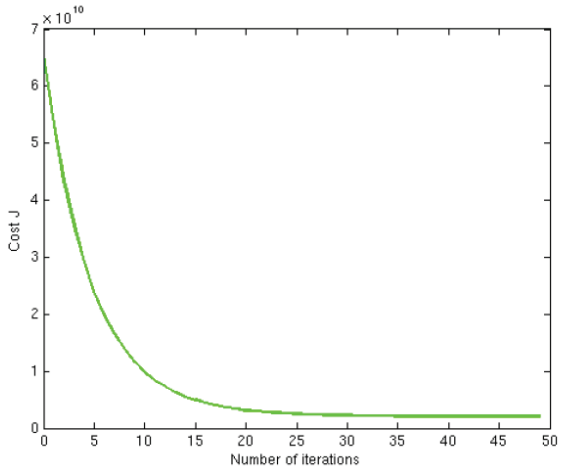


Figure B-7: Convergence of gradient descent with appropriate learning rate [27]

i) Feature scaling and mean normalization

To speed up gradient descent, it helps to have all the input values in roughly the same range. The reason for this is because θ descends fast on small ranges and slowly on large ranges. Thus, gradient descent oscillates inefficiently towards the optimum value when variables are extremely uneven [27].

In feature scaling, the input ranges are modified to be between -1 and 1 or -0.5 and 0.5. The inputs are divided by the range (i.e. the maximum value minus the minimum value) of the input variable, resulting in a new range. Mean normalization involves subtracting the average value from each input variable resulting in a new average value.

$$x_i := \frac{x_i - \mu_I}{s_i} \quad \text{(B-16)}$$

Where μ_I is the average of all the values for feature (i) and s_i is the standard deviation [27].

b) *Logistic regression*

As introduced in section 1.1, in logistic regression problems, we group discrete categories. This is done by using a sigmoid function (Figure B-8) to determine an output that falls between the two criteria: 1 or 0, true, or false (yes/no).

$$h_\Theta(x) = g(\Theta^T x) \quad \text{((B-17)}$$

$$z = \Theta^T x \quad \text{(B-18)}$$

$$g(z) = \frac{1}{1 + e^{-z}} \quad \text{(B-19)}$$

Where: $0 \leq h_\Theta(x) \leq 1$ and $h_\Theta(x) = \frac{1}{1+e^{-\Theta^T x}}$

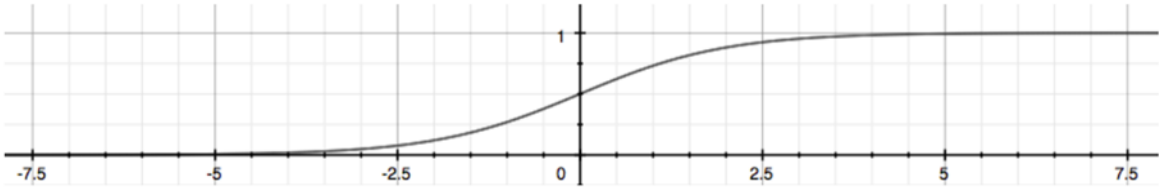


Figure B-8: Sigmoid function [27]

$h_\Theta(x)$ gives a probability that the output is a 1/0. For example, if $h_\Theta(x) = 0.8,$ the probability that the output is 1 is 80% and that the output is 0 is 20%.

$$h_\Theta(x) = P(y = 1|x;\ \Theta) = 1 - P(y = 0|x;\ \Theta) \quad \text{(B-20)}$$

$$P(y = 0|x;\ \Theta) + P(y = 1|x;\ \Theta) = 1 \quad \text{(B-21)}$$

To establish a decision boundary (the area that separates where y = 0 and y = 1), the hypothesis function is set up as follows:

$$h_\Theta(x) \geq 0.5 \rightarrow y = 1 \quad \text{(B-22)}$$

$$h_\Theta(x) < 0.5 \rightarrow y = 0 \quad \text{(B-23)}$$

This is supported by the sigmoid function, $g(z) \geq 0\ when\ z \geq 0$. Given that the input is $\Theta^T x$ therefore:

$$h_\Theta(x) = g(\Theta^T x) > 0.5 \quad when\ \Theta^T x\ \geq 0 \quad \text{(B-24)}$$

And

$$\Theta^T x\ \geq 0 \rightarrow y = 1 \quad \text{(B-25)}$$

$$\Theta^T x\ < 0 \rightarrow y = 0 \quad \text{(B-26)}$$

The cost function is introduced as:

$$J(\Theta) = \frac{1}{m}\sum_{t=1}^{m} Cost(h_\Theta(x^{(i)}), y^{(i)}) \quad \text{(B-27)}$$

$$Cost(h_\Theta(x), y) = -\log(h_\Theta(x))\ if\ y = 1 \quad \text{(B-28)}$$

$$Cost(h_\Theta(x), y) = -\log(1 - h_\Theta(x))\ if\ y = 0 \quad \text{(B-29)}$$

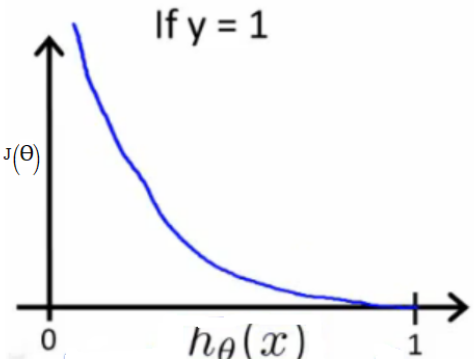


Figure B-9: Relationship between the cost function and hypothesis function (J vs h graph) [27]

Therefore, the cost function is represented as follows:

$$Cost(h_\Theta(x), y) = 0\ if\ h_\Theta(x) = y \quad \text{(B-30)}$$

$$Cost(h_\Theta(x), y) \rightarrow \infty \; if \; y = 0 \; and \; h_\Theta(x) \rightarrow 1 \quad \text{(B-31)}$$

$$Cost(h_\Theta(x), y) \rightarrow \infty \; if \; y = 1 \; and \; h_\Theta(x) \rightarrow 0 \quad \text{(B-32)}$$

The cost functions (equation B-30 to B-32) are then compressed into:

$$Cost(h_\Theta(x), y) = -ylog(\, h_\Theta(x)) - (1 - ylog(1 - h_\Theta(x)) \quad \text{(B-33)}$$

$$J(\Theta) = -\frac{1}{m}\sum_{t=1}^{m}[y^{(i)}\log\,(h_\Theta(x^{(i)})) + (1 - y^{(i)})\log\,(1 - h_\Theta(x^{(i)}))] \quad \text{(B-34)}$$

A vectorized implementation of equation B-34 is:

$$h = g(X\Theta),$$

$$J(\Theta) = \frac{1}{m}.(-y^T\log(h) - (1 - y)^T\log\,(1 - h)] \quad \text{(B-35)}$$

For gradient descent:

$$Repeat\;\{\;\Theta_j \coloneqq -\frac{\alpha}{m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})\,x^{(i)}{}_j\;\} \quad \text{(B-36)}$$

1.2. Machine learning techniques

a) Regularization

A regularization parameter determines the magnitude in which a theta parameter (Θ) is inflated. This is done to reduce the contribution of high-order polynomials to have the best fit. For example, Figure B-10 shows how the best fit line(pink) can be achieved by the minimization of high order polynomials (regularization).

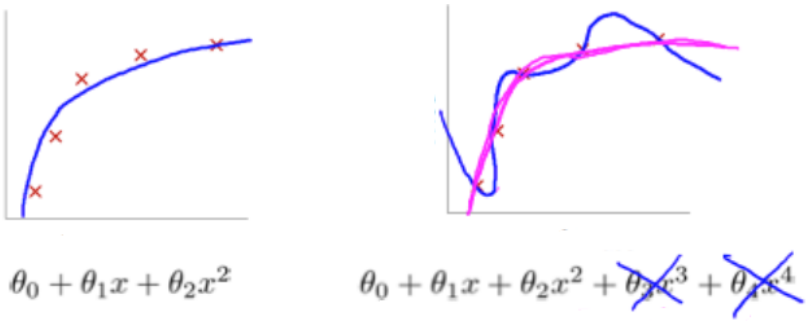


Figure B-10: Regularization example [27]

Lambda (λ), the regularization parameter is included in the cost equation as seen below:

$$min_\Theta \frac{1}{2m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})^2 + \lambda\sum_{j=1}^{n}\Theta^2{}_j \qquad \text{(B-37)}$$

The gradient descent equation with the regularization parameter is represented as follows:

$$\Theta_j := \Theta_j\left(1 - \alpha\frac{\lambda}{m}\right) - \alpha\frac{1}{m}\sum_{t=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})\,x^{(i)}{}_j \qquad \text{(B-38)}$$

b) *Classifier performance*

To evaluate the performance of a classifier, a 2x2 confusion matrix can be constructed:

Hypothesized class \ True class	p	n	
Y	True Positives	False Positives	fp rate $= \frac{FP}{N}$ tp rate $= \frac{TP}{P}$
N	False Negatives	True Negatives	precision $= \frac{TP}{TP+FP}$ recall $= \frac{TP}{P}$ accuracy $= \frac{TP+TN}{P+N}$
Column totals:	P	N	F-measure $= \frac{2}{1/\text{precision}+1/\text{recall}}$

Figure B-11: Confusion matrix and common performance metrics calculated from it [39]

$$\text{Specificity} = \frac{\text{True negatives}}{\text{False positives}+\text{True negatives}} = 1 - \text{fp rate}$$

The measure to focus on depends on the objective. The main aim is to identify between false negative and false positive which is more important to reduce in order to improve the classifier's precision or accuracy.

The terms in Figure B-11 can be explained as follows:

a) Recall commonly known as Sensitivity or True Positive (TP) rate: when the actual value is positive, how often is the prediction correct or how *sensitive* is the classifier in detecting positive instances.
b) Specificity: when the actual value is negative, how often is the prediction correct or how *specific* (or "selective") is the classifier in predicting positive instances.
c) Precision (Positive predictive value): When a positive value is predicted, how often is the prediction correct [40].

In classifying whether a recloser/breaker in a network will experience protection miscoordination at a certain DG penetration threshold (True Positive), it is important to optimize for **precision and accuracy.** False-positive (recloser/breaker not experiencing protection miscoordination and classified as true) is more acceptable than false negative (recloser/breaker experiencing protection miscoordination and classified as false). Hence, false negatives must be as low as possible.

c) Fitting

Two problems exist when trying to fit a hypothesis line through the data set: underfitting and overfitting. Underfitting or high bias is when the hypothesis function maps poorly to the trend of the data. The common cause is a simple hypothesis function or a few features (input variables). Overfitting or high variance is a situation whereby the hypothesis function fits the data perfectly but not generally, to predict new data points correctly. The common cause is a complex function with unnecessary curves unrelated to the data [27]. Figure B-12, Figure B-13, and Figure B-14 showcases of under- and overfitting as a result of features (candidate polynomials of different degrees (d)) and regularization.

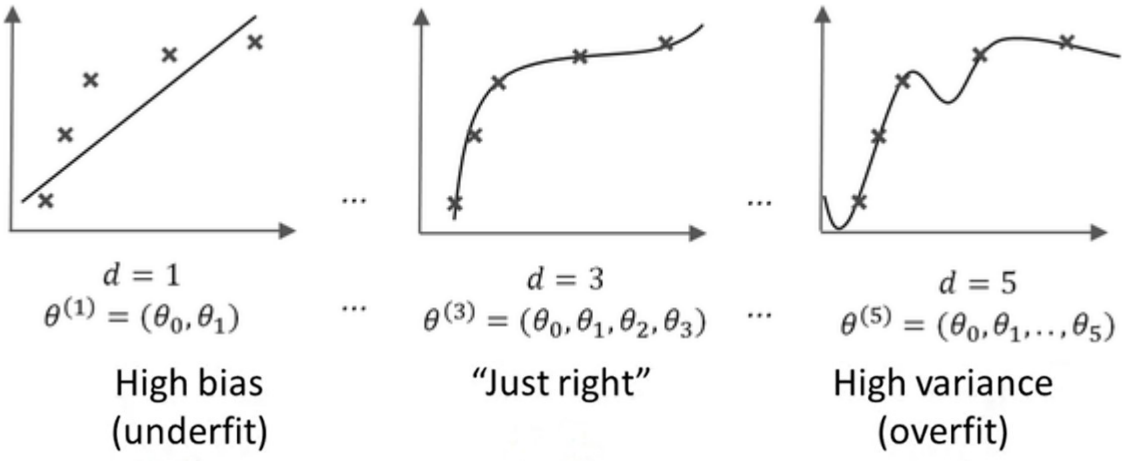


Figure B-12: Bias and variance as a result of features (candidate polynomials of different degrees (d)) [27] [28]

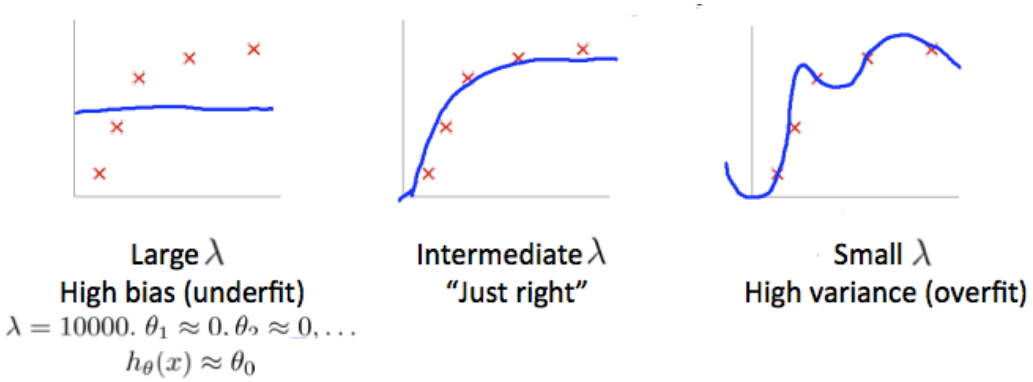


Figure B-13: Bias and variance as a result of regularization [27]

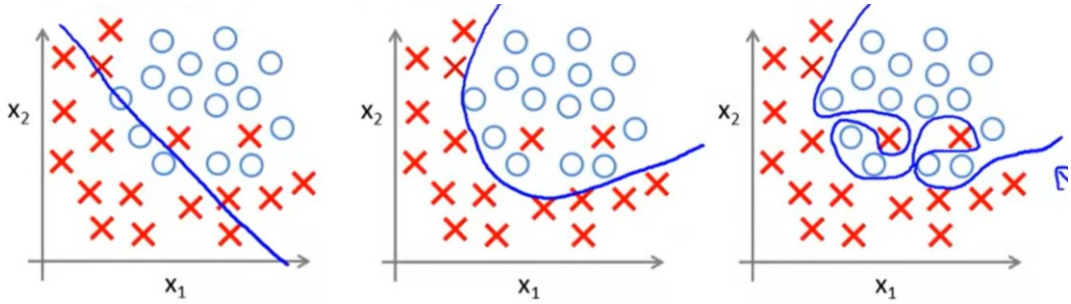


Figure B-14: Bias vs variance boundary in logistic regression [27]

To address overfitting, two main options exist:

- Reduction of the number of features or input variables- manually select features to keep.
- Regularization- keep all features but reduce the magnitude of the parameters (θ) [27].

i. Improving the Learning: learning curves

To evaluate a hypothesis, the training data set is split into three, training set, cross-validation, and test set. One way to breakdown the data: training set consist of **60%**, the cross-validation set consists of **20%** and the test set consists of **20%** of the total data set [27] (illustrated in Figure B-15). The training set is to learn how the inputs map to the outputs, cross-validation is to determine the regularization parameter and the test data set is to evaluate the performance of the model on "unseen" examples.

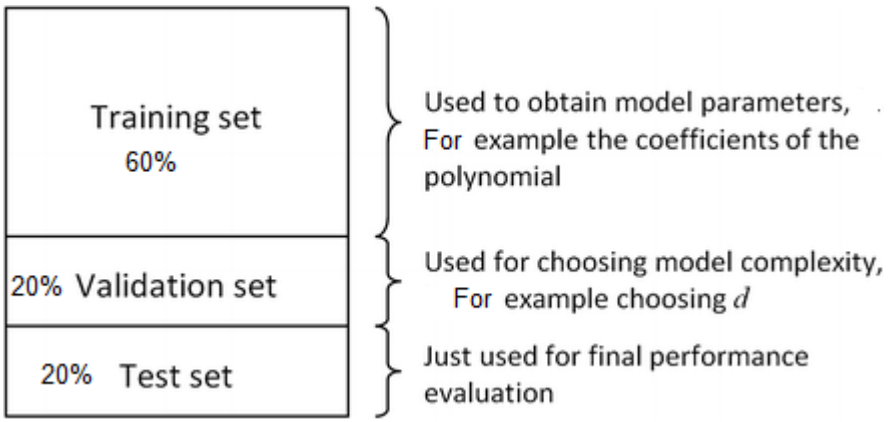


Figure B-15: Typical distribution the dataset for training, cross-validation, and test sets [28]

A hypothesis can have a low error for the training examples but still be inaccurate due to overfitting [27]. To visualise the performance of the hypothesis function, the training and cross-validation error need to be calculated. The method is as follows:

1. Train Θ and minimize the training set cost function [$J_{train}(\Theta)$].
2. Find the polynomial degree with the least error using the cross-validation set.
3. Compute the test set error [$J_{test}(\Theta)$].

The test set error is calculated as follows [27]:

Linear regression

$$J_{test}(\Theta) = \frac{1}{2m_{test}} \sum_{t=1}^{m_{test}} (h_\Theta(x_{test}^{(i)}) - y_{test}^{(i)})^2 \quad \text{(B-39)}$$

Logistic regression ~ Misclassification

$$err(h_\Theta(x), y) = 1 \; if \; h_\Theta(x) \geq 0.5 \; and \; y = 0 \; or \; h_\Theta(x) < 0.5 \; and \; y = 1$$
$$err(h_\Theta(x), y) = 0 \; otherwise \quad \text{(B-40)}$$

The error is evaluated by plotting a graph of error vs degree of the polynomial:

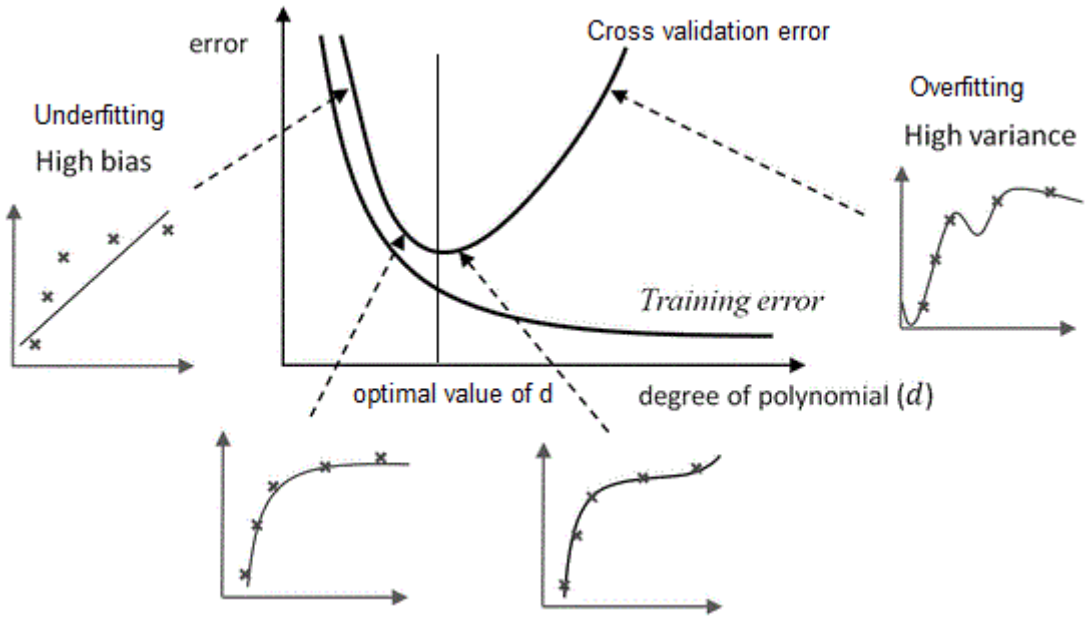


Figure B-16: Cross-validation and training error function vs degree of polynomial [28]

As indicated by Figure B-16, as the complexity of the model increases (from a simple linear fit to a high order polynomial), the training error decreases (model fits the data better). The cross-validation error decreases up to a point, which is the optimal value of d and then increases as the model becomes complex [28]. If both cross-validation and training error functions are high, then the model is suffering from a high bias problem. If the cross-validation error function is high while the training error function is low, the model is suffering from a high variance problem. Learning curves help us select which option to take to improve the accuracy of the model. By plotting error vs training set size curves, we can debug the model.

Figure B-17: Error vs training set size curves for a high bias model [27]

Figure B-18: Error vs training set size for a high variance model [27]

The options to take in order to improve the learning depends on the performance of the hypothesis function on the test data set. The decision process on selecting which option is as follows:

a) Getting more training examples: **Fixes high variance**
b) Trying smaller sets of features: **Fixes high variance**
c) Adding features: **Fixes high bias**
d) Adding polynomial features: **Fixes high bias**
e) Increasing λ: **Fixes high variance**
f) Decreasing λ: **Fixes high bias**

d) Recommended approach

It is important to understand the effects of model complexity:

- The simple model (lower model complexity: lower-order polynomials) has high bias and low variance. Therefore, the model will perform poorly consistently.
- Complicated model (high model complexity: higher-order polynomials) fits the training data very well and performs poorly on the test data. Thus, it has a low bias but high variance.

C. Appendix C – Case files data points

This appendix holds the data points discussed in section 3.2 and 4.

1. MV distribution network case files

1.1. Network load data

Table C-1 contains the MV distribution network actual half-hourly load data for one year (2015) (for complete data points, see attachment 1 "all data for ML.xlsx" sheet 1-7), Table C-2 contains three-phase and single-phase fault levels, section 1.2, section 1.3 and section 1.4 presents the distribution network DIgSILENT files with real network parameters.

Table C-1: Load data (30 min intervals)

Network A		Network C		Network B	
Date & Time	**kVA**	**Date & Time**	**kVA**	**Date & Time**	**kVA**
01 August 2015 00:00	101.52	01-Aug-2015 12:00 AM	291.00	8/1/2015 0:30	376.92
01 August 2015 00:30	94.81	01-Aug-2015 12:30 AM	278.04	8/1/2015 1:00	378.89
01 August 2015 01:00	96.18	01-Aug-2015 01:00 AM	266.05	8/1/2015 1:30	367.40
01 August 2015 01:30	92.13	01-Aug-2015 01:30 AM	268.00	8/1/2015 2:00	369.17
01 August 2015 02:00	93.48	01-Aug-2015 02:00 AM	262.00	8/1/2015 2:30	370.27
01 August 2015 02:30	96.18	01-Aug-2015 02:30 AM	269.09	8/1/2015 3:00	376.43
01 August 2015 03:00	92.13	01-Aug-2015 03:00 AM	258.02	8/1/2015 3:30	365.83
01 August 2015 03:30	93.02	01-Aug-2015 03:30 AM	249.01	8/1/2015 4:00	365.55
01 August 2015 04:00	96.18	01-Aug-2015 04:00 AM	241.00	8/1/2015 4:30	364.16
01 August 2015 04:30	94.02	01-Aug-2015 04:30 AM	246.00	8/1/2015 5:00	371.39
01 August 2015 05:00	98.62	01-Aug-2015 05:00 AM	253.00	8/1/2015 5:30	371.39
01 August 2015 05:30	95.46	01-Aug-2015 05:30 AM	242.00	8/1/2015 6:00	385.74
01 August 2015 06:00	99.02	01-Aug-2015 06:00 AM	259.00	8/1/2015 6:30	385.19
01 August 2015 06:30	115.13	01-Aug-2015 06:30 AM	283.00	8/1/2015 7:00	384.76
01 August 2015 07:00	130.36	01-Aug-2015 07:00 AM	291.00	8/1/2015 7:30	398.08
01 August 2015 07:30	130.51	01-Aug-2015 07:30 AM	282.00	8/1/2015 8:00	415.01
01 August 2015 08:00	130.00	01-Aug-2015 08:00 AM	294.00	8/1/2015 8:30	418.66
01 August 2015 08:30	135.16	01-Aug-2015 08:30 AM	315.00	8/1/2015 9:00	429.04
01 August 2015 09:00	138.32	01-Aug-2015 09:00 AM	326.00	8/1/2015 9:30	479.27
01 August 2015 09:30	146.74	01-Aug-2015 09:30 AM	305.00	8/1/2015 10:00	450.02
01 August 2015 10:00	131.32	01-Aug-2015 10:00 AM	310.00	8/1/2015 10:30	464.29
01 August 2015 10:30	123.66	01-Aug-2015 10:30 AM	298.00	8/1/2015 11:00	485.27
01 August 2015 11:00	119.85	01-Aug-2015 11:00 AM	314.00	8/1/2015 11:30	522.16
01 August 2015 11:30	116.81	01-Aug-2015 11:30 AM	338.00	8/1/2015 12:00	509.23
01 August 2015 12:00	127.24	01-Aug-2015 12:00 PM	318.00	8/1/2015 12:30	497.87
01 August 2015 12:30	120.34	01-Aug-2015 12:30 PM	301.00	8/1/2015 13:00	467.62
01 August 2015 13:00	122.54	01-Aug-2015 01:00 PM	290.00	8/1/2015 13:30	440.44
01 August 2015 13:30	121.64	01-Aug-2015 01:30 PM	284.00	8/1/2015 14:00	416.16
01 August 2015 14:00	113.74	01-Aug-2015 02:00 PM	285.00	8/1/2015 14:30	418.70
01 August 2015 14:30	111.14	01-Aug-2015 02:30 PM	290.00	8/1/2015 15:00	437.10
01 August 2015 15:00	112.77	01-Aug-2015 03:00 PM	287.00	8/1/2015 15:30	447.93

01 August 2015 15:30	102.96	01-Aug-2015 03:30 PM	295.00	8/1/2015 16:00	430.17
01 August 2015 16:00	101.97	01-Aug-2015 04:00 PM	298.00	8/1/2015 16:30	438.49
01 August 2015 16:30	111.49	01-Aug-2015 04:30 PM	306.00	8/1/2015 17:00	442.77
01 August 2015 17:00	99.40	01-Aug-2015 05:00 PM	282.00	8/1/2015 17:30	384.05
01 August 2015 17:30	102.88	01-Aug-2015 05:30 PM	294.00	8/1/2015 18:00	410.86
01 August 2015 18:00	110.39	01-Aug-2015 06:00 PM	315.00	8/1/2015 18:30	361.89
01 August 2015 18:30	125.87	01-Aug-2015 06:30 PM	358.00	8/1/2015 19:00	370.36
01 August 2015 19:00	115.43	01-Aug-2015 07:00 PM	382.00	8/1/2015 19:30	359.47
01 August 2015 19:30	114.02	01-Aug-2015 07:30 PM	382.00	8/1/2015 20:00	364.43
01 August 2015 20:00	122.78	01-Aug-2015 08:00 PM	373.00	8/1/2015 20:30	373.47
01 August 2015 20:30	124.31	01-Aug-2015 08:30 PM	360.00	8/1/2015 21:00	347.73
01 August 2015 21:00	122.43	01-Aug-2015 09:00 PM	340.00	8/1/2015 21:30	390.09
01 August 2015 21:30	125.36	01-Aug-2015 09:30 PM	324.00	8/1/2015 22:00	415.20
01 August 2015 22:00	122.43	01-Aug-2015 10:00 PM	287.00	8/1/2015 22:30	406.12
01 August 2015 22:30	114.02	01-Aug-2015 10:30 PM	276.00	8/1/2015 23:00	405.46
01 August 2015 23:00	97.51	01-Aug-2015 11:00 PM	277.00	8/1/2015 23:30	405.35
01 August 2015 23:30	91.24	01-Aug-2015 11:30 PM	261.00	8/2/2015	387.11
02 August 2015 00:00	86.77	02-Aug-2015 12:00 AM	261.00	8/2/2015 0:30	381.41
02 August 2015 00:30	74.24	02-Aug-2015 12:30 AM	247.00	8/2/2015 1:00	377.81
02 August 2015 01:00	77.37	02-Aug-2015 01:00 AM	241.00	8/2/2015 1:30	383.10
02 August 2015 01:30	82.37	02-Aug-2015 01:30 AM	224.00	8/2/2015 2:00	374.75
02 August 2015 02:00	85.00	02-Aug-2015 02:00 AM	233.00	8/2/2015 2:30	350.89
02 August 2015 02:30	78.29	02-Aug-2015 02:30 AM	238.00	8/2/2015 3:00	354.89
02 August 2015 03:00	84.08	02-Aug-2015 03:00 AM	236.00	8/2/2015 3:30	358.09
02 August 2015 03:30	84.08	02-Aug-2015 03:30 AM	234.00	8/2/2015 4:00	356.55
02 August 2015 04:00	83.63	02-Aug-2015 04:00 AM	233.00	8/2/2015 4:30	364.26
02 August 2015 04:30	85.49	02-Aug-2015 04:30 AM	229.00	8/2/2015 5:00	366.34
02 August 2015 05:00	84.31	02-Aug-2015 05:00 AM	244.00	8/2/2015 5:30	369.95
02 August 2015 05:30	87.32	02-Aug-2015 05:30 AM	234.00	8/2/2015 6:00	376.44
02 August 2015 06:00	86.16	02-Aug-2015 06:00 AM	231.00	8/2/2015 6:30	333.74
02 August 2015 06:30	88.02	02-Aug-2015 06:30 AM	251.00	8/2/2015 7:00	315.89
02 August 2015 07:00	91.76	02-Aug-2015 07:00 AM	257.00	8/2/2015 7:30	330.12
02 August 2015 07:30	109.90	02-Aug-2015 07:30 AM	287.00	8/2/2015 8:00	367.16
02 August 2015 08:00	113.76	02-Aug-2015 08:00 AM	321.00	8/2/2015 8:30	351.66
02 August 2015 08:30	133.20	02-Aug-2015 08:30 AM	338.00	8/2/2015 9:00	356.50
02 August 2015 09:00	151.80	02-Aug-2015 09:00 AM	326.00	8/2/2015 9:30	342.18
02 August 2015 09:30	134.17	02-Aug-2015 09:30 AM	311.00	8/2/2015 10:00	341.12
02 August 2015 10:00	127.09	02-Aug-2015 10:00 AM	297.00	8/2/2015 10:30	370.86
02 August 2015 10:30	122.80	02-Aug-2015 10:30 AM	298.00	8/2/2015 11:00	373.53
02 August 2015 11:00	123.97	02-Aug-2015 11:00 AM	288.00	8/2/2015 11:30	376.11
02 August 2015 11:30	129.03	02-Aug-2015 11:30 AM	290.00	8/2/2015 12:00	390.52
02 August 2015 12:00	121.12	02-Aug-2015 12:00 PM	265.00	8/2/2015 12:30	377.58
02 August 2015 12:30	112.54	02-Aug-2015 12:30 PM	287.00	8/2/2015 13:00	443.76
02 August 2015 13:00	118.85	02-Aug-2015 01:00 PM	292.00	8/2/2015 13:30	480.65
02 August 2015 13:30	103.32	02-Aug-2015 01:30 PM	291.00	8/2/2015 14:00	468.18
02 August 2015 14:00	104.65	02-Aug-2015 02:00 PM	278.00	8/2/2015 14:30	428.49

02 August 2015 14:30	98.49	02-Aug-2015 02:30 PM	279.00	8/2/2015 15:00	441.43
02 August 2015 15:00	95.26	02-Aug-2015 03:00 PM	272.00	8/2/2015 15:30	428.21
02 August 2015 15:30	98.84	02-Aug-2015 03:30 PM	303.00	8/2/2015 16:00	430.84
02 August 2015 16:00	107.94	02-Aug-2015 04:00 PM	287.00	8/2/2015 16:30	473.36
02 August 2015 16:30	115.16	02-Aug-2015 04:30 PM	291.00	8/2/2015 17:00	481.24
02 August 2015 17:00	123.49	02-Aug-2015 05:00 PM	299.00	8/2/2015 17:30	487.71
02 August 2015 17:30	126.49	02-Aug-2015 05:30 PM	293.00	8/2/2015 18:00	486.04
02 August 2015 18:00	133.24	02-Aug-2015 06:00 PM	330.00	8/2/2015 18:30	427.03
02 August 2015 18:30	146.56	02-Aug-2015 06:30 PM	358.00	8/2/2015 19:00	367.36
02 August 2015 19:00	155.77	02-Aug-2015 07:00 PM	371.00	8/2/2015 19:30	353.95
02 August 2015 19:30	166.48	02-Aug-2015 07:30 PM	351.00	8/2/2015 20:00	364.82
02 August 2015 20:00	145.46	02-Aug-2015 08:00 PM	348.00	8/2/2015 20:30	355.02
02 August 2015 20:30	147.72	02-Aug-2015 08:30 PM	343.00	8/2/2015 21:00	346.95
02 August 2015 21:00	132.23	02-Aug-2015 09:00 PM	332.00	8/2/2015 21:30	355.21
02 August 2015 21:30	125.80	02-Aug-2015 09:30 PM	315.00	8/2/2015 22:00	334.48
02 August 2015 22:00	111.89	02-Aug-2015 10:00 PM	296.00	8/2/2015 22:30	341.14
02 August 2015 22:30	102.00	02-Aug-2015 10:30 PM	257.00	8/2/2015 23:00	327.29
02 August 2015 23:00	97.62	02-Aug-2015 11:00 PM	227.00	8/2/2015 23:30	317.91
02 August 2015 23:30	88.19	02-Aug-2015 11:30 PM	215.00	8/3/2015	308.97
03 August 2015 00:00	86.77	03-Aug-2015 12:00 AM	223.00	8/3/2015 0:30	302.89
03 August 2015 00:30	88.59	03-Aug-2015 12:30 AM	216.00	8/3/2015 1:00	306.20
03 August 2015 01:00	85.00	03-Aug-2015 01:00 AM	200.00	8/3/2015 1:30	316.47
03 August 2015 01:30	89.07	03-Aug-2015 01:30 AM	208.00	8/3/2015 2:00	305.37

Table C-2: Three-phase and single-phase fault levels

Network A		**Network B**		**Network C**	
Fault Level Three Phase kA	**Fault Level Single Phase kA**	**Fault Level Three Phase kA**	**Fault Level Single Phase kA**	**Fault Level Three Phase kA**	**Fault Level Single Phase kA**
2.71	0.201	3.51	0.366	2.877	0.165

1.2. Network A

Figure C-1 shows the complete short circuit parameters of network A with RO=0.1, X0=1 X1=1 and R=0.1. The fault levels at each node (based on conductor/equipment rating and length) are shown in Table C-3. Applying considerations in section 4.3 and 4.4 to Figure 4-9, the pick-up current settings are shown in Table C-4 for each node. The three-phase and earth fault protection pick-up current settings outlined in section 4.4 are applied to network A and recorded in Figure C-3 and Figure C-2. For the furthest point on the backbone, the first auto-recloser was given a TMS of 0.01 and a TMS of 0.05 was assigned to the furthest point on the T-offs/branch and a TMS of 0.06 for a T-off of a T-off to start the grading calculations. A minimum grading margin of 200ms was used between the protection devices.

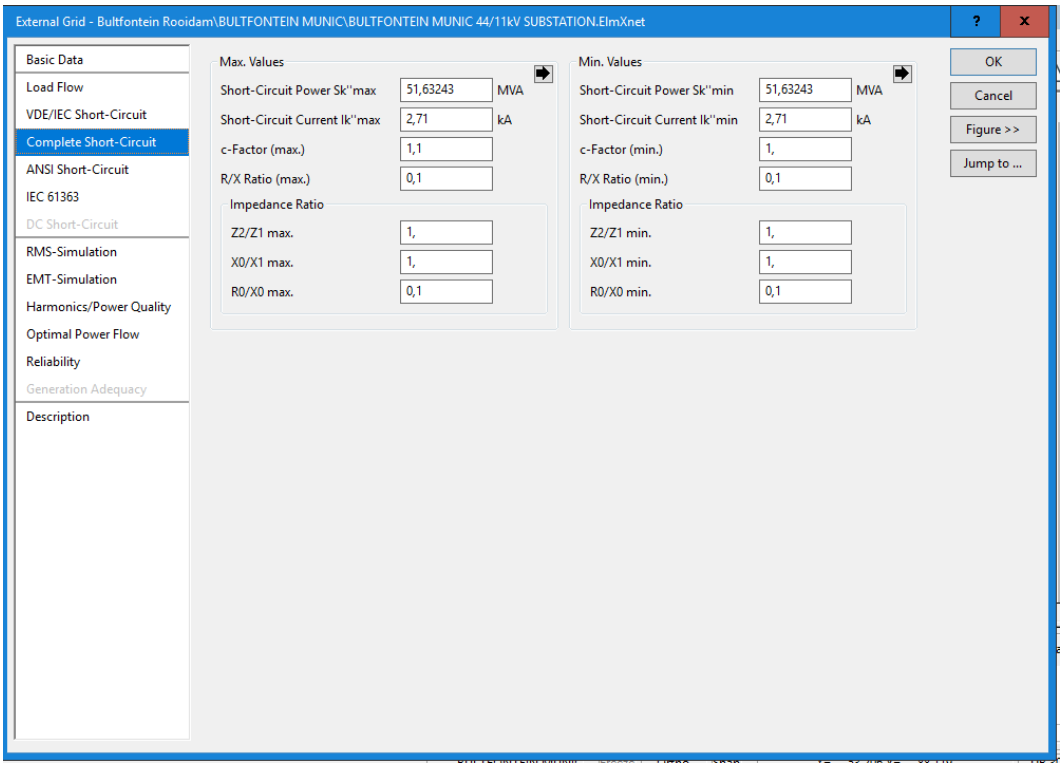


Figure C-1: Network A complete short circuit parameters

Table C-3: Fault levels at each node for network A

Fault levels at each node		
Node	**Three-phase fault level (A)**	**Single-phase fault level (A)**
65/238	211	164
65/176	264	204
65/164/34	243	188
65/164/1	277	214
65/138/81	225	174
65/138/36	266	206
65/138/1	311	240
65/128	328	254
65/127	330	255
65/67	473	365
65/64/93	286	221
65/64/3	473	365
65/39/62	379	293
65/39/58	379	293
65/39/56	393	304
65/39/53/12	373	288
65/39/53/1	398	307
65/39/54	398	307
65/39/39/12/5	393	304
65/39/39/1	435	336
65/39/2	582	445
65/39/1	587	453
65/2	888	684
65/1	901	693
209	350	257
148	519	370
114	630	459
111/43	409	316
111/1	555	428
67	895	688
65T/5	860	661
65T/1	907	698
48	1204	926
1	5207	5115
CB	2710	201

Table C-4: Three-phase and earth fault pick-up current settings for network A

Fault levels at each node		
Node	3phase Fault pick-up current setting (A)	Earth Fault pick-up current setting (A)
65/176	91.5057	25
65/164/1	91.5057	20
65/138/36	82.355	20
65/138/1	91.5057	25
65/128	101.673	30
65/127	112.97	35
65/67	125.522	40
65/64/3	125.522	40
65/39/58	82.3551	20
65/39/56	91.5057	25
65/39/53/1	91.5057	20
65/39/54	101.673	30
65/39/39/1	101.673	30
65/39/2	112.97	35
65/39/1	125.522	40
65/2	139.469	45
65/1	154.965	50
148	125.522	25
114	139.469	35
111/1	139.469	35
67	154.965	45
65T/1	154.965	45
48	172.183	65
1	**191.314**	80
CB	**212.571**	90

3 Phase Time Multiplier				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
148	0.01	370	25.00	0.03
114	0.11	370	35	0.33
114	0.11	459	35	0.30
67	0.14	459	45	0.60
67	0.14	688	45	0.36
48	0.23	688	65	0.66
48	0.23	926	65	0.58
1	0.32	926	80	0.88
1	0.32	5115	80	0.51
CB	0.49	5115	90	0.81
CB	0.49	201	90	4.21
Trf	0.15	201	100.00	1.50
Trf	0.15	201	100.00	1.50
* compare and use the slowest time				

T-offs 3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
111/1	0.05	428	35.00	0.14
67	0.14	428	45	0.44
67	0.14	688	45	0.36

T-offs 2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65T/1	0.05	698	45.00	0.12
48	0.23	698	65	0.42
48	0.23	926	65	0.58

T-off 1				
	TMS	Fault level	Pickup current	Operating time
65/176	0.05	204	25.00	0.16
65/128	0.13	204	30	0.46
65/128	0.13	254	30	0.41
65/127	0.21	254	35	0.71
65/127	0.21	255	35	0.71
65/67	0.27	255	40	1.01
65/67	0.27	365	40	0.85
65/2	0.35	365	45	1.15
65/2	0.35	684	45	0.88
65/1	0.45	684	50	1.18
65/1	0.45	693	50	1.17
48	0.23	693	65	1.47
48	0.23	926	65	0.58

T-offs 1/1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/39/58	0.06	293	20.00	0.15
65/39/56	0.16	293	25	0.45
65/39/56	0.16	304	25	0.45
65/39/54	0.25	304	30	0.75
65/39/54	0.25	307	30	0.74
65/39/2	0.33	307	35	1.04
65/39/2	0.33	445	35	0.89
65/39/1	0.42	445	40	1.19
65/39/1	0.42	453	40	1.18
65/2	0.35	453	45	1.48
65/2	0.35	684	45	0.88

T-offs 1/2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/64/3	0.06	365	40.00	0.19
65/2	0.35	365	45	0.49
65/2	0.35	684	45	0.88

T-offs 1/3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/138/1	0.06	240	25	0.18
65/128	0.13	240	30	0.48
65/128	0.13	254	30	0.41

T-offs 1/4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/164/1	0.06	214	20.00	0.17
65/128	0.13	214	30	0.47
65/128	0.13	254	30	0.41

T-offs 1/1/1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/39/39/1	0.08	336	30.00	0.23
65/39/2	0.33	336	35	0.53
65/39/2	0.33	445	35	0.89

T-offs 1/1/2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/39/53/1	0.08	307	20.00	0.20
65/39/54	0.25	307	30	0.50
65/39/54	0.25	307	30	0.74

Figure C-2: Earth fault operating times and TMS values for the protective devices on network A

3 Phase Time Multiplier				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
148	0.01	519	125.52	0.05
114	0.07	519	139.469	0.35
114	0.07	630	139.469	0.30
67	0.10	630	154.965	0.60
67	0.10	895	154.965	0.40
48	0.17	895	172.183	0.70
48	0.17	1204	172.183	0.59
1	0.24	1204	191.314	0.89
1	0.24	5207	191.314	0.49
CB	0.37	5207	212.571	0.79
CB	0.37	2710	212.571	1.00
Trf	0.54	2710	59.05	0.94
Trf	0.54	678	59.05	1.50
* compare and use the slowest time				

T-offs 3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
111/1	0.05	555	139.47	0.25
67	0.10	555	154.965	0.55
67	0.10	895	154.965	0.40

T-offs 2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65T/1	0.05	860	154.97	0.20
48	0.17	860	172.183	0.50
48	0.17	1204	172.183	0.59

T-off 1				
	TMS	Fault level	Pickup current	Operating time
65/176	0.05	264	91.51	0.33
65/128	0.09	264	101.673	0.63
65/128	0.09	328	101.673	0.51
65/127	0.12	328	112.97	0.81
65/127	0.12	330	112.97	0.80
65/67	0.15	330	125.522	1.10
65/67	0.15	473	125.522	0.80
65/2	0.19	473	139.469	1.10
65/2	0.19	888	139.469	0.72
65/1	0.26	888	154.965	1.02
65/1	0.26	901	154.965	1.01
48	0.17	901	172.183	1.31
48	0.17	1204	172.183	0.59

T-offs 1/1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/39/58	0.06	379	82.36	0.27
65/39/56	0.12	379	91.5057	0.57
65/39/56	0.12	393	91.5057	0.56
65/39/54	0.17	393	101.673	0.86
65/39/54	0.17	398	101.673	0.85
65/39/2	0.21	398	112.97	1.15
65/39/2	0.21	582	112.97	0.88
65/39/1	0.26	582	125.522	1.18
65/39/1	0.26	587	125.522	1.17
65/2	0.19	587	139.469	1.47
65/2	0.19	888	139.469	0.72

T-offs 1/2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/64/3	0.06	473	125.52	0.31
65/2	0.19	473	139.469	0.61
65/2	0.19	888	139.469	0.72

T-offs 1/3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/138/1	0.06	311	91.5057	0.34
65/128	0.09	311	101.673	0.64
65/128	0.09	328	101.673	0.51

T-offs 1/4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/164/1	0.06	277	91.51	0.38
65/128	0.09	277	101.673	0.68
65/128	0.09	328	101.673	0.51

T-offs 1/1/1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
65/39/39/1	0.08	435	101.67	0.38
65/39/2	0.21	435	112.97	0.68
65/39/2	0.21	582	112.97	0.88

T-offs 1/1/2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
665/39/53/1	0.08	398	91.51	0.38
65/39/54	0.17	398	101.673	0.68
65/39/54	0.17	398	101.673	0.85

Figure C-3:Three phase fault operating times & TMS values for the protective devices on network A

1.3. Network B

Figure C-4 shows the complete short circuit parameters of network B with RO=0.1, X0=1 X1=1 and R=0.1. The fault levels at each node (based on conductor/equipment rating and length) are shown in Table C-5. Applying considerations in section 4.3 and 4.4 to Figure 4-10, the pick-up current settings are shown in Table C-6 for each node. The three-phase and earth fault protection pick-up current settings outlined in section 4.4 are applied to network B and recorded in Figure C-5 and Figure C-6. For the furthest point on the backbone, the first auto-recloser was given a TMS of 0,01 and a TMS of 0.05 was assigned to the furthest point on the T-offs/branch and a TMS of 0.06 for a T-off of a T-off to start the grading calculations. A minimum grading margin of 200ms was used between the protection devices.

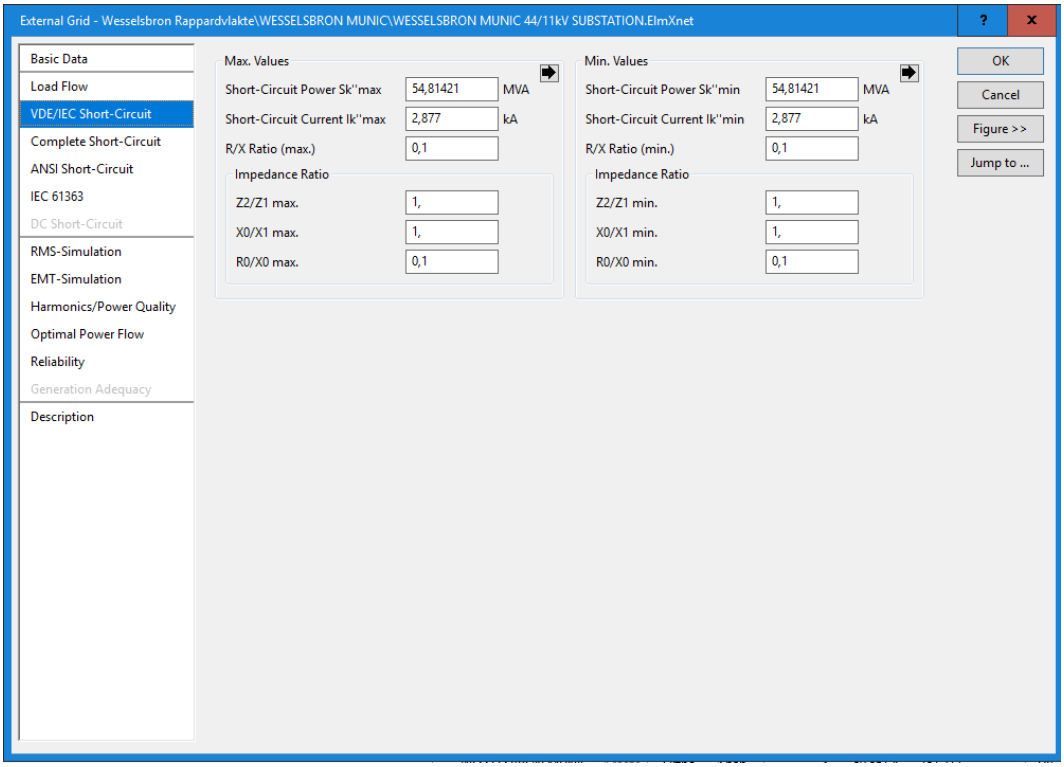


Figure C-4: Network B complete short circuit parameters

Table C-5: Fault levels at each node for network B

Fault levels at each node		
Node	**Three-phase fault level (A)**	**Single fault level(A)**
49/29/36	680	480
49/29/21	803	557
49/29/1	1052	709
49/63/47	465	338
49/63/11	623	442
49/110/14/10	513	334
49/110/1	621	392
49/140/17	437	291
49/140/1	487	320
49/147/9	440	293
49/147/1	464	307
49/166/7	487	295
49/166/1	509	306
214/102/10		
49/175/20	421	263
49/175/1	477	291
49/195/11	394	250
49/195/7/1	401	253
49/195/7	404	254
49/195/1	419	262
49/252/1	393	227
49/252/30/6	328	198
49/357/22	250	158
49/357/18/1	234	150
49/357/18	235	150
49/357/1	250	158
49/471/93	178	109
49/471/1	235	134
49/478/8/3	222	129
49/478/1	229	132
3		
327/69		
49/470	236	135
49/373	288	165
49/261	387	222
49/250	400	230
49/148	590	342
49/79	867	511
49/1	1833	1147
2/6/9	2695	2124
2/2	4649	4113

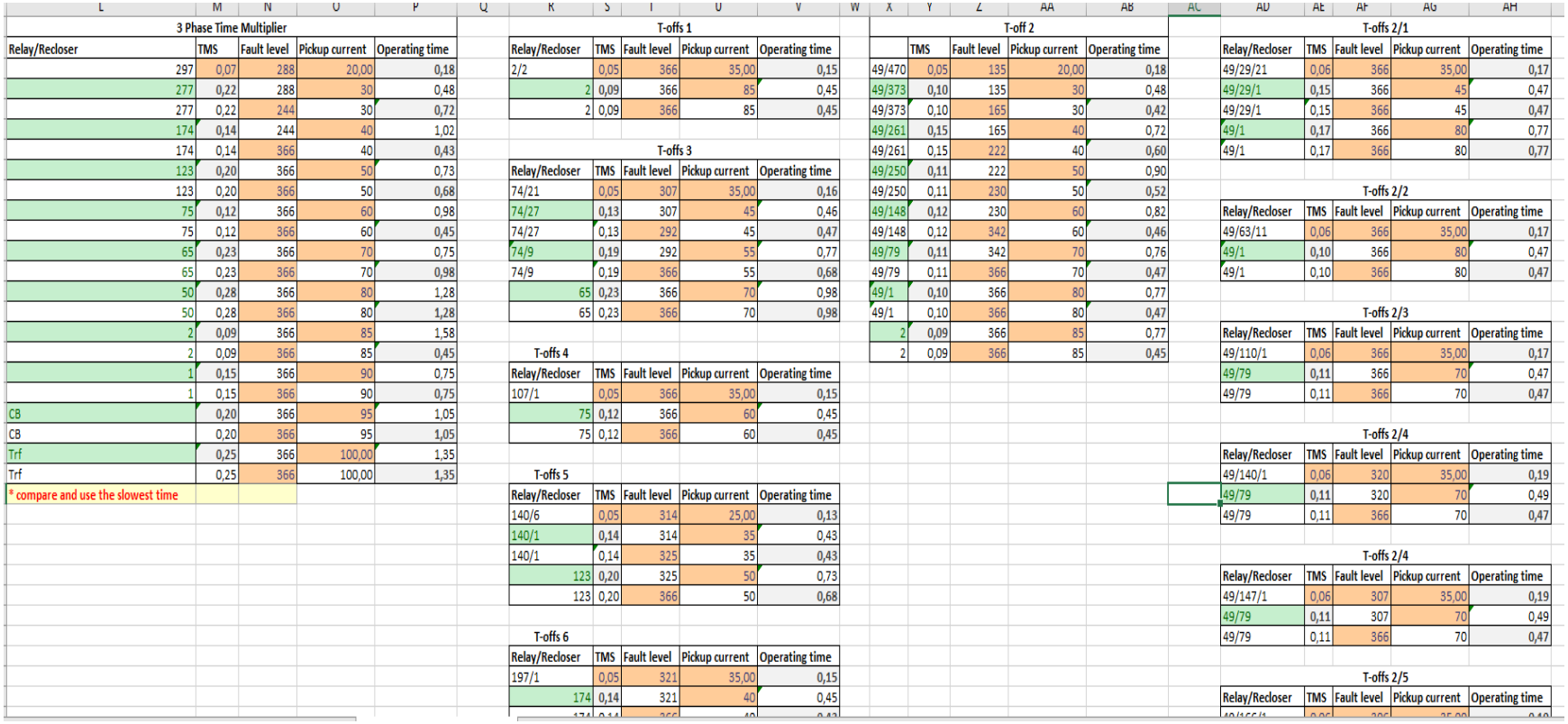
74/39/2	275	251
74/27	292	268
74/21	337	307
74/9	490	426
107/28	584	410
107/1	764	523
140/33	338	269
140/6	393	314
140/21	360	288
140/1	405	325
197/34/6	395	259
197/1	511	321
198/14	463	296
198/1	507	319
220/6/4	343	244
220/6/1	370	256
220/1	436	281
221/10	409	267
221/1	433	280
235/21	355	237
235/1	397	260
259/1	303	216
259/5/1	383	242
259/5/5	375	238
280/57/4	280	183
280/57/1	284	185
280/8/1	358	224
280/8/2	356	223
280/77	263	173
280/1	372	232
394/25/2	281	164
394/9	293	171
394/9/1	292	170
394/8/1	292	170
394/8	293	171
394/1	298	174
394/25/2	281	164
312/1	370	217
312/5	365	214
330/80	288	168
330/1	308	180
408/1	289	168
408/2	288	168
432	274	159
297	388	288

277	413	244
174	622	375
123	829	511
75	1208	777
65	1335	872
50	1833	1147
2	5057	4721
1	5215	5031
CB	3510	366

Table C-6: Three-phase and earth fault pick-up current settings for network B

Fault levels at each node		
Node	**Three-phase fault pick-up current setting (A)**	**Earth fault pick-up current setting (A)**
49/29/21	125.52	35
49/29/1	139.469	45
49/63/11	139.469	35
49/110/1	125.52	35
49/140/1	125.52	35
49/147/1	125.52	35
49/166/1	112.97	35
49/175/1	112.97	20
49/195/7	101.67	20
49/195/1	112.97	35
49/252/1	101.67	35
49/357/18	82.36	20
49/357/1	91.506	30
49/471/1	74.12	15
49/478/1	74.12	15
49/470	82.36	20
49/373	91.506	30
49/261	101.673	40
49/250	112.97	50
49/148	125.522	60
49/79	139.469	70
49/1	154.965	80
2/2	154.97	35
74/27	112.963	45
74/21	101.67	35
74/9	125.52	55
107/1	112.97	35
140/6	91.51	25
140/1	101.67	35
197/1	91.51	35
198/1	91.51	35

220/6/1	82.36	25
220/1	91.51	35
221/1	91.51	35
235/1	91.51	35
259/1	91.51	35
259/5/1	82.35	25
280/57/1	74.12	20
280/1	82.35	25
394/9	66.71	20
394/8	66.71	20
394/1	74.12	25
312/1	74.12	20
330/1	74.12	20
408/1	74.12	25
297	82.355	20
277	91.5057	30
174	101.673	40
123	112.97	50
75	125.522	60
65	139.469	70
50	154.965	80
2	172.183	85
1	**191.314**	90
CB	**212.571**	95



3 Phase Time Multiplier				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
297	0,07	288	20,00	0,18
277	0,22	288	30	0,48
277	0,22	244	30	0,72
174	0,14	244	40	1,02
174	0,14	366	40	0,43
123	0,20	366	50	0,73
123	0,20	366	50	0,68
75	0,12	366	60	0,98
75	0,12	366	60	0,45
65	0,23	366	70	0,75
65	0,23	366	70	0,98
50	0,28	366	80	1,28
50	0,28	366	80	1,28
2	0,09	366	85	1,58
2	0,09	366	85	0,45
1	0,15	366	90	0,75
1	0,15	366	90	0,75
CB	0,20	366	95	1,05
CB	0,20	366	95	1,05
Trf	0,25	366	100,00	1,35
Trf	0,25	366	100,00	1,35
* compare and use the slowest time				

T-offs 1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
2/2	0,05	366	35,00	0,15
2	0,09	366	85	0,45
2	0,09	366	85	0,45

T-offs 3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
74/21	0,05	307	35,00	0,16
74/27	0,13	307	45	0,46
74/27	0,13	292	45	0,47
74/9	0,19	292	55	0,77
74/9	0,19	366	55	0,68
65	0,23	366	70	0,98
65	0,23	366	70	0,98

T-offs 4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
107/1	0,05	366	35,00	0,15
75	0,12	366	60	0,45
75	0,12	366	60	0,45

T-offs 5				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
140/6	0,05	314	25,00	0,13
140/1	0,14	314	35	0,43
140/1	0,14	325	35	0,43
123	0,20	325	50	0,73
123	0,20	366	50	0,68

T-offs 6				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
197/1	0,05	321	35,00	0,15
174	0,14	321	40	0,45

T-off 2				
	TMS	Fault level	Pickup current	Operating time
49/470	0,05	135	20,00	0,18
49/373	0,10	135	30	0,48
49/373	0,10	165	30	0,42
49/261	0,15	165	40	0,72
49/261	0,15	222	40	0,60
49/250	0,11	222	50	0,90
49/250	0,11	230	50	0,52
49/148	0,12	230	60	0,82
49/148	0,12	342	60	0,46
49/79	0,11	342	70	0,76
49/79	0,11	366	70	0,47
49/1	0,10	366	80	0,77
49/1	0,10	366	80	0,47
2	0,09	366	85	0,77
2	0,09	366	85	0,45

T-offs 2/1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/29/21	0,06	366	35,00	0,17
49/29/1	0,15	366	45	0,47
49/29/1	0,15	366	45	0,47
49/1	0,17	366	80	0,77
49/1	0,17	366	80	0,77

T-offs 2/2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/63/11	0,06	366	35,00	0,17
49/1	0,10	366	80	0,47
49/1	0,10	366	80	0,47

T-offs 2/3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/110/1	0,06	366	35,00	0,17
49/79	0,11	366	70	0,47
49/79	0,11	366	70	0,47

T-offs 2/4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/140/1	0,06	320	35,00	0,19
49/79	0,11	320	70	0,49
49/79	0,11	366	70	0,47

T-offs 2/4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/147/1	0,06	307	35,00	0,19
49/79	0,11	307	70	0,49
49/79	0,11	366	70	0,47

T-offs 2/5				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time

Figure C-5: Earth fault operating times and TMS values for the protective devices on network B

3 Phase Time Multiplier

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
297	0,07	388	82,36	0,31
277	0,16	388	91,5057	0,61
277	0,16	413	91,5057	0,75
174	0,12	413	101,673	1,05
174	0,12	622	101,673	0,44
123	0,15	622	112,97	0,74
123	0,15	829	112,97	0,52
75	0,13	829	125,522	0,82
75	0,13	1208	125,522	0,38
65	0,16	1208	139,469	0,68
65	0,16	1335	139,469	0,47
50	0,24	1335	154,965	0,77
50	0,24	1833	154,965	0,67
2	0,18	1833	172,183	0,97
2	0,18	3510	172,183	0,41
1	0,30	3510	191,314	0,71
1	0,30	3510	191,314	0,71
CB	0,42	3510	212,571	1,01
CB	0,42	3510	212,571	1,01
Trf	0,80	3510	59,05	1,31
Trf	0,80	788	59,05	2,10
* compare and use the slowest time				

T-offs 1

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
2/2	0,05	3510	154,97	0,11
2	0,18	3510	172,183	0,41
2	0,18	3510	172,183	0,41

T-offs 3

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
74/21	0,05	490	101,67	0,22
74/27	0,11	490	112,963	0,52
74/27	0,11	337	112,963	0,70
74/9	0,14	337	125,522	1,00
74/9	0,14	292	125,522	1,17
65	0,16	292	139,469	1,47
65	0,16	1335	139,469	0,47

T-offs 4

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
107/1	0,05	764	112,97	0,18
75	0,13	764	125,522	0,48
75	0,13	1208	125,522	0,38

T-offs 5

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
140/6	0,05	393	91,51	0,24
140/1	0,11	393	101,673	0,54
140/1	0,11	405	101,673	0,52
123	0,15	405	112,97	0,82
123	0,15	829	112,97	0,52

T-offs 6

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
197/1	0,05	511	91,51	0,20
174	0,12	511	101,673	0,50

T-off 2

	TMS	Fault level	Pickup current	Operating time
49/470	0,05	236	82,36	0,33
49/373	0,09	236	91,5057	0,63
49/373	0,09	288	91,5057	0,52
49/261	0,12	288	101,673	0,82
49/261	0,12	387	101,673	0,64
49/250	0,11	387	112,97	0,94
49/250	0,11	400	112,97	0,60
49/148	0,12	400	125,522	0,90
49/148	0,12	590	125,522	0,52
49/79	0,12	590	139,469	0,82
49/79	0,12	867	139,469	0,45
49/1	0,12	867	154,965	0,75
49/1	0,12	1833	154,965	0,32
2	0,18	1833	172,18	0,62
2	0,18	3510	172,18	0,41

T-offs 2/1

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/29/21	0,06	803	125,52	0,22
49/29/1	0,13	803	139,469	0,52
49/29/1	0,13	1052	139,469	0,45
49/1	0,21	1052	154,965	0,75
49/1	0,21	1833	154,965	0,58

T-offs 2/2

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/63/11	0,06	623	139,47	0,28
49/1	0,12	623	154,965	0,58
49/1	0,12	1833	154,965	0,32

T-offs 2/3

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/110/1	0,06	621	125,52	0,26
49/79	0,12	621	139,469	0,56
49/79	0,12	867	139,469	0,45

T-offs 2/4

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/140/1	0,06	487	125,52	0,31
49/79	0,12	487	139,469	0,61
49/79	0,12	867	139,469	0,45

T-offs 2/4

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
49/147/1	0,06	464	125,52	0,32
49/79	0,12	464	139,469	0,62
49/79	0,12	867	139,469	0,45

T-offs 2/5

Relay/Recloser	TMS	Fault level	Pickup current	Operating time

Figure C-6: Three-phase Operating times and TMS values for the protective devices on network B

1.4. Network C

Figure C-7 shows the complete short circuit parameters of network C with RO=0.1, X0=1 X1=1 and R=0.1. The fault levels at each node (based on conductor/equipment rating and length) are shown in Table C-7. Applying considerations in section 4.3 and 4.4 to Figure 4-11, the pick-up current settings are shown in Table C-8 for each node. The three-phase and earth fault protection pick-up current settings outlined in section 4.4 are applied to network C and recorded in Figure C-8 and Figure C-9. For the furthest point on the backbone, the first auto-recloser was given a TMS of 0,01 and a TMS of 0.05 was assigned to the furthest point on the T-offs/branch and a TMS of 0.06 for a T-off of a T-off to start the grading calculations. A minimum grading margin of 200ms was used between the protection devices.

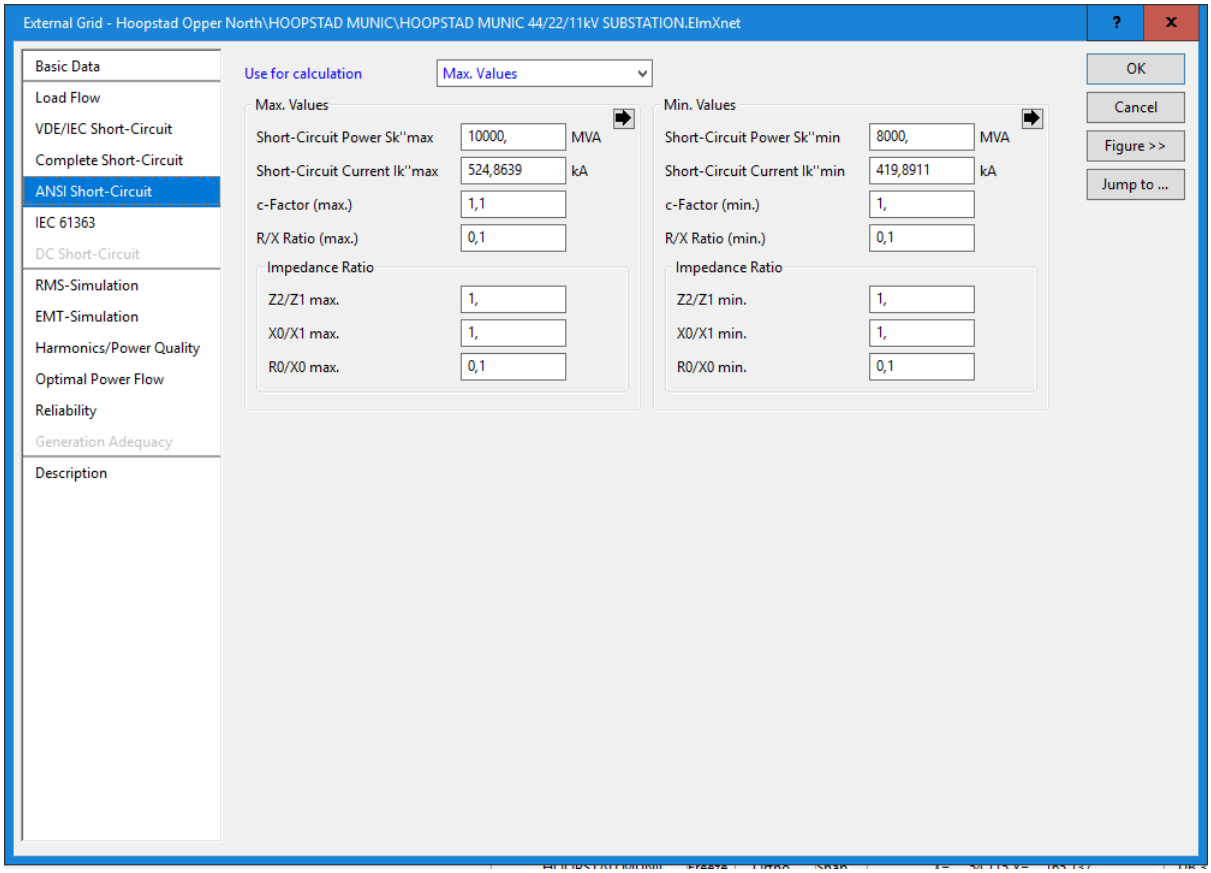


Figure C-7: Network C complete short circuit parameters

Table C-7: Fault levels at each node for network C

Fault levels at each node		
Node	**Three-phase fault level (A)**	**Single fault level(A)**
86/15	624	478
86/1	719	551
61/42	612	470
61/40/5	595	457
61/40/1	618	474
61/7/8	818	626
61/7/1	895	684
61/2	974	744
111/131	266	205
111/98	307	237
111/29/3	446	343
111/29/1	452	347
111/1	565	433
200/1	390	278
200/26/22	305	221
146/58	384	273
146/1	572	393
146/12/8	492	343
146/12/1	519	360
314/1	261	183
314/40	226	161
304/97	195	140
304/3	270	188
356/1	265	174
387/4	233	156
387/1	236	158
339	285	185
202	468	305
113	801	528
5	4588	4183
4	4742	4445
3	4889	4650
2	5044	4873
CB	2877	165

Table C-8: Three-phase and earth fault pick-up current settings for network C

Fault levels at each node		
Node	**Three-phase fault pick-up current setting (A)**	**Earth fault pick-up current setting (A)**
86/1	125.52	35
61/40/1	112.97	30
61/7/1	112.97	30
61/2	125.52	40
111/98	112.97	35
111/29/1	112.97	30
111/1	125.522	45
200/1	112.97	25
146/1	154.97	25
146/12/1	101.67	20
314/1	91.51	20
304/3	**91.51**	25
356/1	**91.51**	20
387/1	**91.51**	20
339	**101,67**	35
202	**112.97**	40
113	**125.522**	50
5	**139.469**	60
4	**154.965**	70
3	**172.183**	80
2	**191.314**	85
CB	**212.571**	90

3 Phase Time Multiplier				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
339	0,01	165	35,00	0,04
202	0,07	165	40	0,34
202	0,07	165	40	0,34
113	0,11	165	50	0,64
113	0,11	165	50	0,64
5	0,08	165	60	0,94
5	0,08	165	60	0,54
4	0,10	165	70	0,84
4	0,10	165	70	0,84
3	0,12	165	80	1,14
3	0,12	165	80	1,14
2	0,14	165	85,00	1,44
2	0,14	165	85	1,44
CB	0,15	165	90,00	1,74
CB	0,15	165	90	1,74
Trf	0,15	165	100,00	2,04
Trf	0,15	165	100,00	2,04
* compare and use the slowest time				

T-offs 1				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
61/2	0,05	165	40,00	0,24
5	0,08	165	60	0,54
5	0,08	165	60	0,54

T-offs 2				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
86/1	0,05	165	35,00	0,22
5	0,08	165	60	0,52
5	0,08	165	60	0,54

T-offs 3				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
111/98	0,05	165	35,00	0,22
111/1	0,16	165	45	0,84
111/1	0,16	165	45	0,84
5	0,08	165	60	1,14
5	0,08	165	60	0,54

T-offs 4				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
146/1	0,05	165	25,00	0,18
113	0,08	165	50	0,48
113	0,08	165	50	0,48

T-offs 5				
Relay/Recloser	TMS	Fault level	Pickup current	Operating time
200/1	0,05	165	25,00	0,18

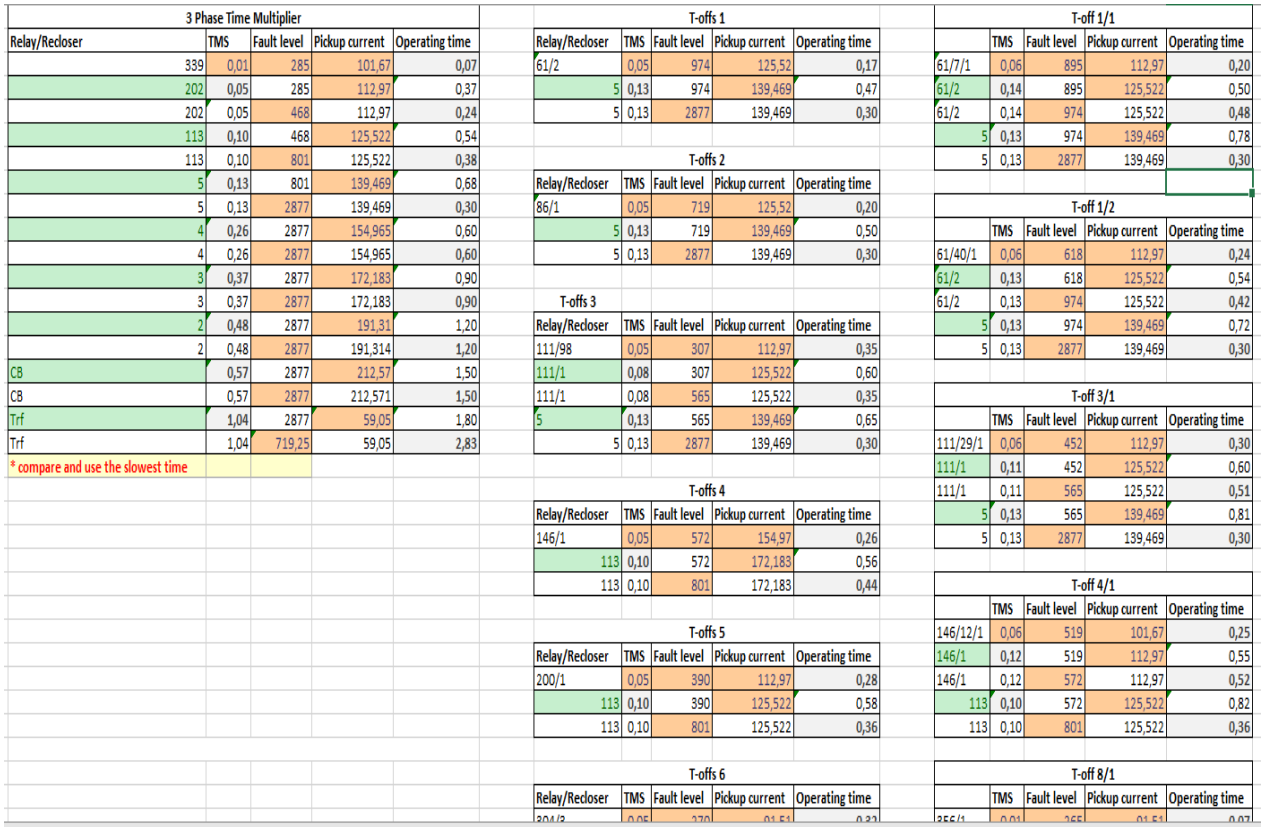
T-off 1/1				
	TMS	Fault level	Pickup current	Operating time
61/7/1	0,06	165	30,00	0,24
61/2	0,11	165	40	0,54
61/2	0,11	165	40	0,54
5	0,08	165	60	0,84
5	0,08	165	60	0,54

T-off 1/2				
	TMS	Fault level	Pickup current	Operating time
61/40/1	0,06	165	30,00	0,24
61/2	0,11	165	40	0,54
61/2	0,11	165	40	0,54
5	0,08	165	60	0,84
5	0,08	165	60	0,54

T-off 3/1				
	TMS	Fault level	Pickup current	Operating time
111/29/1	0,06	165	30,00	0,24
111/1	0,10	165	45	0,54
111/1	0,10	165	45	0,54
5	0,08	165	60	0,84
5	0,08	165	60	0,54

T-off 4/1				
	TMS	Fault level	Pickup current	Operating time
146/12/1	0,06	165	20,00	0,19
146/1	0,14	165	25	0,49
146/1	0,14	165	25	0,49

Figure C-8: Earth fault operating times and TMS values for the protective devices on network C



3 Phase Time Multiplier

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
339	0,01	285	101,67	0,07
202	0,05	285	112,97	0,37
202	0,05	468	112,97	0,24
113	0,10	468	125,522	0,54
113	0,10	801	125,522	0,38
5	0,13	801	139,469	0,68
5	0,13	2877	139,469	0,30
4	0,26	2877	154,965	0,60
4	0,26	2877	154,965	0,60
3	0,37	2877	172,183	0,90
3	0,37	2877	172,183	0,90
2	0,48	2877	191,31	1,20
2	0,48	2877	191,314	1,20
CB	0,57	2877	212,57	1,50
CB	0,57	2877	212,571	1,50
Trf	1,04	2877	59,05	1,80
Trf	1,04	719,25	59,05	2,83
* compare and use the slowest time				

T-offs 1

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
61/2	0,05	974	125,52	0,17
5	0,13	974	139,469	0,47
5	0,13	2877	139,469	0,30

T-offs 2

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
86/1	0,05	719	125,52	0,20
5	0,13	719	139,469	0,50
5	0,13	2877	139,469	0,30

T-offs 3

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
111/98	0,05	307	112,97	0,35
111/1	0,08	307	125,522	0,60
111/1	0,08	565	125,522	0,35
5	0,13	565	139,469	0,65
5	0,13	2877	139,469	0,30

T-offs 4

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
146/1	0,05	572	154,97	0,26
113	0,10	572	172,183	0,56
113	0,10	801	172,183	0,44

T-offs 5

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
200/1	0,05	390	112,97	0,28
113	0,10	390	125,522	0,58
113	0,10	801	125,522	0,36

T-offs 6

Relay/Recloser	TMS	Fault level	Pickup current	Operating time
304/2	0,05	270	91,51	0,32

T-off 1/1

	TMS	Fault level	Pickup current	Operating time
61/7/1	0,06	895	112,97	0,20
61/2	0,14	895	125,522	0,50
61/2	0,14	974	125,522	0,48
5	0,13	974	139,469	0,78
5	0,13	2877	139,469	0,30

T-off 1/2

	TMS	Fault level	Pickup current	Operating time
61/40/1	0,06	618	112,97	0,24
61/2	0,13	618	125,522	0,54
61/2	0,13	974	125,522	0,42
5	0,13	974	139,469	0,72
5	0,13	2877	139,469	0,30

T-off 3/1

	TMS	Fault level	Pickup current	Operating time
111/29/1	0,06	452	112,97	0,30
111/1	0,11	452	125,522	0,60
111/1	0,11	565	125,522	0,51
5	0,13	565	139,469	0,81
5	0,13	2877	139,469	0,30

T-off 4/1

	TMS	Fault level	Pickup current	Operating time
146/12/1	0,06	519	101,67	0,25
146/1	0,12	519	112,97	0,55
146/1	0,12	572	112,97	0,52
113	0,10	572	125,522	0,82
113	0,10	801	125,522	0,36

T-off 8/1

	TMS	Fault level	Pickup current	Operating time
356/1	0,01	265	91,51	0,07

Figure C-9: Three-phase Operating times and TMS values for the protective devices on network C

2. PV DG Data

The DG data consists of; postal code, city, PV name, kW rating, date of installation (Table C-9), and actual power generation of a 5.3 kW PV system recorded in 15-minute intervals over a year (Table C-10) [28]. For complete data points, see attachment 1 “all data for ML.xlsx” sheet 1-7.

Table C-9: PV systems data publicly available on Sunny Portal [28]

PV System Name	Country	Zip Code	City	Power (kWp)	Date of installation
Rozanne	South Africa	1	Centurion	2.80	20/02/2015
Vettie	South Africa	1	Pretoria	5.30	
LG BABERTON	South Africa	1	Barberton	15.90	07/05/2011
15kWp-Fairview-Barberton	South Africa	1	Fairview, Barberton	15.90	10/07/2011
Lephalale Canadian Farm	South Africa	14	Lephalale	93.60	11/08/2018
Berndt Merkman	South Africa	27	Kensington	5.30	21/11/2015
HS Van Der Walt	South Africa	27	wesselsbron	5.30	
Tandt Family	South Africa	27	PRETORIA	5.30	
ICAT	South Africa	27	Pretoria	51.50	10/02/2015
Gal01	South Africa	36	Rhenosterfontein	8.20	
Old Farm School	South Africa	40	Pretoria	9.04	
HENKPV	South Africa	46	Olifantsfontein	3.20	
7 Lock Street	South Africa	54	Pretoria	16.30	07/10/2015
Plot 123	South Africa	56	Pretoria	5.30	
Warrick	South Africa	59	Mooikloof	14.50	24/06/2017
Brother South Africa-Centurion	South Africa	67	Centurion	28.05	26/10/2015
Delta BEC Greenhouse MenloPark	South Africa	81	Pretoria	1.55	03/12/2017
45 Gift Acres	South Africa	81	Pretoria	4.20	29/07/2015
Verryn	South Africa	81	Pretoria	5.00	13/05/2015
Sunnyboy 5000	South Africa	81	Pretoria	5.30	
Genpower Solar	South Africa	81	Pretoria	7.92	
Sargon House	South Africa	81	Pretoria	19.62	04/01/2016
SB5000	South Africa	82	Pretoria	5.00	06/11/2016
UDP	South Africa	82	Pretoria	205.25	04/11/2016
R&G Sheetmetal	South Africa	100	JHB	154.50	15/12/2015
20 Golf Close	South Africa	123	Pretoria	8.25	

Woodhill	South Africa	127	Pretoria	7.28	03/07/2015
Raslouw Dierekliniek	South Africa	157	Centurion	2.50	10/05/2015
House Booysen	South Africa	157	Centurion	2.55	
40 Koggelaar	South Africa	157	Pretoria	2.97	08/01/2018
Plaas SMA FV	South Africa	157	Centurion	3.30	19/12/2017
Steinmann Heritage Hill House	South Africa	157	Centurion	4.20	06/10/2018
4375 Southdowns Estate	South Africa	157	Irene	4.88	13/12/2018
2605 Southdowns Estate	South Africa	157	Irene	5.17	13/12/2018
HJ Blair 14	South Africa	157	Centurion	5.30	27/01/2018
5kW SMA - Gavin Polson	South Africa	157	Constantia Park	5.30	11/03/2016
Herman vd Merwe	South Africa	157	Centurion	8.22	17/06/2015
Ampath Centurion	South Africa	157	Centurion	360.50	13/07/2016
51 Centurion Drive	South Africa	158	Centurion	5.04	14/09/2013
FINANCIAL ALLIANCE	South Africa	172	Johannesburg	7.60	
Nortje PV plant	South Africa	181	Pretoria	5.00	30/09/2014
RWeiss_5kw	South Africa	181	Pretoria	5.40	03/05/2016
EndlessPower Development	South Africa	181	Pretoria	8.40	
Cameron Marais	South Africa	181	Pretoria	16.90	30/09/2015
LUIS MARQUES	South Africa	182	Montana	0.00	21/09/2017
Sunny Boy 4.0	South Africa	182	Pretoria	4.20	16/12/2017
SAR Electronic SA (Pty) Ltd.	South Africa	182	Pretoria	10.80	14/09/2010
Heatherdale Solar System	South Africa	182	Heatherdale	16.50	01/10/2017
Total Annlin	South Africa	182	Annlin	27.47	16/01/2018
Diogenes Barrel	South Africa	184	Pretoria	3.40	13/09/2014
DII PTA	South Africa	184	Pretoria	109.20	
GSK Consulting Engineering	South Africa	186	Pretoria	10.20	15/01/2010
MULTICHOICE SAMRAND	South Africa	187	Centurion	224.00	06/01/2018
SMA Demo	South Africa	216	Hartbeespoort	8.60	08/04/2017
Kalkheuvel	South Africa	240	Mogale City	3.20	10/02/2013
Lift Master Harties	South Africa	250	Brits	103.00	
Midas MidTown	South Africa	299	Rustenburg	16.11	24/01/2018
MAEMO CV Rustenburg	South Africa	299	RUSTENBURG	30.60	12/10/2015

Willie Jones	South Africa	348	Koster	10.71	10/05/2015
Willie Jones HM	South Africa	348	Koster	10.71	10/05/2015
Jones Masjiene HM	South Africa	348	Koster	42.84	10/02/2015
Jones Masjiene	South Africa	348	Koster	42.84	10/02/2015
Projexconsult 30kW SB5s	South Africa	350	Kroondal	33.21	05/08/2015
Ottermann Farms	South Africa	350	Kroondal	103.00	12/02/2016
Warmbad Rusoord - Grid tie	South Africa	480	Bella Bella	53.76	14/12/2016
Hennie Fourie	South Africa	600	Mokopane	5.20	
Fruit & Veg City Mokopane	South Africa	600	Mokopane	103.00	23/01/2017
PEV Land Polokwane	South Africa	699	Polokwane	6.20	29/05/2014
SBSA 14 Thorpe	South Africa	1000	Johannesburg	235.95	03/10/2018
KragDag	South Africa	1001	Cullinan	27.75	26/10/2016
Lepelle-Phalaborwa Plant	South Africa	1014	Phalaborwa	39.89	12/01/2016
South Africa		1020	Bronkhorstspruit	28.00	
Hilton Glenn	South Africa	1200	Nelspruit	2.75	08/01/2015
Patel Residence	South Africa	1200	Nelspruit	15.39	14/01/2014
Lionsrock Workshop	South Africa	1234	Bethal	33.00	18/10/2018
lhollman@iafrica.com	South Africa	1320	Malelane	10.40	24/04/2015
Tshokwane Battery Containers	South Africa	1350	Tshokwane Picnic Site, Kruger National Park	0.00	27/02/2018
Nkuhlu Battery Container	South Africa	1350	Nkuhlu Picnic Site, Kruger National Park	0.00	31/08/2018
Nkuhlu Off-Grid System	South Africa	1350	Nkuhlu Picnic Site, Kruger National Park	130.65	30/08/2018
Tshokwane Off-Grid system	South Africa	1350	Tshokwane Picnic Site, Kruger National Park	404.30	
TSR PV PLANT	South Africa	1378	Elandsfontein	289.85	06/06/2016
Dale Pudney	South Africa	1401	Johannesburg	5.00	24/05/2018
AL PV Carports	South Africa	1451	Alberton	35.20	04/01/2013
AL South	South Africa	1451	Alrode	266.50	19/03/2014
Gregory Bishop	South Africa	1459	Boksburg	5.30	19/10/2018
EnercoGen	South Africa	1459	Johannesburg	6.30	20/08/2018
Turbo Fastners	South Africa	1459	Boksburg	41.10	09/11/2018
Fountain Liquor Boksburg	South Africa	1460	Boksburg	11.28	25/01/2013
francois@fgelectrical.co.za	South Africa	1501	Benoni	7.20	16/03/2018
Grantham Air	South Africa	1501	Benoni	13.50	05/06/2013

Bluesun Benoni	South Africa	1501	Benoni	17.00	10/07/2014
Billy Jacobs PV plant	South Africa	1559	SPRINGS	2.65	30/11/2017
15 Tony Street, Harmelia.	South Africa	1609	Germiston	2.60	06/03/2015
Voestalpine VAE South Africa	South Africa	1609	Isando	352.92	08/06/2014
Improvon Ext 11	South Africa	1610	Edenvale	14.85	14/05/2013
Cylinder Test Building	South Africa	1619	Kempton Park	100.00	
Total Hyper	South Africa	1666	Olifantsfontein	28.57	28/10/2018
2466 Waterfall Country Village	South Africa	1682	Midrand	2.24	13/02/2018
995 Waterfall Country Estate	South Africa	1682	Midrand	4.40	22/02/2019

Table C-10: 5.3 kW PV full day output data

Date	Time	Output Power (kW)
14/08/2015	6:45 AM -7:00 AM	0.01 -0.08
14/08/2015	7:00 AM -7:15 AM	0.08 -0.15
14/08/2015	7:15 AM -7:30 AM	0.15 -0.43
14/08/2015	7:30 AM -7:45 AM	0.43 -0.71
14/08/2015	7:45 AM -8:00 AM	0.71 -0.99
14/08/2015	8:00 AM -8:15 AM	0.99 -1.25
14/08/2015	8:15 AM -8:30 AM	1.25 -0.86
14/08/2015	8:30 AM -8:45 AM	0.86 -1.01
14/08/2015	8:45 AM -9:00 AM	1.01 -1.77
14/08/2015	9:00 AM -9:15 AM	1.77 -2.45
14/08/2015	9:15 AM -9:30 AM	2.45 -3.09
14/08/2015	9:30 AM -9:45 AM	3.09 -3.27
14/08/2015	9:45 AM -10:00 AM	3.27 -2.50
14/08/2015	10:00 AM -10:15 AM	2.50 -3.15
14/08/2015	10:15 AM -10:30 AM	3.15 -3.53
14/08/2015	10:30 AM -10:45 AM	3.53 -3.92
14/08/2015	10:45 AM -11:00 AM	3.92 -3.18
14/08/2015	11:00 AM -11:15 AM	3.18 -3.56
14/08/2015	11:15 AM -11:30 AM	3.56 -3.72
14/08/2015	11:30 AM -11:45 AM	3.72 -3.78
14/08/2015	11:45 AM -12:00 PM	3.78 -3.53
14/08/2015	12:00 PM -12:15 PM	3.53 -3.84
14/08/2015	12:15 PM -12:30 PM	3.84 -2.51
14/08/2015	12:30 PM -12:45 PM	2.51 -3.63
14/08/2015	12:45 PM -1:00 PM	3.63 -3.75
14/08/2015	1:00 PM -1:15 PM	3.75 -4.23
14/08/2015	1:15 PM -1:30 PM	4.23 -2.15
14/08/2015	1:30 PM -1:45 PM	2.15 -2.72

14/08/2015	1:45 PM -2:00 PM	2.72 -3.30
14/08/2015	2:00 PM -2:15 PM	3.30 -3.68
14/08/2015	2:15 PM -2:30 PM	3.68 -2.80
14/08/2015	2:30 PM -2:45 PM	2.80 -3.05
14/08/2015	2:45 PM -3:00 PM	3.05 -3.10
14/08/2015	3:00 PM -3:15 PM	3.10 -2.60
14/08/2015	3:15 PM -3:30 PM	2.60 -1.93
14/08/2015	3:30 PM -3:45 PM	1.93 -2.22
14/08/2015	3:45 PM -4:00 PM	2.22 -1.88
14/08/2015	4:00 PM -4:15 PM	1.88 -1.77
14/08/2015	4:15 PM -4:30 PM	1.77 -1.48
14/08/2015	4:30 PM -4:45 PM	1.48 -1.10
14/08/2015	4:45 PM -5:00 PM	1.10 -0.69
14/08/2015	5:00 PM -5:15 PM	0.69 -0.43
14/08/2015	5:15 PM -5:30 PM	0.43 -0.15
14/08/2015	5:30 PM -5:45 PM	0.15 -0.04

3. Region and postal code range

Table C-11 shows the full data set of the region and postal code range data [29] in Table 3-3.

Table C-11: Region and postal code data [29]

Zip Range		Region
1	299	Gauteng—Pretoria/Tshwane
300	499	North West—the northern part
500	698	Limpopo—south and west
699	999	Limpopo—north and east
1000	1399	Mpumalanga—northern half
1400	1699	Gauteng—East Rand / Ekurhuleni Metro
1700	1799	Gauteng—West Rand, Mogale City/Krugersdorp, Roodepoort (now part of Johannesburg)
1800	1999	Gauteng—Soweto and Vereeniging/Vanderbijlpark Region (Motsweding)
2000	2199	Gauteng—Johannesburg (original Johannesburg, Randburg, Sandton)
2200	2499	Mpumalanga—southern half
2500	2899	Northwest Province—southern and central
2900	3199	KwaZulu-Natal—Northern KwaZulu-Natal
3200	3299	KwaZulu-Natal—Pietermaritzburg and surrounds
3300	3599	KwaZulu-Natal—Midlands
3600	3799	KwaZulu-Natal—Region between Durban and Pietermaritzburg, including satellite towns, such as Westville.
3800	3999	KwaZulu-Natal—Zululand Region (including Richards Bay)
4000	4099	KwaZulu-Natal—Durban / Ethekwini (original area)
4100	4299	KwaZulu-Natal—South Coast
4300	4499	KwaZulu-Natal—North Coast
4500	4730	KwaZulu-Natal—Griqualand East and Umzimkulu
4731	5199	Eastern Cape—Former Transkei
5200	5299	Eastern Cape—East London
5300	5499	Eastern Cape—historical "Border" region
5500	5999	Eastern Cape—Northern part

6000	6099	Eastern Cape—Port Elizabeth
6100	6499	Eastern Cape—Western part
6500	6699	Western Cape—Garden Route and Oudtshoorn area
6700	6899	Western Cape—Klein Karoo
6900	7099	Western Cape—Great Karoo
7100	7299	Western Cape—Area south-east of Cape Town
7300	7399	Western Cape—West Coast
7400	7599	Western Cape—Northern parts of Cape Metropole
7600	7699	Western Cape—Areas East of Cape Town, such as Stellenbosch
7700	8099	Western Cape—Cape Town and Cape Peninsula
8100	8299	Northern Cape—Namaqualand Region
8300	8799	Northern Cape—Eastern Part
8800	8999	Northern Cape—Gordonia Region
9000	9299	formerly assigned to South West Africa[6]
9300	9399	Free State—Bloemfontein and surrounds
9400	9699	Free State—Northern Free State
9700	9899	Free State—Eastern Free State
9900	9999	Free State—Southern Free State

D. Appendix D– MV protection philosophy calculation tools

The fault levels at each node are based on the 11 kV three-wire fox conductor system i.e. conducting area mm^2 Cu equivalent, pu R, X, B values, and conductor length shown in Table D-1 and Table D-2. Fault levels at each node were calculated using the fault level tool shown in Figure D-1 and Figure D-2. See attachment 2 "Fault level rev1_One.xls" sheet 1.

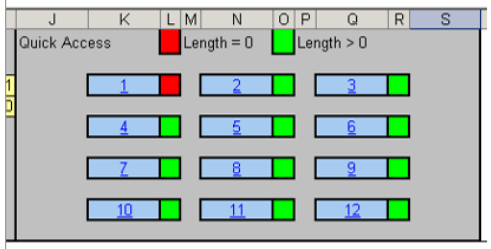
Table D-1: Conductor properties

DESCRIPTION	Normal [A]	Emergency [A]	1s Short Circuit [kA]	R1, R2 [Ω]	X1, X2 [Ω]	R0 [Ω]	X0 [Ω]
Bantam	57	68.5	1.164	5.114830	0.420988	5.261738	1.777166
Bear	529	715	30.969	0.116644	0.336380	0.420596	0.811668
C185P3	410	410	26.455	0.081688	0.058739	0.222662	0.079230
C35P3	170	170	5.005	0.425449	0.074449	0.689844	0.099037
C70P3	240	240	10.01	0.213032	0.066936	0.446007	0.090840
C95P3	290	290	13.585	0.157230	0.063521	0.371560	0.086743
Chickadee	433	576	21.040	0.154832	0.353030	0.301822	1.662153
Copper 0.1HDC	280.1	321.8	9.750	0.292578	0.371591	0.623634	1.015432
Fox	148	192	4.131	0.929570	0.375928	1.076469	1.771290
Goat	607	822	37.983	0.095348	0.331056	0.399300	0.806344
Gopher	122.2	147.2	2.956	1.299293	0.386440	1.446187	1.781798
Hare	292	380	11.824	0.387452	0.341089	0.560351	1.709909
Mink	209	272	7.110	0.543029	0.367085	0.734494	1.703632
Rabbit	186	226.6	5.956	0.650903	0.372622	0.842368	1.709173
Squirrel	106	135	2.363	1.477691	0.393487	1.624580	1.788840
Wolf	378	501	18.512	0.208101	0.347541	0.433006	1.608743

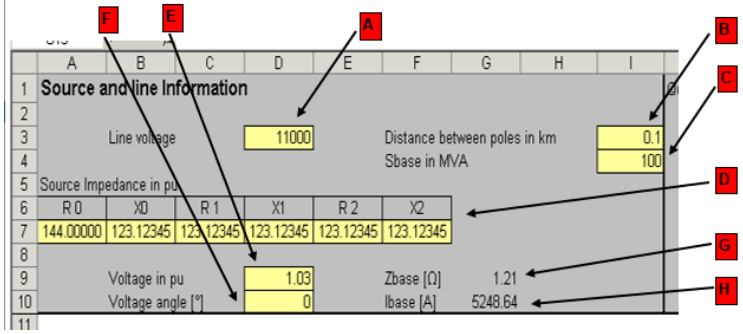
Table D-2: Conductor properties with pu impedances per km

Per Unit Impedances per Km based on 100 MVA and 22kV									50 Degree Probabilistic Rating Amps		Deterministic Rating Amps		
DESCRIPTION	**kV**	**Conducting Area mm² Cu Equiv**	**R_1**	**X_1**	**B_1**	**R_0**	**X_0**	**B_0**	**Normal**	**Emergency**	**75 DEG**	**90 DEG**	**1 Second Short Circuit Rating in kA**
Bantam	22	4.18	1.056783	0.086981	0.000013	1.087136	0.367183	0.000006	57	68.5	76	91	1.164
Squirrel	22	12.90	0.305308	0.081299	0.000014	0.335657	0.369595	0.000006	106	135	107	128	2.363
Gopher	22	16.30	0.268449	0.079843	0.000014	0.298799	0.368140	0.000006	122.2	147.2	124	155	2.956
Fox	22	22.58	0.192060	0.077671	0.000015	0.222411	0.365969	0.000006	148	192	151	190	4.131
Rabbit	22	32.26	0.134484	0.076988	0.000015	0.174043	0.353135	0.000007	186	226.6	185	230	5.956
Mink	22	38.71	0.112196	0.075844	0.000015	0.151755	0.351990	0.000007	209	272	204	255	7.110
Hare	22	64.52	0.080052	0.070473	0.000017	0.115775	0.353287	0.000007	292	380	268	340	11.824
Copper 0.1HDC	22	65.44	0.060450	0.076775	0.000015	0.128850	0.209800	0.000010	280.1	321.8	275	350	9.750
Wolf	22	96.77	0.042996	0.071806	0.000016	0.089464	0.332385	0.000008	378	501	370	480	18.512
Chickadee	22	126.70	0.031990	0.072940	0.000020	0.062360	0.343420	0.000010	433	576	425	556	21.040
Bear	22	161.30	0.024100	0.069500	0.000020	0.086900	0.167700	0.000010	529	715	502	660	30.969
Goat	22	193.50	0.019700	0.068400	0.000020	0.082500	0.166600	0.000010	607	822	565	749	37.983

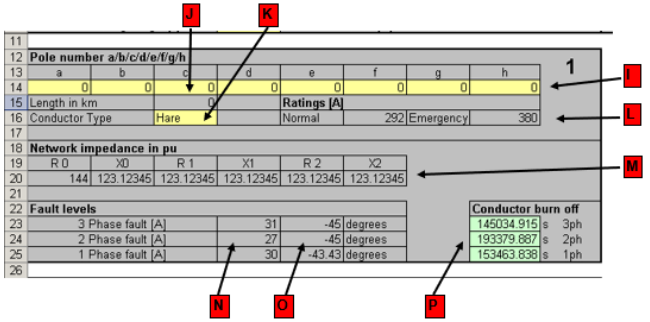
Conductor fault level tool user guide



The quick access bar provides the user a simple, quick way of jumping to the different network nodes.
This is achieved by simply clicking on the desired node number.
The red and green indication lights provides information with regards to the length of the conductor at that node.
If the conductor is longer than 0 km, the light will be green and when the conductor length is 0km, the light will be red.



The nominal line voltage at the busbar can be chosen by using the dropdown list named line voltage (A)
The standard distance between poles can be set by typing the distance between poles in km in B.
The base value of apparent power used for the per unit calculations can be set by typing the base value in MVA in C.
The source impedance in per unit can be set by typing the resistive and reactive, sequence component values in per unit in D.
The per unit busbar voltage can should be set in E.
The busbar voltage angle can be set in F (degrees).
The impedance base in ohms is provided in G for the selected values.
The current base in amps is provided in H for the selected values.



The pole number is inserted in I.
The line length is indicating km in J.
The type of conductor is selected by the dropdown list in K.
The probabilistic normal and emergency ratings of the conductor when template at 50°C is shown in L.
The network impedance is shown in M when looking back into the network at that node.
The three phase, phase to phase and single phase faults are indicated in ampere at N.
The angle of the different faults are indicated in degrees at O.
The conductor burn off times in seconds are indicated at P.
If the burn off time cell is green, it indicates that the burn off time is will not infringe on the one second safety margin. If the cell is red, it is infringing on the safety margin and the conductor can not be used at that fault level.

Figure D-1: Fault level calculating tool user guide

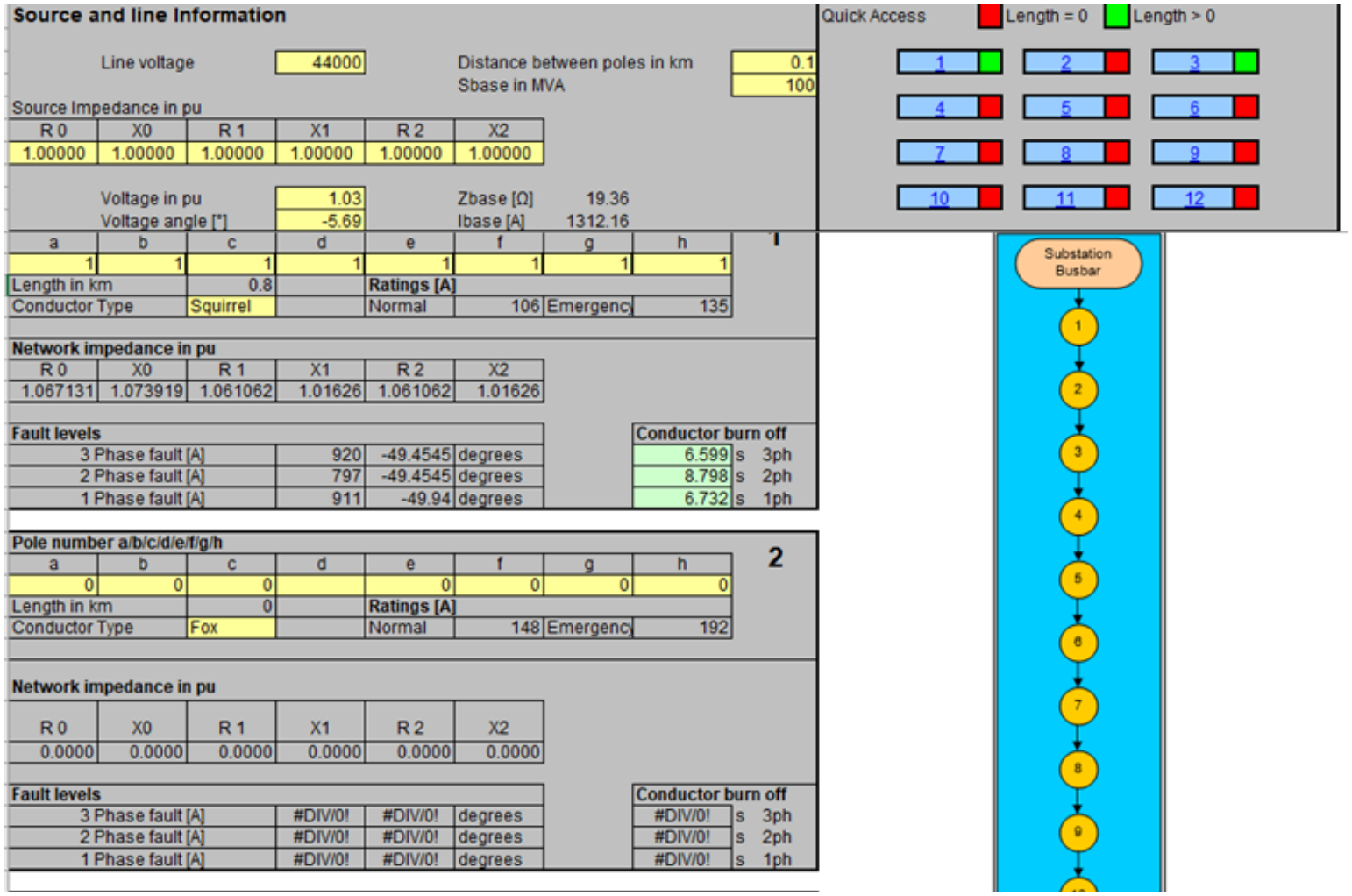


Figure D-2: Fault level calculating tool landing page

E. Appendix E – Training data

This appendix holds the results from a distribution network where detailed studies on the impact of DG penetration on the protection system was done [22].

1. Training data

i. Protection Coordination and Relay Selection before the integration of DG [22] [37]

Using DIgSILENT fault analysis (IEC 60909) the minimum and maximum fault currents on each node, busbar, out-goer or incomer were recorded.

Table E-1: Fault analysis and current transformer selection before DG penetration

Relay/Fuse Location	Relay Label	Phase Element		Earth Element Maximum Fault current (A)
		Minimum Fault current(A)	Maximum Fault current(A)	
A6	15	520	587	340
A6/1	14	13789	15556	15599
A6/6/6	1	305	4745	2713
A1	13	1610	1817	1061
A1/1	12	42677	48170	48712
A1/1/4	7	800	7540	4016
A2	28	632	749	433
A2/1	27	16744	19844	19856
A2/2/5	23	111	926	472

The total burden for each current transformer was set to 1 Ω.

Table E-2: Relay settings, type, operating time and grading

Location	Relay Label	Characteristic	PS (%)		TMS		Operating time (s)	
			OC	Earth	OC	Earth	OC	Earth
A6	15	OC-IDMT-Very inverse	1.5	0.8	0.11	0.05	0.50	0.58
A6/1	14	OC-IDMT-Very inverse	1.5	0.8	0.43	0.55	0.30	0.38
A1	13	OC-IDMT-Very inverse	1.5	0.8	0.67	0.38	0.81	0.85
A1/1	12	OC-IDMT-Very inverse	1.5	0.8	1.17	1.26	0.61	0.65
A2	28	OC-IDMT-Very inverse	1.5	0.8	0.25	0.17	0.86	1.23
A2/1	27	OC-IDMT-Very inverse	1.5	0.8	1.03	1.61	0.66	1.03

ii. Assessment of different fault cases at various locations [37]

Two types of faults were investigated; three phase-fault and a single phase to ground to a neutral fault [22]. Only the actual relay operating times obtained from faults simulated at points close to the main zone circuit breaker (points that are 50% of the feeder length) were used as training data for this research.

Case 1- Parallel Distribution fault location

- Fault on the main busbar

Table E-3: Case 1-before DG

Tripping sequence	Relay 12	Relay 13
Time (s)	0.817	1.095

Case 2- Radial distribution fault location

- Fault on the main busbar.

Table E-4: Case 2-before DG

Tripping sequence	Relay 14	Relay 15
Time (s)	0.320	0.464

Case 3- Ring distribution fault location

- Fault on the main Busbar.

Table E-5: Case 3-prior DG

Tripping sequence	Relay 27	Relay 28
Time (s)	0.746	0.846

iii. Case Study: Protection system after the integration of DG (PV) and Induction Machines

Case 1- Parallel Distribution fault location

- Fault on the main busbar

Table E-6: Case 1-after DG

Tripping sequence Time (s)	Relay 12	Relay 13
38%	0.817	1.017
NRS	0.817	1.094

--Fault current higher than minimum pick-up current setting

Case 2- Radial distribution fault location

- Fault on the main busbar.

Table E-7: Case 2-after DG

Tripping sequence Time (s)	Relay 14	Relay 15
35%	0.320	0.461
NRS	0.320	0.461

Case 3- Ring distribution fault location

- Fault on the main Busbar.

Table E-8: Case 3-after DG

Tripping sequence	Relay 27	Relay 28
NRS Time (s)	0.746	0.644

**Secondary relay trips before Primary relay*

iv. Protection issues and triggers [37]

Table E-9: Parallel distribution Protection issue triggers

	Installation Size(kVA)			
Issue / Technology	Loss of Coordination	De-sensitisation	False tripping	Bi-directionality
Inverter-based	750kVA (~75% of DG)	380kVA (~35% of DG)	No effect	380kVA (~35% of DG)
Induction Machine DG	---	340kVA (~35% of DG)	No effect	340kVA (~35% of DG)

Table E-10: Radial distribution Protection issue triggers

	Installation Size(kVA)			
Issue / Technology	Loss of Coordination	De-sensitisation	False tripping	Bi-directionality
Inverter-based	110kVA (~35% DG)	110kVA (~35% DG)	236.25kVA (~75% of DG)	110kVA (~35% DG)
Induction Machine DG	110kVA (~35% DG)	110kVA (~35% DG)	No effect	110kVA (~35% DG)

Table E-11: Ring Distribution Protection issue triggers

	Installation Size (kVA)			
Issue / Technology	Loss of Coordination	De-sensitisation	False tripping	Bi-directionality
Inverter-based	375kVA (~75% of DG)	200kVA (~38% DG)	No effect	200kVA (~38% DG)
Induction Machine DG	--	143kVA (~35% DG)	No effect	143kVA (~38% DG)

Table E-12: PV integration Protection issue triggers

	Installation Size (kVA)			
Issue / Distribution	Loss of Coordination	De-sensitisation	False tripping	Bidirectionality
Parallel	750kVA	380kVA	No effect	380kVA
Radial	110kVA	110kVA	236.25kVA	110kVA
Ring	375kVA	200kVA	No effect	200kVA

Table E-13 Induction machine DG Protection issue triggers

	Installation Size (kVA)			
Issue	Loss of Coordination	De-sensitisation	False tripping	Bidirectionality

Distribution				
Parallel	---	340kVA	No effect	340kVA
Radial	110kVA	110kVA	No effect	110kVA
Ring	--	143kVA	No effect	143kVA

F. Appendix F – Machine learning models

This appendix holds the regression and classification model parameters (theta values) for all the scenarios.

1. Linear regression as a relay operating time predictor

Model parameters (theta values) for four combinations of the network protection parameters (features) as inputs i.e. all features, fitting (adding polynomials), eliminating features, adding polynomials and eliminating other features.

i. All features as inputs

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 587 148 0 ], y = 7
 x = [6 0 587 29 0 ], y = 1
 x = [2 0 749 152 1 ], y = 80
 x = [2 1 749 34 1 ], y = 61
 x = [1 1 1817 149 1 ], y = 11
 x = [1 1 1817 68 1 ], y = 5
 x = [0 1 1817 157 1 ], y = 1
 x = [0 1 1817 157 1 ], y = 1
 x = [6 0 587 148 0 ], y = 2
 x = [6 0 587 29 0 ], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 7.314468
 -8.427974
 -0.684909
 -10.923147
 -0.739823
 0.975197

Program paused. Press enter to continue.
Predicted 0 percent penetration operating time of recloser with features [1,
204,    191.314,        0.889931937](using gradient descent):
 seconds 12.032394
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 37.974102
 -3.717085
 -0.062685
 -0.019235
 -0.002331
 1.516335


Predicted 0 per penetration operating time of recloser with features [1,        0.2382
204,    191.314,        0.889931937](using normal equations):
 seconds 11.986953
```

Figure F-1: Linear regression all feature model parameters before DG penetration

```
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.718243
 -1.118146
 -1.161114
 -0.647853
 0.767601
 -0.878060

Program paused. Press enter to continue.
Predicted 35 per penetration operating time of recloser with features [1,
204,    191.314,        0.889931937](using gradient descent):
 seconds 7.586719
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 15.810096
 -1.595390
 3.277861
 -0.005227
 0.070281
 -13.314624

Predicted 35 per penetration operating time of recloser with features [1,
204,    191.314,        0.889931937](using normal equations):
 seconds 10.298572
```

Figure F-2: Linear regression all feature model parameters after 35% of DG penetration

```
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.308368
 -1.950330
 -0.870586
 -1.380690
 0.051742
 -0.623236

Program paused. Press enter to continue.
Predicted 65 per penetration operating time of recloser with features [1,
204,    191.314,        0.889931937](using gradient descent):
 seconds 6.295922
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 26.255264
 -2.804466
 -3.124493
 -0.006382
 0.003794
 -7.033670

Predicted 65 per penetration operating time of recloser with features [1,
204,    191.314,        0.889931937](using normal equations):
 seconds 9.488531
```

Figure F-3: Linear regression all feature model parameters after 65% of DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 587 148 0 ], y = 22
 x = [6 0 587 29 0 ], y = 4
 x = [2 0 749 152 1 ], y = 10
 x = [2 1 749 34 1 ], y = 8
 x = [1 1 1817 149 1 ], y = 10
 x = [1 1 1817 68 1 ], y = 5
 x = [0 1 1817 157 1 ], y = 1
 x = [0 1 1817 157 1 ], y = 1
 x = [6 0 587 148 0 ], y = 4
 x = [6 0 587 29 0 ], y = 2
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.125261
 -2.166391
 -1.101013
 -1.253631
 0.215144
 -0.945459

Program paused. Press enter to continue.
Predicted 75 per penetration operating time of recloser with features [1,       0.2382
204,    191.314,        0.889931937](using gradient descent):
 seconds 6.641049
```

```
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 28.059865
 -3.115583
 -3.492624
 -0.006128
 0.013524
 -10.187595

Predicted 75 per penetration operating time of recloser with features [1,       0.2382
204,    191.314,        0.889931937](using normal equations):
 seconds 10.254995
Program paused. Press enter to continue.
```

Figure F-4: Linear regression all-feature model parameters after 75% of DG penetration

ii. Fitting (adding polynomials)

```
First 10 examples from the dataset:
 x = [0 0 0 587 344569 148 21862 0 ], y = 7
 x = [1 0 0 587 344569 29 832 0 ], y = 1
 x = [2 0 0 749 561001 152 23134 1 ], y = 80
 x = [1 1 1 749 561001 34 1152 1 ], y = 61
 x = [1 1 0 1817 3301489 149 22303 1 ], y = 11
 x = [1 1 1 1817 3301489 68 4565 1 ], y = 5
 x = [0 1 1 1817 3301489 157 24793 1 ], y = 1
 x = [0 1 1 1817 3301489 157 24793 1 ], y = 1
 x = [6 0 0 587 344569 148 21862 0 ], y = 2
 x = [1 0 0 587 344569 29 832 0 ], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 7.313173
 -0.979838
 1.580953
 0.814356
 -3.464830
 -5.855405
 1.655740
 0.092737
 2.632824

Program paused. Press enter to continue.
Predicted 0 PV penetration operating time of recloser with features [1, 0.238
        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gradient d
 seconds 9.854579
```

Figure F-5: Linear regression fitted (second-degree polynomials) model parameters before DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 0 587 344569 148 21862 0  0], y = 11
 x = [1 0 0 587 344569 29 832 0  0], y = 2
 x = [2 0 0 749 561001 152 23134 1  1], y = 19
 x = [1 1 1 749 561001 34 1152 1  0], y = 14
 x = [1 1 0 1817 3301489 149 22303 1  1], y = 13
 x = [1 1 1 1817 3301489 68 4565 1  0], y = 6
 x = [0 1 1 1817 3301489 157 24793 1  1], y = 1
 x = [0 1 1 1817 3301489 157 24793 1  2], y = 1
 x = [6 0 0 587 344569 148 21862 0  0], y = 9
 x = [1 0 0 587 344569 29 832 0  0], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.593037
 -1.117605
 3.469767
 0.116095
 -0.101139
 -3.301361
 3.758626
 2.100578
 3.025225
 -8.138356

Program paused. Press enter to continue.
Predicted 35 percent PV penetration operating time of recloser with features
17, 0.056746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937,
adient descent):
 seconds 12.517941
```

Figure F-6: Linear regression fitted (second-degree polynomials) model parameters after 35% of DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 0 587 344569 148 21862 0  0], y = 14
 x = [1 0 0 587 344569 29 832 0  0], y = 2
 x = [2 0 0 749 561001 152 23134 1  1], y = 15
 x = [1 1 1 749 561001 34 1152 1  0], y = 12
 x = [1 1 0 1817 3301489 149 22303 1  1], y = 11
 x = [1 1 1 1817 3301489 68 4565 1  0], y = 5
 x = [0 1 1 1817 3301489 157 24793 1  1], y = 1
 x = [0 1 1 1817 3301489 157 24793 1  2], y = 1
 x = [6 0 0 587 344569 148 21862 0  0], y = 4
 x = [1 0 0 587 344569 29 832 0  0], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.107167
 -3.507030
 1.929948
 3.165036
 -1.230118
 -4.802562
 5.422191
 1.574330
 3.297076
 -9.079718

Program paused. Press enter to continue.
Predicted 65 percent PV penetration operating time of recloser with features
17, 0.056746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937,
adient descent):
 seconds 12.762355
```

Figure F-7: Linear regression fitted (second-degree polynomials) model parameters after 65% of DG penetration

```
First 10 examples from the dataset:
 x = [0 0 0 587 344569 148 21862 0 ], y = 22
 x = [1 0 0 587 344569 29 832 0 ], y = 4
 x = [2 0 0 749 561001 152 23134 1 ], y = 10
 x = [1 1 1 749 561001 34 1152 1 ], y = 8
 x = [1 1 0 1817 3301489 149 22303 1 ], y = 10
 x = [1 1 1 1817 3301489 68 4565 1 ], y = 5
 x = [0 1 1 1817 3301489 157 24793 1 ], y = 1
 x = [0 1 1 1817 3301489 157 24793 1 ], y = 1
 x = [6 0 0 587 344569 148 21862 0 ], y = 4
 x = [1 0 0 587 344569 29 832 0 ], y = 2
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.889246
 -2.569465
 -0.447171
 0.932662
 -0.970796
 -1.602676
 2.175635
 0.768824
 -2.363279

Program paused. Press enter to continue.
Predicted 75 per penetration operating time of recloser with features [1,
6746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gr
 seconds 9.483618
```

Figure F-8: Linear regression fitted (second-degree polynomials) model parameters after 75% of DG penetration

iii. Eliminating features

```
Loading data ...
First 10 examples from the dataset:
 x = [0 148 0], y = 7
 x = [0 29 0], y = 1
 x = [0 152 1], y = 80
 x = [1 34 1], y = 61
 x = [1 149 1], y = 11
 x = [1 68 1], y = 5
 x = [1 157 1], y = 1
 x = [1 157 1], y = 1
 x = [0 148 0], y = 2
 x = [0 29 0], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 7.313173
 -4.378063
 -3.296820
 3.290035

Predicted loaded operating time of recloser with features [1,   0.238215017,
        0.889931937](using gradient descent):
 seconds 8.571796
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 18.627674
 -19.914325
 -0.138850
 23.040365

Predicted loaded operating time of recloser with features [423  0.01    231
84](using normal equations):
 seconds 7.824229
```

Figure F-9: Linear regression eliminated (eliminated relay position) model parameters before DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 148 0], y = 11
 x = [0 29 0], y = 2
 x = [0 152 1], y = 19
 x = [1 34 1], y = 14
 x = [1 149 1], y = 13
 x = [1 68 1], y = 6
 x = [1 157 1], y = 1
 x = [1 157 1], y = 1
 x = [0 148 0], y = 9
 x = [0 29 0], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.572076
 -0.960325
 1.308865
 -1.199837

Predicted 35 percent PV operating time of recloser with features [1,     0.238
91.314, 0.889931937](using gradient descent):
 seconds 8.300730
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 6.052210
 0.335106
 0.047932
 -7.581723

Predicted 35 percent PV  operating time of recloser with features [423  0.01
.112850284](using normal equations):
 seconds 8.554938
```

Figure F-10: Linear regression eliminated (eliminated relay positions) model parameters after 35% DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 148 0], y = 14
 x = [0 29 0], y = 2
 x = [0 152 1], y = 15
 x = [1 34 1], y = 12
 x = [1 149 1], y = 11
 x = [1 68 1], y = 5
 x = [1 157 1], y = 1
 x = [1 157 1], y = 1
 x = [0 148 0], y = 4
 x = [0 29 0], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.106263
 -1.106616
 0.002657
 -0.202741

Predicted 65 percent PV operating time of recloser with features [1,     0.238
91.314, 0.889931937](using gradient descent):
 seconds 6.371604
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 7.787821
 -3.336949
 -0.004970
 0.317768

Predicted 65 percent PV  operating time of recloser with features [423  0.01
.112850284](using normal equations):
 seconds 6.324915
```

Figure F-11: Linear regression eliminated (eliminated relay position) model parameters after 65% DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 148 0], y = 22
 x = [0 29 0], y = 4
 x = [0 152 1], y = 10
 x = [1 34 1], y = 8
 x = [1 149 1], y = 10
 x = [1 68 1], y = 5
 x = [1 157 1], y = 1
 x = [1 157 1], y = 1
 x = [0 148 0], y = 4
 x = [0 29 0], y = 2
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.889246
 -1.101267
 0.496833
 -0.724381

Predicted 75 percent PV operating time of recloser with features [1,    0.238
91.314, 0.889931937](using gradient descent):
 seconds 6.730926
Program paused. Press enter to continue.
Solving with normal equations...
Theta computed from the normal equations:
 7.293601
 -2.403362
 0.012112
 -2.558834

Predicted 75 percent PV  operating time of recloser with features [423  0.01
.112850284](using normal equations):
 seconds 6.761050
```

Figure F-12: Linear regression eliminated (eliminated relay position) model parameters after 75% DG penetration

iv. Fitting and eliminating of features

```
First 10 examples from the dataset:
 x = [0 0 587 344569 148 21862 0  ], y = 7
 x = [0 0 587 344569 29 832 0  ], y = 1
 x = [0 0 749 561001 152 23134 1  ], y = 80
 x = [1 1 749 561001 34 1152 1  ], y = 61
 x = [1 0 1817 3301489 149 22303 1  ], y = 11
 x = [1 1 1817 3301489 68 4565 1  ], y = 5
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [0 0 587 344569 148 21862 0  ], y = 2
 x = [0 0 587 344569 29 832 0  ], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 7.313173
 1.993877
 0.563532
 -3.321716
 -5.728026
 1.379273
 -0.027064
 3.032117

Program paused. Press enter to continue.
Predicted 0 PV penetration operating time of recloser with features [1, 0.238
        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gradient d
 seconds 8.921099
```

Figure F-13: Linear regression eliminated (eliminated relay position and fitted second-degree polynomials) model parameters before DG penetration

```
First 10 examples from the dataset:
 x = [0 0 587 344569 148 21862 0  ], y = 11
 x = [0 0 587 344569 29 832 0  ], y = 2
 x = [0 0 749 561001 152 23134 1  ], y = 19
 x = [1 1 749 561001 34 1152 1  ], y = 14
 x = [1 0 1817 3301489 149 22303 1  ], y = 13
 x = [1 1 1817 3301489 68 4565 1  ], y = 6
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [0 0 587 344569 148 21862 0  ], y = 9
 x = [0 0 587 344569 29 832 0  ], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.592046
 0.114106
 -0.239380
 0.069821
 -0.777330
 1.488643
 0.550974
 -1.640395

Program paused. Press enter to continue.
Predicted 35 PV penetration operating time of recloser with features [1,
6746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gr
 seconds 8.800915
```

Figure F-14: Linear regression eliminated (eliminated relay position and fitted second-degree polynomials) model parameters after 35% DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 587 344569 148 21862 0  ], y = 14
 x = [0 0 587 344569 29 832 0  ], y = 2
 x = [0 0 749 561001 152 23134 1  ], y = 15
 x = [1 1 749 561001 34 1152 1  ], y = 12
 x = [1 0 1817 3301489 149 22303 1  ], y = 11
 x = [1 1 1817 3301489 68 4565 1  ], y = 5
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [1 1 1817 3301489 157 24793 1  ], y = 1
 x = [0 0 587 344569 148 21862 0  ], y = 4
 x = [0 0 587 344569 29 832 0  ], y = 1
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 5.106263
 0.115090
 0.423766
 -0.496640
 -1.524699
 1.054993
 0.034911
 -0.342342

Program paused. Press enter to continue.
Predicted 65 PV penetration operating time of recloser with features [1,
6746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gr
 seconds 6.477888
```

Figure F-15: Linear regression eliminated (eliminated relay position and fitted second-degree polynomials) model parameters after 65% DG penetration

```
Loading data ...
First 10 examples from the dataset:
 x = [0 0 587 344569 148 21862 0  0 ], y = 22
 x = [0 0 587 344569 29 832 0  0 ], y = 4
 x = [0 0 749 561001 152 23134 1  1 ], y = 10
 x = [1 1 749 561001 34 1152 1  0 ], y = 8
 x = [1 0 1817 3301489 149 22303 1  1 ], y = 10
 x = [1 1 1817 3301489 68 4565 1  0 ], y = 5
 x = [1 1 1817 3301489 157 24793 1  1 ], y = 1
 x = [1 1 1817 3301489 157 24793 1  2 ], y = 1
 x = [0 0 587 344569 148 21862 0  0 ], y = 4
 x = [0 0 587 344569 29 832 0  0 ], y = 2
Program paused. Press enter to continue.
Normalizing Features ...
Running gradient descent ...
Theta computed from gradient descent:
 4.890112
 -1.183611
 2.773212
 0.337711
 -2.585628
 2.157382
 0.218800
 4.132801
 -6.208093

Program paused. Press enter to continue.
Predicted 75 PV penetration operating time of recloser with features [1,
6746394,        1204, 1449616,  191.314, 36601.0466,    0.889931937](using gr
 seconds 8.205732
```

Figure F-16: Linear regression eliminated (eliminated relay position and fitted second-degree polynomials) model parameters after 75% DG penetration

2. Linear regression debugging: learning curves

This section holds the table of values and learning curves for each model i.e. all features as inputs, fitting, eliminating features, fitting and eliminating features. See attachment 3 Code_scripts.zip: *ex5.m, ex5_3, ex5_2.m and ex5_4.m.*

i. All features as inputs

1. Relay operating time prediction- before DG penetration

Table F-1: Training error and cross-validation error for a given number of training examples (all feature linear regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.407198
2	0.000000	0.388907
3	0.000000	0.477747
4	0.000000	0.477759
5	0.000000	3.612971
6	0.000000	0.670008
7	0.001796	511.298140
8	0.248992	0.341564
9	0.221542	0.343781
10	0.217299	0.292130
11	0.197545	0.292137
12	0.181147	0.289547
13	0.175236	0.267474
14	0.167069	0.220413
15	0.162752	0.202230
16	0.156378	0.190312
17	0.259628	0.192505
18	0.349875	0.409665
19	0.333835	0.392152

Table F-2: Training error and cross-validation error for a given number of training examples (all feature polynomial regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.480407
2	0.000000	0.501002
3	0.000000	0.501176
4	0.000000	0.500718
5	0.963060	3.613947
6	0.802550	0.888343
7	0.690043	0.916331
8	0.612154	0.916114
9	0.545823	0.916315
10	0.491247	0.915408
11	0.447520	0.915577
12	0.416700	0.962560
13	0.384648	0.962249
14	0.357304	0.962341
15	0.334357	0.947212
16	0.313461	0.947071
17	0.335994	0.646355
18	0.319560	0.651700
19	0.303083	0.651609

Table F-3: Training error and cross-validation error for different lambda values on the all feature model (regularization parameter: before DG penetration)

lambda	Train Error	Validation Error
0.000000	0.303083	0.651609
0.001000	0.303084	0.645176
0.003000	0.303091	0.634492
0.010000	0.303183	0.623613
0.030000	0.303894	0.549356
0.100000	0.309239	0.515161
0.300000	0.327868	0.493118
1.000000	0.362280	0.473780
3.000000	0.387245	0.460349
10.000000	0.406095	0.443437
11.000000	0.407615	0.442120
15.000000	0.412759	0.438100
20.000000	0.417765	0.434852
25.000000	0.421723	0.432711
50.000000	0.433428	0.428140
100.000000	0.442601	0.425881

2. *Relay operating time prediction after 35% of DG penetration*

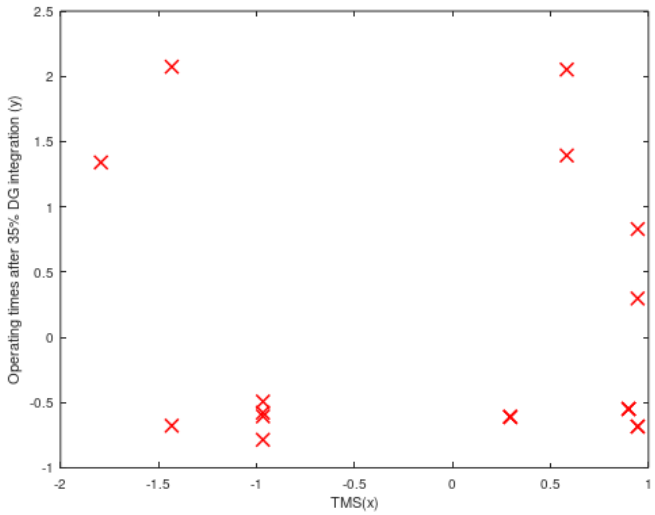


Figure F-17: Normalised TMS vs relay operating times after 35% DG penetration

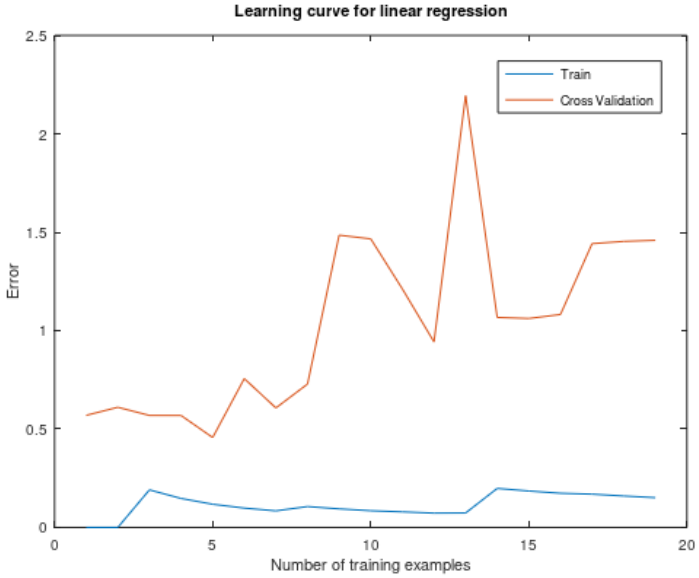


Figure F-18: All features as inputs model linear regression learning curves (relay operating time after 35% of DG penetration)

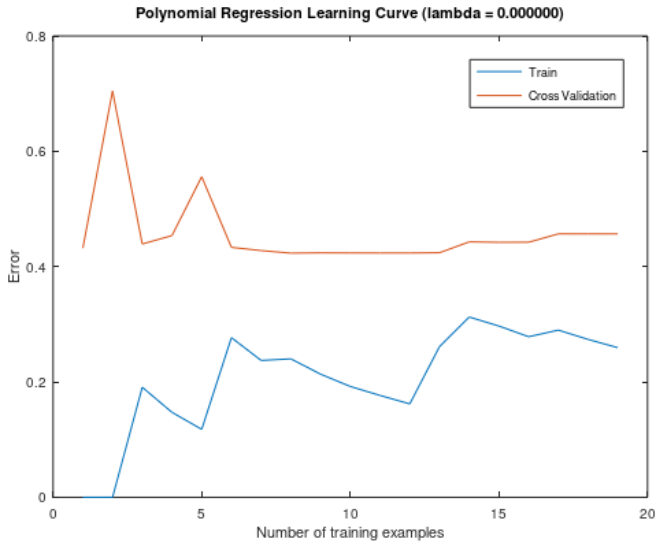


Figure F-19: All-feature model polynomial regression learning curves (relay operating time after 35% DG penetration)

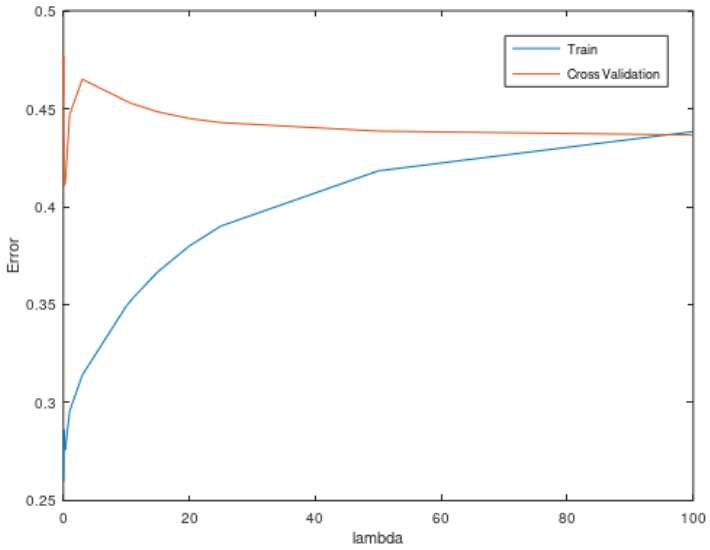


Figure F-20: All-feature model regularization learning curves (relay operating time after 35% DG penetration)

Table F-4: Training error and cross-validation error for a given number of training examples (all feature linear regression model- after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.569541
2	0.000000	0.610612
3	0.190848	0.569133
4	0.147338	0.568877
5	0.117871	0.456848
6	0.098226	0.756142
7	0.084193	0.607197
8	0.106080	0.728527
9	0.094298	1.485149
10	0.084872	1.466314
11	0.079143	1.211965
12	0.072708	0.942830
13	0.073337	2.193499
14	0.198263	1.066917
15	0.185099	1.061770
16	0.173660	1.081991
17	0.169105	1.441994
18	0.159753	1.453429
19	0.151360	1.458995

Table F-5: Training error and cross-validation error for a given number of training examples (polynomial regression all feature model- after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.432101
2	0.000000	0.704500
3	0.190848	0.439268
4	0.147338	0.453878
5	0.117871	0.556020
6	0.276895	0.433454
7	0.237339	0.427878
8	0.240082	0.423389
9	0.213511	0.423887
10	0.192191	0.423720
11	0.176668	0.423631
12	0.161976	0.423635
13	0.261227	0.424102
14	0.312620	0.442954
15	0.296988	0.442176
16	0.278465	0.442318
17	0.289925	0.456852
18	0.273839	0.456798
19	0.259456	0.456766

Table F-6: Training error and cross-validation error for different lambda values on the all feature model (regularization parameter: after 35% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.259456	0.456766
0.001000	0.259457	0.477064
0.003000	0.259464	0.454840
0.010000	0.259656	0.451947
0.030000	0.260092	0.442132
0.100000	0.286235	0.410769
0.300000	0.275816	0.411782
1.000000	0.295578	0.446911
3.000000	0.313972	0.465034
10.000000	0.349279	0.453853
11.000000	0.353198	0.452506
15.000000	0.366757	0.448352
20.000000	0.379923	0.445031
25.000000	0.390083	0.442893
50.000000	0.418273	0.438552
100.000000	0.438300	0.436580

3. *Relay operating time prediction after 65% DG penetration*

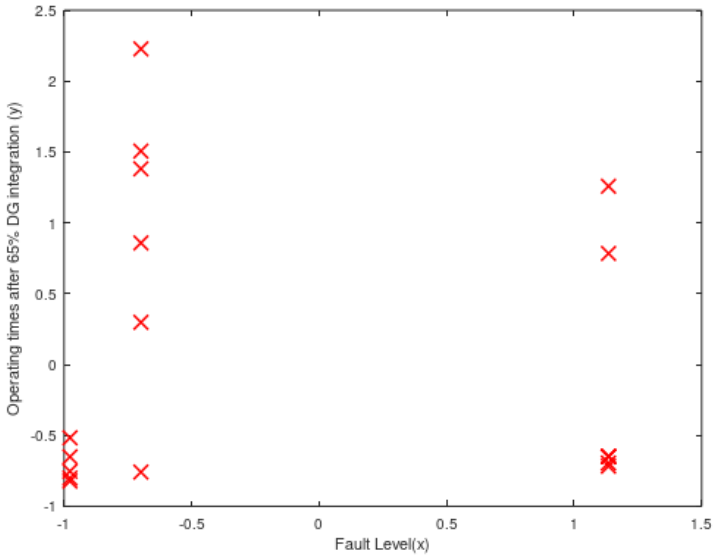


Figure F-21: Normalised fault vs relay operating times after 65% DG penetration

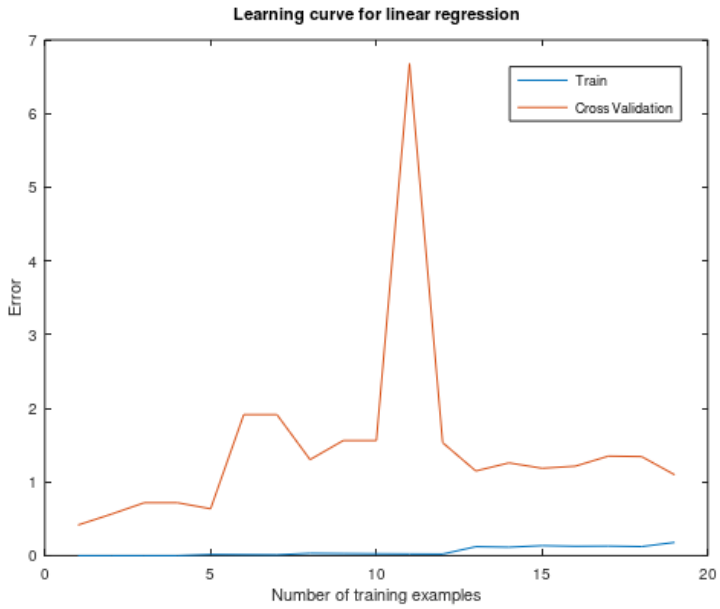


Figure F-22: All features as inputs model linear regression learning curves (relay operating time after 65% of DG penetration)

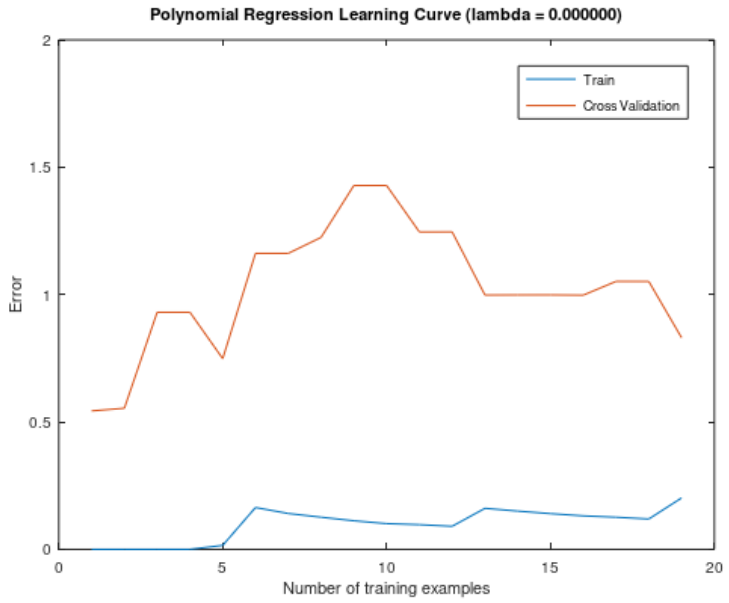


Figure F-23: All-feature model polynomial regression learning curves (relay operating time after 65% DG penetration)

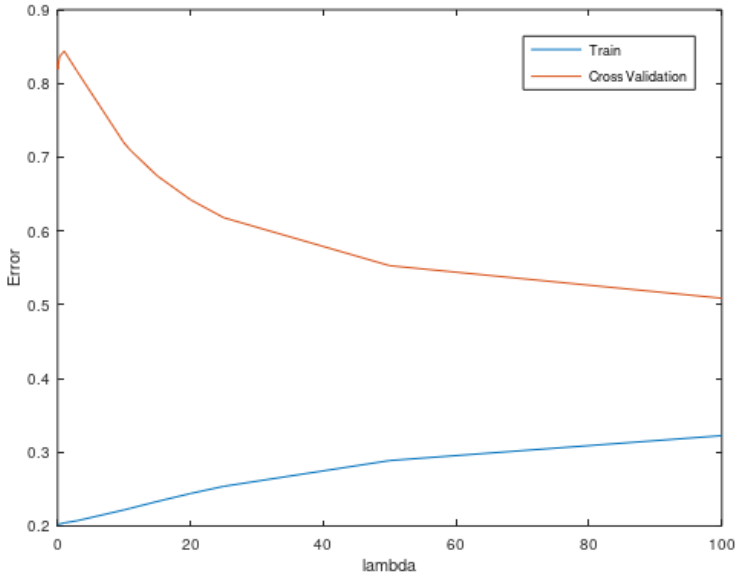


Figure F-24: All-feature model regularization learning curves (relay operating time after 65% DG penetration)

Table F-7: Training error and cross-validation error for a given number of training examples (all feature linear regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.415764
2	0.000000	0.562571
3	0.000000	0.718383
4	0.000000	0.718383
5	0.015694	0.636469
6	0.013079	1.914454
7	0.011210	1.914454
8	0.032147	1.303466
9	0.028575	1.562366
10	0.025717	1.562366
11	0.023387	6.676600
12	0.021701	1.533480
13	0.122127	1.149186
14	0.114865	1.259304
15	0.136111	1.186174
16	0.127845	1.214391
17	0.130516	1.351194
18	0.123284	1.344901
19	0.180336	1.097337

Table F-8: Training error and cross-validation error for a given number of training examples (polynomial regression all feature model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.543355
2	0.000000	0.553949
3	0.000000	0.930614
4	0.000000	0.930614
5	0.015694	0.748554
6	0.164215	1.161892
7	0.140755	1.161892
8	0.126188	1.224377
9	0.112167	1.428023
10	0.100950	1.428141
11	0.096878	1.245924
12	0.090492	1.246194
13	0.160998	0.998222
14	0.150137	0.998602
15	0.140182	0.998711
16	0.131458	0.997782
17	0.126069	1.051765
18	0.119182	1.051507
19	0.201790	0.830569

Table F-9: Training error and cross-validation error for different lambda values on the all feature model (regularization parameter: after 65% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.201790	0.830569
0.001000	0.201790	0.830519
0.003000	0.201790	0.830427
0.010000	0.201793	0.826055
0.030000	0.201812	0.819037
0.100000	0.201967	0.824505
0.300000	0.202590	0.836349
1.000000	0.204148	0.843809
3.000000	0.207216	0.815585
10.000000	0.221741	0.719375
11.000000	0.224049	0.709171
15.000000	0.233196	0.674742
20.000000	0.243969	0.642396
25.000000	0.253717	0.618039
50.000000	0.288681	0.552893
100.000000	0.322436	0.508794

4. *Relay operating time prediction after 75% DG penetration*

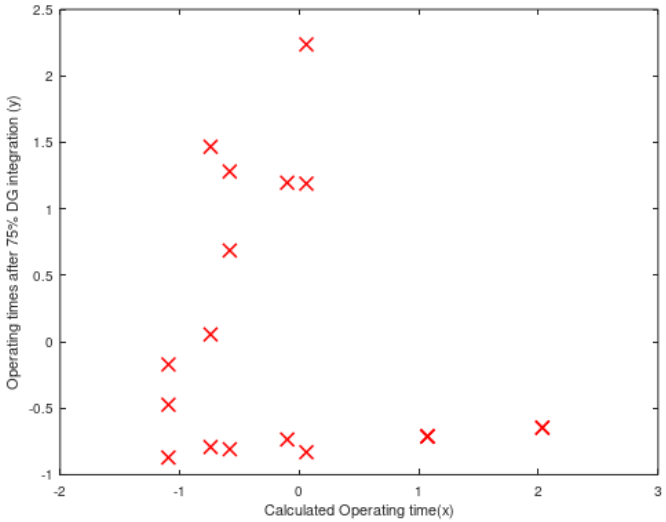


Figure F-25: Normalised calculated operating times vs relay operating time after 75% DG penetration

Table F-10: Training error and cross-validation error for a given number of training examples (all feature linear regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.195186
2	0.000000	0.195186
3	0.000000	0.614722
4	0.000000	7.803144
5	0.000000	9.128664
6	0.000000	8.777602
7	0.039130	8.395615
8	0.034239	8.395615
9	0.042492	9.496785
10	0.051699	461.489961
11	0.058790	144.658261
12	0.057098	79.843062
13	0.126290	81.221726
14	0.287350	62.109920
15	0.269890	60.024764
16	0.309485	65.369026
17	0.291345	62.804705
18	0.281583	49.714669
19	0.319066	23.379090

Table F-11: Training error and cross-validation error for a given number of training examples (polynomial regression all feature model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.762973
2	0.000000	0.762973
3	0.000000	5.618403
4	0.000366	5.531523
5	0.000293	0.814080
6	0.054412	2.994637
7	0.085769	0.764082
8	0.075071	0.757063
9	0.073645	0.540680
10	0.133309	3.263536
11	0.147726	1.952370
12	0.135415	0.494993
13	0.268179	0.439189
14	0.341243	0.425272
15	0.320137	0.419287
16	0.339379	0.498286
17	0.319379	0.515843
18	0.307224	0.521470
19	0.322215	0.417875

Table F-12: Training error and cross-validation error for different lambda values on the all feature model (regularization parameter: after 75% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.322215	0.417875
0.001000	0.322365	0.414608
0.003000	0.322215	0.418050
0.010000	0.322216	0.418495
0.030000	0.322222	0.419883
0.100000	0.322292	0.423472
0.300000	0.327762	0.424131
1.000000	0.325622	0.460290
3.000000	0.328830	0.439257
10.000000	0.349858	0.435512
11.000000	0.352627	0.435215
15.000000	0.362629	0.434459
20.000000	0.372900	0.434096
25.000000	0.381167	0.434039
50.000000	0.405528	0.434755
100.000000	0.424010	0.435917

ii. Fitting (adding polynomials)

1. Relay operating time prediction before DG penetration

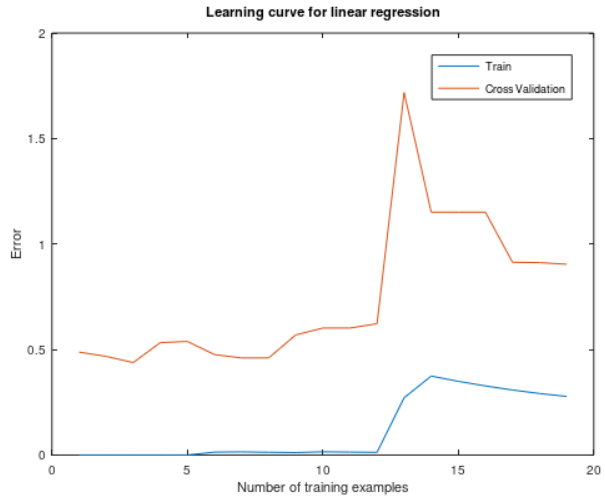


Figure F-26: Fitted model linear regression learning curves (relay operating time before DG penetration)

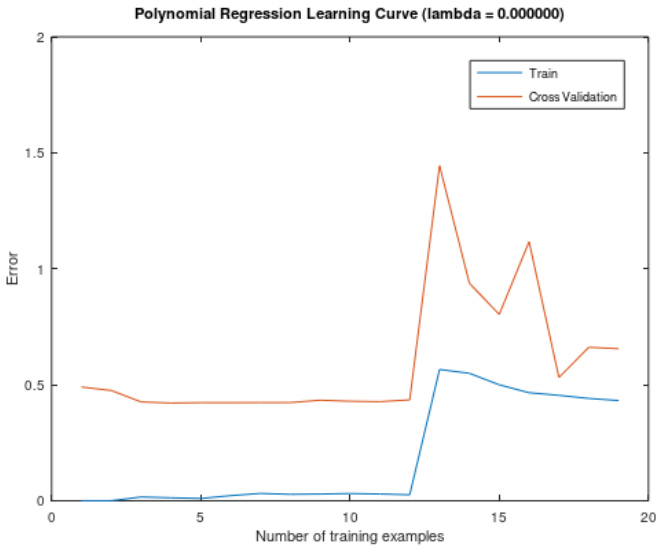


Figure F-27: Fitted model polynomial regression learning curves (relay operating time before DG penetration)

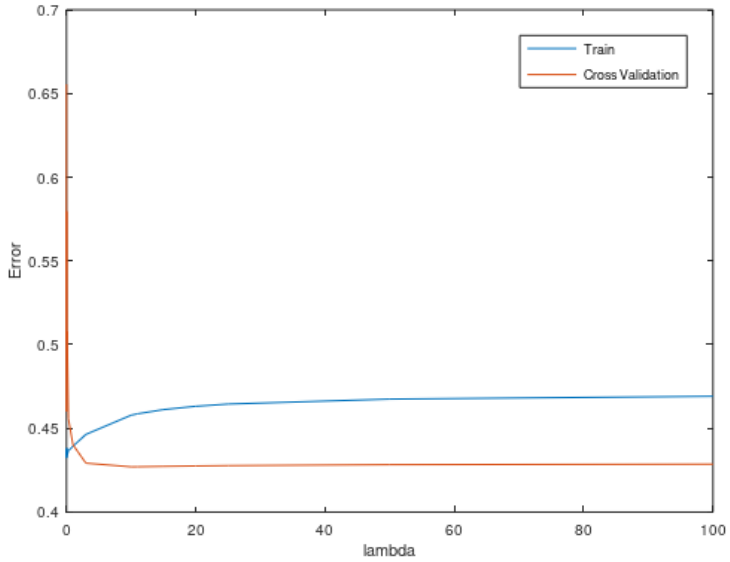


Figure F-28: Fitted model regularization learning curves (relay operating time before DG penetration)

Table F-13: Training error and cross-validation error for a given number of training examples (fitted linear regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.487834
2	0.000000	0.467722
3	0.000000	0.438401
4	0.000000	0.532854
5	0.000000	0.538558
6	0.013885	0.476058
7	0.015004	0.460321
8	0.013128	0.460321
9	0.011670	0.569337
10	0.015359	0.601847
11	0.013963	0.601838
12	0.012799	0.622440
13	0.271634	1.718289
14	0.374349	1.151145
15	0.349393	1.151043
16	0.327556	1.151140
17	0.308310	0.913855
18	0.291581	0.911964
19	0.277783	0.904405

Table F-14: Training error and cross-validation error for a given number of training examples (fitted polynomial regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.490291
2	0.000000	0.475398
3	0.016399	0.426054
4	0.012742	0.421224
5	0.009863	0.422971
6	0.022104	0.422906
7	0.031652	0.423213
8	0.027701	0.423243
9	0.028911	0.433465
10	0.031111	0.428915
11	0.029146	0.427176
12	0.025778	0.435007
13	0.565942	1.444530
14	0.549427	0.937708
15	0.499892	0.803158
16	0.465797	1.116482
17	0.454500	0.531357
18	0.441134	0.661612
19	0.432091	0.655504

Table F-15: Training error and cross-validation error for different lambda values on the fitted model (regularization parameter: before DG penetration)

lambda	Train Error	Validation Error
0.000000	0.432091	0.655504
0.001000	0.437536	0.556896
0.003000	0.436375	0.460071
0.010000	0.432526	0.579569
0.030000	0.438248	0.502722
0.100000	0.433976	0.507843
0.300000	0.436626	0.455821
1.000000	0.439296	0.440211
3.000000	0.446389	0.429224
10.000000	0.457868	0.427031
11.000000	0.458705	0.427074
15.000000	0.461209	0.427294
20.000000	0.463199	0.427552
25.000000	0.464518	0.427759
50.000000	0.467481	0.428311
100.000000	0.469156	0.428646

2. Relay operating time prediction after 35% of DG penetration

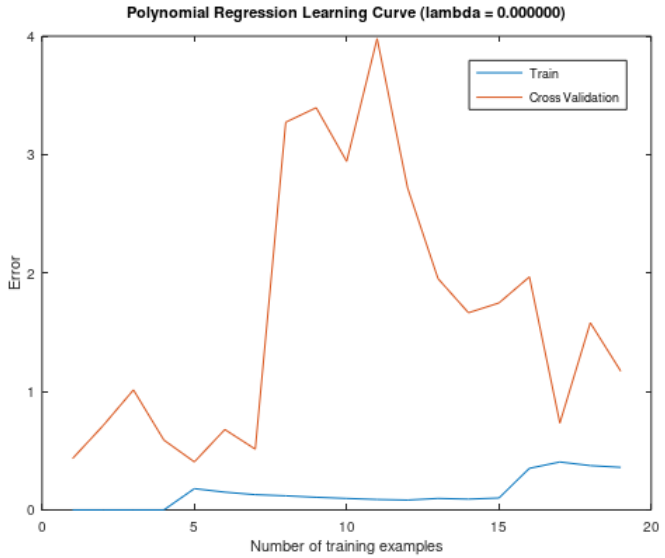


Figure F-29: Feature elimination model polynomial regression learning curves (relay operating time after 35% DG penetration)

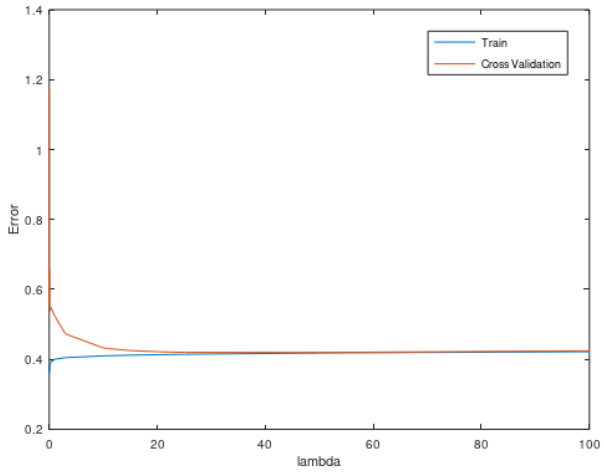


Figure F-30: Feature elimination model regularization learning curves (relay operating time after 35% DG penetration)

Table F-16: Training error and cross-validation error for a given number of training examples (fitted linear regression model after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.365973
2	0.000000	0.425739
3	0.000000	0.453394
4	0.000000	0.471058
5	0.000000	0.574530
6	0.000000	0.575258
7	0.000000	0.575258
8	0.001171	0.586290
9	0.001041	0.379888
10	0.000939	0.379656
11	0.000884	0.400263
12	0.002871	0.384644
13	0.003939	0.387596
14	0.048093	0.306767
15	0.045008	0.304155
16	0.042406	0.303240
17	0.104805	0.334863
18	0.209290	0.377341
19	0.217244	0.387857

Table F-17: Training error and cross-validation error for a given number of training examples (fitted polynomial regression model- after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.428460
2	0.000000	0.828891
3	0.000226	0.804528
4	0.000169	0.803463
5	0.119773	0.495853
6	0.111658	0.449148
7	0.095707	0.448874
8	0.110122	0.500647
9	0.105774	0.569478
10	0.123315	0.469562
11	0.181801	1.205748
12	0.179671	1.164006
13	0.167236	1.167279
14	0.296089	0.622237
15	0.289185	0.819930
16	0.276243	0.715918
17	0.325075	0.709538
18	0.384464	0.465720
19	0.421924	0.376454

Table F-18: Training error and cross-validation error for different lambda values on the fitted model (regularization parameter: after 35% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.421924	0.376454
0.001000	0.385751	0.823224
0.003000	0.385756	0.816211
0.010000	0.387318	0.661052
0.030000	0.386205	0.730733
0.100000	0.419944	0.399239
0.300000	0.407021	0.418271
1.000000	0.408730	0.417760
3.000000	0.418946	0.391817
10.000000	0.441853	0.413608
11.000000	0.443178	0.414412
15.000000	0.447736	0.417341
20.000000	0.451464	0.419679
25.000000	0.454034	0.421275
50.000000	0.460124	0.425003
100.000000	0.463483	0.427364

3. Relay operating time prediction after 65% of DG penetration

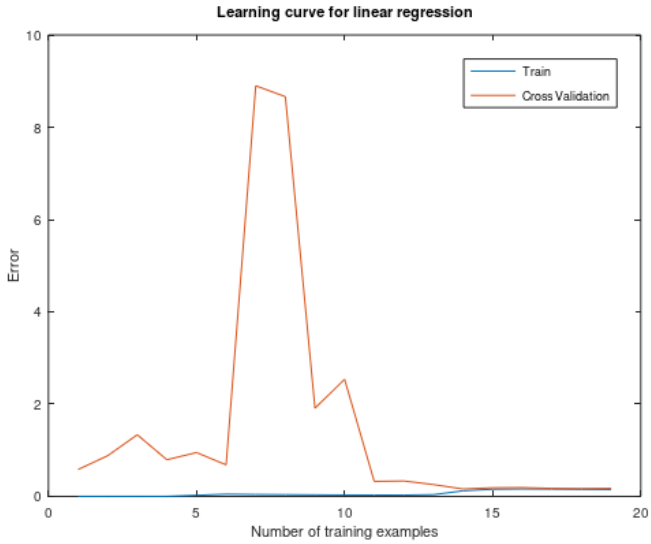


Figure F-31: Fitted model linear regression learning curves (relay operating time after 65% DG penetration)

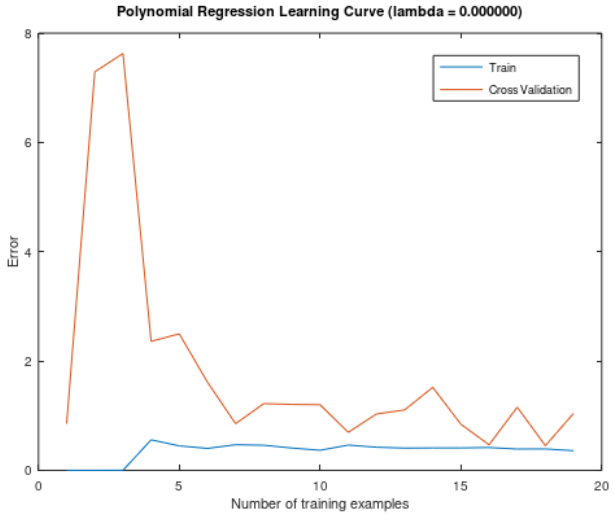


Figure F-32: Fitted model polynomial regression learning curves (relay operating time after 65% DG penetration)

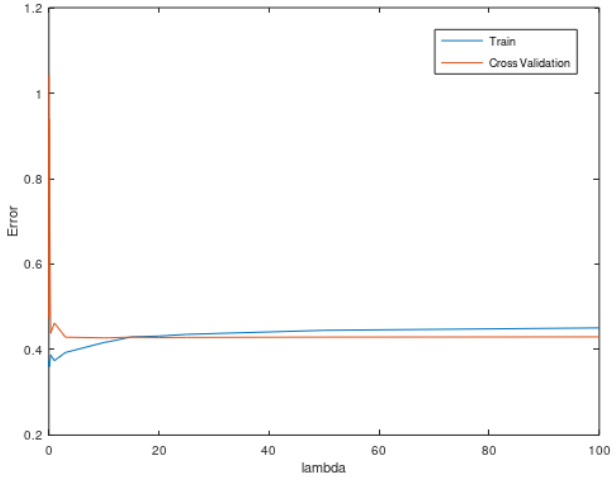


Figure F-33: Fitted model regularization learning curves (relay operating time after 65% DG penetration)

Table F-19: Training error and cross-validation error for a given number of training examples (fitted linear regression model- after 65% DG penetration)

Training Examples	Train Error	Cross Validation Error
1	0.000000	0.581918
2	0.000000	0.881416
3	0.000000	1.336222
4	0.000000	0.794064
5	0.023738	0.950225
6	0.048078	0.683007
7	0.041209	8.899143
8	0.037233	8.663265
9	0.033096	1.911794
10	0.029787	2.538392
11	0.028995	0.325436
12	0.027621	0.334181
13	0.037523	0.256013
14	0.120798	0.161260
15	0.155352	0.187126
16	0.160992	0.191239
17	0.158403	0.173316
18	0.153353	0.167968
19	0.149111	0.174013

Table F-20: Training error and cross-validation error for a given number of training examples (fitted polynomial regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.855090
2	0.000000	7.293010
3	0.000587	7.627914
4	0.558412	2.360030
5	0.448027	2.498660
6	0.401652	1.613338
7	0.468971	0.854184
8	0.459674	1.221662
9	0.408599	1.206810
10	0.367742	1.200800
11	0.462367	0.695128
12	0.423519	1.031391
13	0.407386	1.105341
14	0.410361	1.518933
15	0.410615	0.843560
16	0.417624	0.467293
17	0.390601	1.155133
18	0.392042	0.451104
19	0.360023	1.042197

Table F-21: Training error and cross-validation error for different lambda values on the fitted model (regularization parameter: after 65% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.360023	1.042197
0.001000	0.366658	0.558607
0.003000	0.374035	0.473465
0.010000	0.364782	0.606325
0.030000	0.360150	0.939038
0.100000	0.361076	0.787459
0.300000	0.387723	0.438265
1.000000	0.374253	0.462291
3.000000	0.393173	0.429062
10.000000	0.416657	0.427443
11.000000	0.418895	0.427518
15.000000	0.429720	0.429242
20.000000	0.431673	0.428224
25.000000	0.435675	0.428512
50.000000	0.445124	0.429268
100.000000	0.450735	0.429756

4. *Relay operating time prediction after 75% of DG penetration*

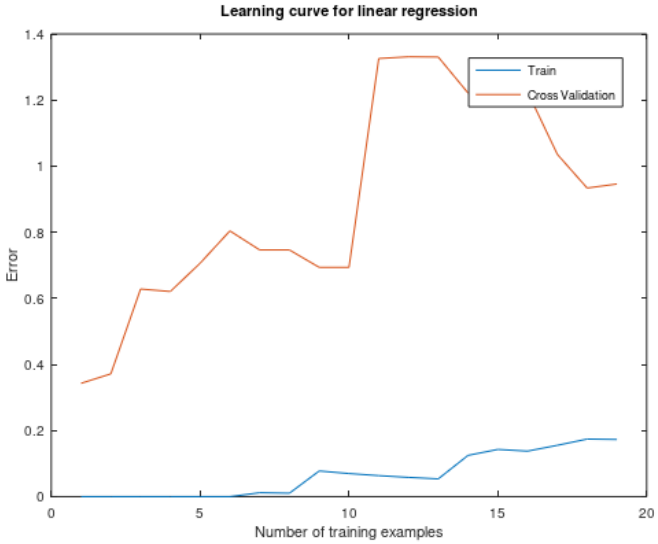


Figure F-34: Fitted model linear regression learning curves (relay operating time after 75% DG penetration)

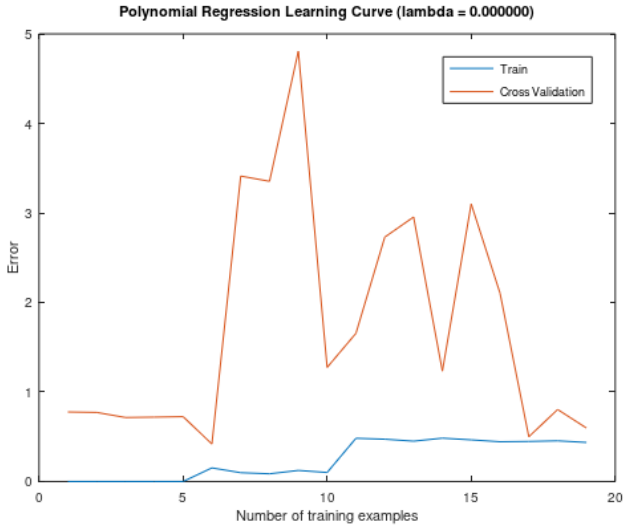


Figure F-35: Fitted model polynomial regression learning curves (relay operating time after 75% DG penetration)

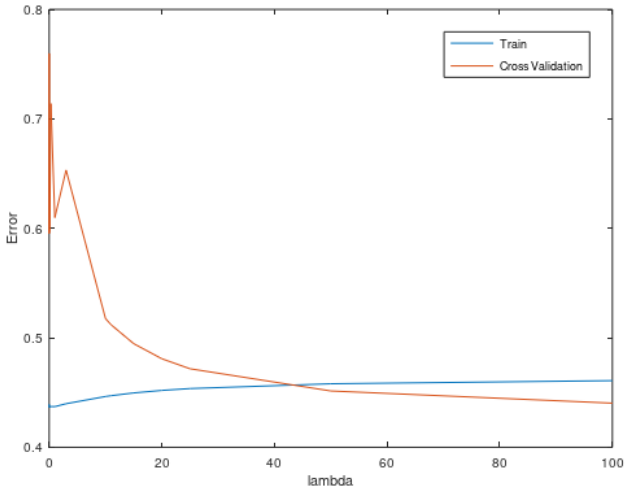


Figure F-36: Fitted model regularization learning curves (relay operating time after 75% DG penetration)

Table F-22: Training error and cross-validation error for a given number of training examples (fitted linear regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.342875
2	0.000000	0.371510
3	0.000000	0.628188
4	0.000000	0.620737
5	0.000000	0.706386
6	0.000000	0.804134
7	0.011844	0.746491
8	0.010364	0.746490
9	0.077668	0.693783
10	0.069901	0.693784
11	0.063547	1.325739
12	0.058251	1.330955
13	0.053771	1.330187
14	0.124922	1.221373
15	0.142830	1.250340
16	0.137749	1.224284
17	0.155336	1.035703
18	0.174173	0.933829
19	0.172965	0.946014

Table F-23: Training error and cross-validation error for a given number of training examples (fitted polynomial regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.776879
2	0.000000	0.771870
3	0.000000	0.714912
4	0.000000	0.719014
5	0.000113	0.724480
6	0.152297	0.419010
7	0.098570	3.411963
8	0.086276	3.353423
9	0.123862	4.805476
10	0.101128	1.274782
11	0.483038	1.655605
12	0.472825	2.729594
13	0.450836	2.954408
14	0.484309	1.233708
15	0.465485	3.100272
16	0.444341	2.104606
17	0.447737	0.499876
18	0.454577	0.804561
19	0.436601	0.597406

Table F-24: Training error and cross-validation error for different lambda values on the fitted model (regularization parameter: after 75% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.436601	0.597406
0.001000	0.436601	0.595178
0.003000	0.439091	0.759851
0.010000	0.436601	0.596673
0.030000	0.436602	0.596688
0.100000	0.436611	0.601134
0.300000	0.437271	0.713974
1.000000	0.437134	0.609805
3.000000	0.439806	0.653227
10.000000	0.446391	0.517577
11.000000	0.447144	0.511894
15.000000	0.449652	0.494683
20.000000	0.451952	0.480876
25.000000	0.453645	0.471741
50.000000	0.458017	0.451450
100.000000	0.460879	0.440319

iii. Eliminating of features

1. *Relay operating time prediction before DG penetration*

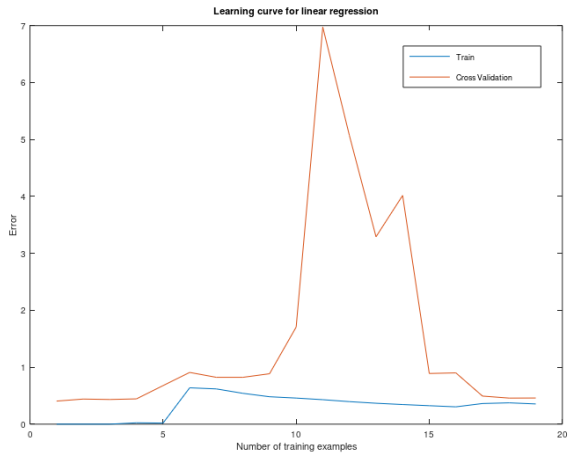


Figure F-37: Feature elimination model linear regression learning curves (relay operating time before DG penetration)

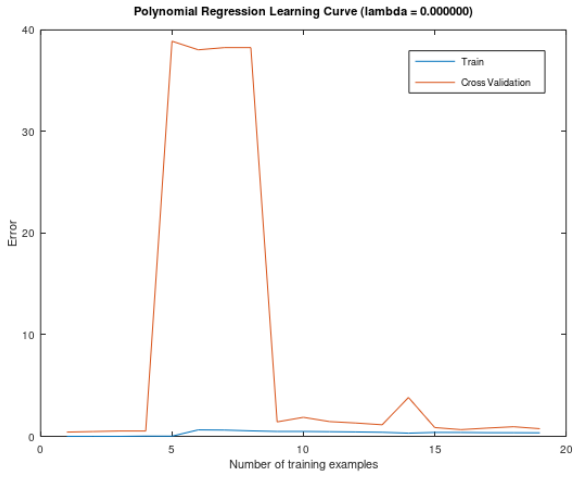


Figure F-38: Feature elimination model polynomial regression learning curves (relay operating time before DG penetration)

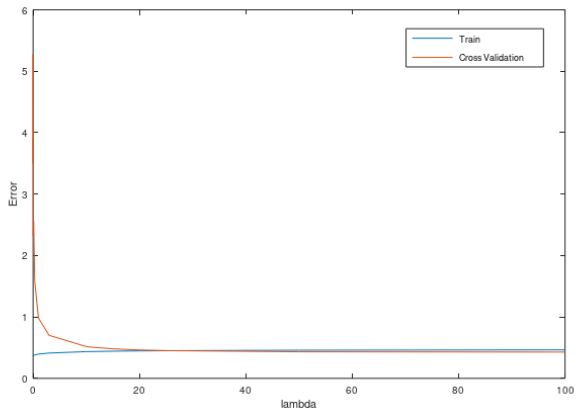


Figure F-39: Feature elimination model regularization learning curves (relay operating time before DG penetration)

Table F-25: Training error and cross-validation error for a given number of training examples (feature eliminated linear regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.405140
2	0.000000	0.441089
3	0.000000	0.433367
4	0.024039	0.444325
5	0.019231	0.677598
6	0.638385	0.908476
7	0.619062	0.822850
8	0.541679	0.822850
9	0.481492	0.885358
10	0.458399	1.704916
11	0.430004	6.974919
12	0.395683	5.070160
13	0.367225	3.291115
14	0.344615	4.016966
15	0.324399	0.889984
16	0.304377	0.900471
17	0.362249	0.493029
18	0.374452	0.457998
19	0.354860	0.459510

Table F-26: Training error and cross-validation error for a given number of training examples (feature eliminated polynomial regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.418812
2	0.000000	0.473204
3	0.000000	0.527152
4	0.024039	0.526436
5	0.019231	38.893064
6	0.638385	38.047466
7	0.619062	38.263778
8	0.541679	38.263800
9	0.481757	1.423905
10	0.484176	1.873223
11	0.459713	1.452729
12	0.434511	1.312936
13	0.400817	1.140272
14	0.314403	3.826452
15	0.387780	0.865456
16	0.383285	0.660441
17	0.361688	0.815290
18	0.358793	0.946545
19	0.349296	0.752742

Table F-27: Training error and cross-validation error for different lambda values on the feature eliminated model (regularization parameter: before DG penetration)

lambda	Train Error	Validation Error
0.000000	0.369082	3.898441
0.001000	0.370574	3.500716
0.003000	0.365782	5.156875
0.010000	0.365870	5.274882
0.030000	0.376303	2.314694
0.100000	0.373461	2.561388
0.300000	0.381301	1.589339
1.000000	0.394770	0.982662
3.000000	0.411787	0.700990
10.000000	0.435139	0.517059
11.000000	0.436903	0.507503
15.000000	0.442352	0.481048
20.000000	0.446949	0.462443
25.000000	0.450181	0.451529
50.000000	0.458306	0.432780
100.000000	0.463734	0.427532

2. *Relay operating time prediction after 35% of DG penetration*

Table F-28: Training error and cross-validation error for a given number of training examples (eliminated features linear regression model after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.544840
2	0.000000	0.931037
3	0.000000	1.196970
4	0.000000	1.287517
5	0.179600	0.688946
6	0.149667	0.759914
7	0.150059	0.865083
8	0.132905	0.814919
9	0.118171	0.815494
10	0.107975	0.818429
11	0.098641	0.798243
12	0.090777	0.792593
13	0.088339	0.768487
14	0.082048	0.764137
15	0.083983	0.736495
16	0.431504	0.470951
17	0.427511	0.499295
18	0.407301	0.511320
19	0.386021	0.509155

Table F-29: Training error and cross-validation error for a given number of training examples (eliminated features polynomial regression model- after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.433452
2	0.000000	0.710598
3	0.000000	1.013048
4	0.000000	0.589343
5	0.179600	0.405431
6	0.149860	0.678684
7	0.128550	0.512874
8	0.119059	3.273273
9	0.106691	3.396021
10	0.096803	2.942815
11	0.087973	3.979509
12	0.083164	2.719210
13	0.096878	1.953012
14	0.091008	1.664590
15	0.100512	1.747328
16	0.351068	1.967558
17	0.404608	0.733975
18	0.374136	1.580012
19	0.359428	1.171025

Table F-30: Regularization Table showing Training error and Cross-validation error different lambda values (Elimination: After 35% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.359428	1.171025
0.001000	0.377198	0.563033
0.003000	0.366830	0.659394
0.010000	0.372076	0.589175
0.030000	0.372692	0.587291
0.100000	0.379819	0.537926
0.300000	0.392183	0.549378
1.000000	0.400293	0.524594
3.000000	0.405035	0.472885
10.000000	0.410023	0.432713
11.000000	0.410462	0.430695
15.000000	0.411949	0.425297
20.000000	0.413403	0.421845
25.000000	0.414572	0.420152
50.000000	0.418331	0.419728
100.000000	0.421989	0.424361

3. Relay operating time prediction after 65% of DG penetration

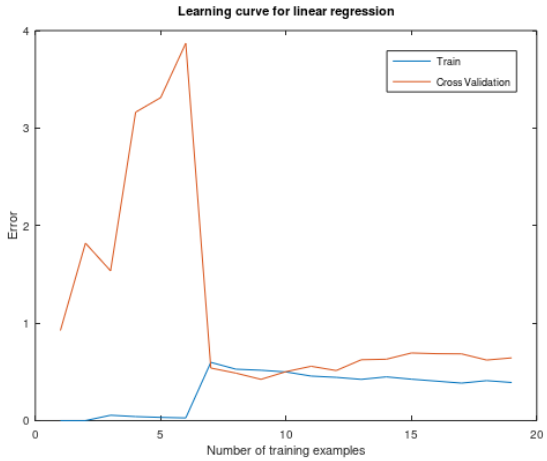


Figure F-40: Feature elimination model linear regression learning curves (relay operating time after 65% DG penetration)

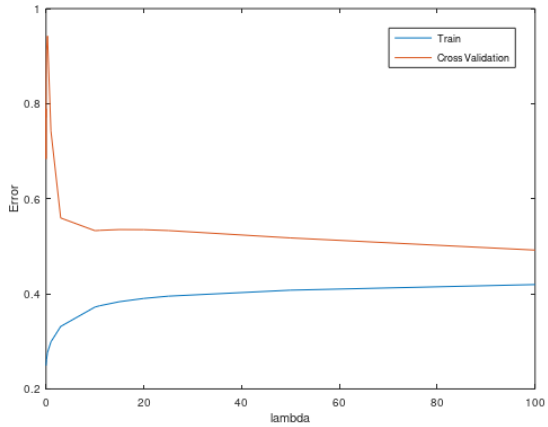


Figure F-41: Feature elimination model regularization learning curves (relay operating time after 65% DG penetration)

Table F-31: Training error and cross-validation error for a given number of training examples (feature eliminated linear regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.924218
2	0.000000	1.819752
3	0.054948	1.536080
4	0.041211	3.163106
5	0.032969	3.312563
6	0.027474	3.869755
7	0.597321	0.539093
8	0.527326	0.486907
9	0.516571	0.423096
10	0.498917	0.502204
11	0.456710	0.557213
12	0.443696	0.513372
13	0.422868	0.624072
14	0.448671	0.629434
15	0.424673	0.693566
16	0.404754	0.686333
17	0.384423	0.684457
18	0.409863	0.621000
19	0.390961	0.643342

Table F-32: Training error and cross-validation error for a given number of training examples (feature eliminated polynomial regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.437395
2	0.000000	0.435928
3	0.000000	0.411299
4	0.000000	0.843401
5	0.113574	0.365998
6	0.125294	0.396361
7	0.238335	0.404965
8	0.208544	1.059790
9	0.196732	0.795536
10	0.301066	0.814626
11	0.291710	0.443544
12	0.376998	0.429346
13	0.406125	0.857597
14	0.368533	1.253940
15	0.352125	1.191276
16	0.342125	1.151319
17	0.336243	0.532771
18	0.324474	0.665949
19	0.312243	0.444007

Table F-33: Training error and cross-validation error for different lambda values on the feature eliminated model (regularization parameter: after 65% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.250261	0.771117
0.001000	0.250473	0.789459
0.003000	0.249592	0.683696
0.010000	0.251186	0.806536
0.030000	0.253528	0.845364
0.100000	0.261359	0.910669
0.300000	0.276386	0.942358
1.000000	0.298969	0.742866
3.000000	0.331435	0.559579
10.000000	0.372191	0.532873
11.000000	0.375004	0.533569
15.000000	0.383469	0.535240
20.000000	0.390368	0.534982
25.000000	0.395139	0.533126
50.000000	0.407730	0.517624
100.000000	0.419553	0.492039

4. Relay operating time prediction after 75% of DG penetration

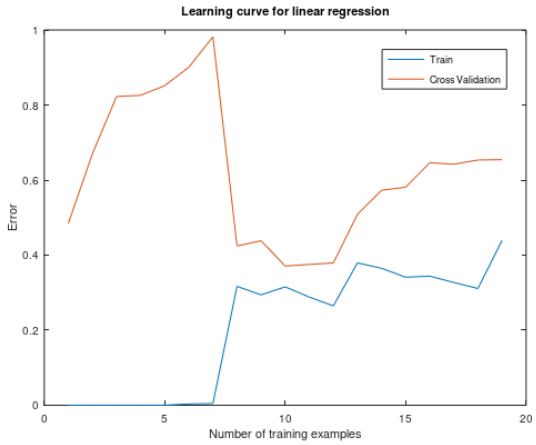


Figure F-42: Feature elimination model linear regression learning curves (relay operating time after 75% DG penetration)

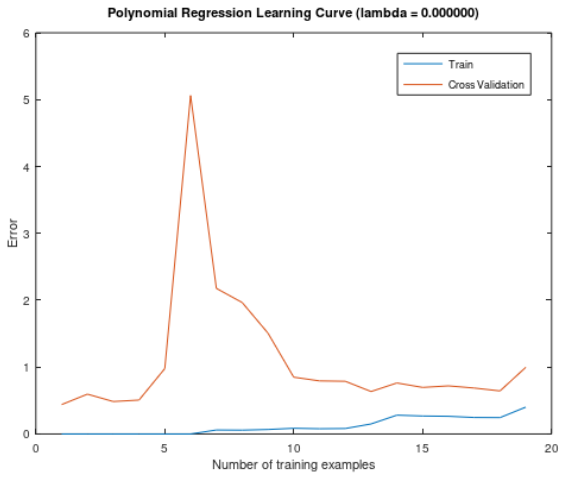


Figure F-43: Feature elimination model polynomial regression learning curves (relay operating time after 75% DG penetration)

Table F-34:Training error and cross-validation error for a given number of training examples (feature eliminated linear regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.484840
2	0.000000	0.670533
3	0.000000	0.822516
4	0.000000	0.825980
5	0.000000	0.851593
6	0.003430	0.900737
7	0.004565	0.981882
8	0.316817	0.424535
9	0.294102	0.438350
10	0.315485	0.370870
11	0.288230	0.375162
12	0.264889	0.379254
13	0.379394	0.508925
14	0.364892	0.572880
15	0.341209	0.581035
16	0.343945	0.646298
17	0.327107	0.642231
18	0.311157	0.653441
19	0.439082	0.654366

Table F-35: Training error and cross-validation error for a given number of training examples (feature eliminated polynomial regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.439602
2	0.000000	0.595890
3	0.000000	0.485323
4	0.000000	0.507208
5	0.000000	0.978055
6	0.000000	5.062254
7	0.059862	2.177535
8	0.057749	1.966356
9	0.067100	1.508896
10	0.085402	0.849507
11	0.078387	0.793700
12	0.081687	0.788211
13	0.150125	0.634789
14	0.281467	0.762949
15	0.268898	0.696314
16	0.264401	0.719624
17	0.246635	0.687069
18	0.244754	0.643815
19	0.402036	0.997483

Table F-36: Training error and cross-validation error for different lambda values on the feature eliminated model (regularization parameter: after 75% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.402036	0.997483
0.001000	0.397799	0.998369
0.003000	0.403566	1.027571
0.010000	0.404277	0.963740
0.030000	0.404177	0.834554
0.100000	0.409267	0.686625
0.300000	0.418283	0.522244
1.000000	0.429070	0.448255
3.000000	0.436688	0.435278
10.000000	0.446251	0.432216
11.000000	0.447111	0.432066
15.000000	0.449908	0.431661
20.000000	0.452413	0.431394
25.000000	0.454234	0.431251
50.000000	0.458887	0.431042
100.000000	0.461933	0.430976

iv. Fitting and eliminating features

1. *Relay operating time prediction before DG penetration*

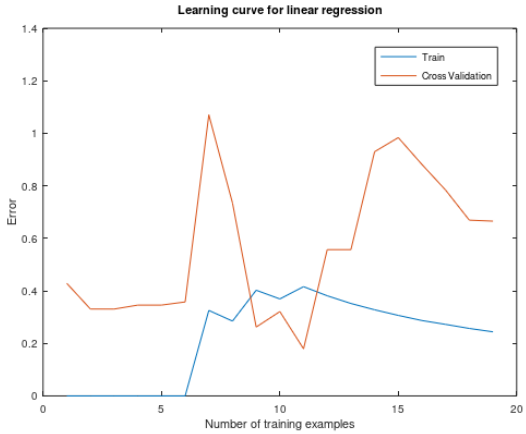


Figure F-44: Feature elimination and fitted model linear regression learning curves (relay operating time before DG penetration)

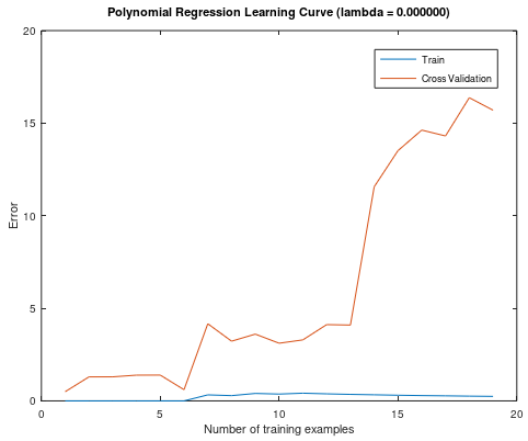


Figure F-45: Feature elimination and fitted model polynomial regression learning curves (relay operating time before DG penetration)

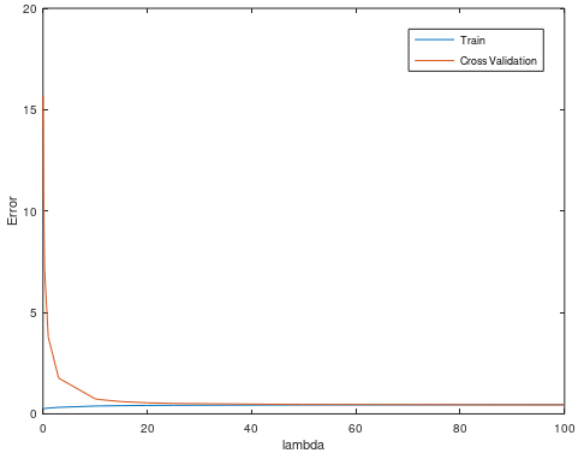


Figure F-46: Feature elimination and fitted model regularization learning curves (relay operating time before DG penetration)

Table F-37: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated linear regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.428819
2	0.000000	0.330898
3	0.000000	0.330765
4	0.000000	0.345834
5	0.000001	0.345874
6	0.000001	0.357900
7	0.325659	1.070335
8	0.284952	0.734464
9	0.402495	0.262267
10	0.369350	0.321174
11	0.416050	0.179337
12	0.381379	0.557070
13	0.352042	0.557074
14	0.328052	0.930027
15	0.306618	0.983728
16	0.287085	0.881405
17	0.272375	0.783722
18	0.256931	0.669379
19	0.244076	0.665585

Table F-38: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated polynomial regression model- before DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.495932
2	0.000000	1.296735
3	0.000000	1.299457
4	0.000000	1.391067
5	0.000001	1.394546
6	0.000001	0.604750
7	0.325659	4.168025
8	0.284952	3.225473
9	0.402495	3.606499
10	0.369350	3.114714
11	0.416050	3.295239
12	0.381386	4.119795
13	0.352044	4.096745
14	0.330256	11.564015
15	0.306327	13.515096
16	0.286844	14.625899
17	0.273626	14.308991
18	0.257099	16.374472
19	0.244609	15.701148

Table F-39: Training error and cross-validation error for different lambda values on the fitted and feature eliminated model (regularization parameter: before DG penetration)

lambda	Train Error	Validation Error
0.000000	0.244609	15.701148
0.001000	0.244998	15.312142
0.003000	0.245571	14.443723
0.010000	0.245699	14.475220
0.030000	0.245700	14.141176
0.100000	0.250392	11.005829
0.300000	0.262257	7.124862
1.000000	0.284223	3.786370
3.000000	0.323356	1.764574
10.000000	0.385589	0.739610
11.000000	0.390236	0.701297
15.000000	0.404323	0.604725
20.000000	0.415786	0.545166
25.000000	0.423561	0.513258
50.000000	0.441897	0.461634
100.000000	0.453167	0.443681

2. Relay operating time prediction after 35% of DG penetration

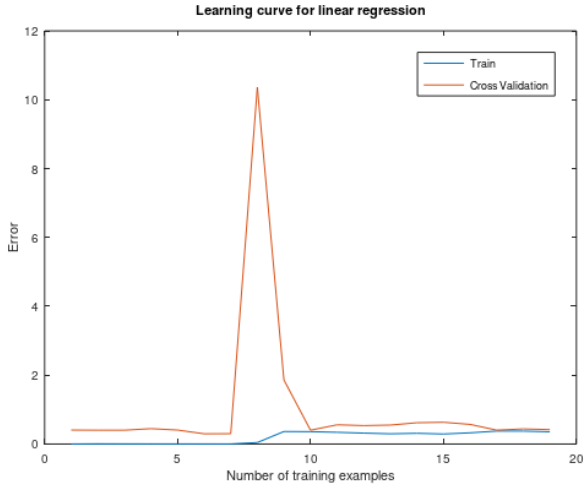


Figure F-47: Feature elimination and fitted model linear regression learning curves (relay operating time after 35% DG penetration)

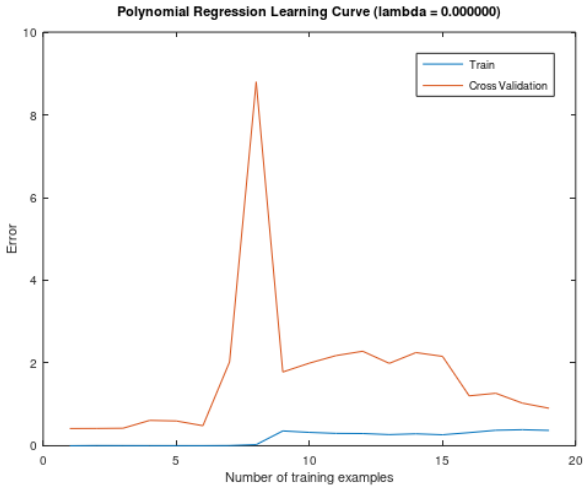


Figure F-48: Feature elimination and fitted model polynomial regression learning curves (relay operating time after 35% DG penetration)

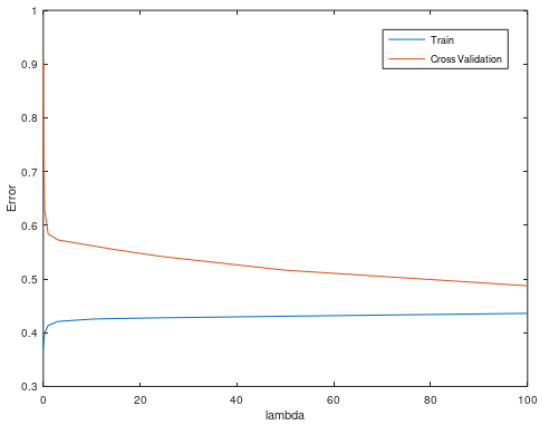


Figure F-49: Feature elimination and fitted model regularization learning curves (relay operating time after 35% DG penetration)

Table F-40: Training error and cross-validation error for a given number of training examples (fitted and eliminated features, linear regression model, after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.406217
2	0.004127	0.402318
3	0.002751	0.401977
4	0.002064	0.446919
5	0.001651	0.406144
6	0.001376	0.296673
7	0.001179	0.299845
8	0.043339	10.365127
9	0.363386	1.861866
10	0.359321	0.401781
11	0.340678	0.558260
12	0.317216	0.535745
13	0.296812	0.550379
14	0.311423	0.618794
15	0.291137	0.631976
16	0.325678	0.570655
17	0.376389	0.403600
18	0.374917	0.439198
19	0.355590	0.419042

Table F-41: Training error and cross-validation error for a given number of training examples (fitted and eliminated feature polynomial regression model- after 35% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.416580
2	0.004127	0.418432
3	0.002751	0.423627
4	0.002064	0.613315
5	0.001651	0.599608
6	0.001378	0.484456
7	0.004274	2.025846
8	0.026463	8.797233
9	0.358229	1.782952
10	0.323296	1.998100
11	0.299296	2.180081
12	0.295311	2.285018
13	0.269707	1.992456
14	0.288152	2.252915
15	0.265863	2.161220
16	0.317207	1.206801
17	0.374434	1.269070
18	0.387981	1.030586
19	0.371147	0.906628

Table F-42: Regularization Table showing Training error and Cross-validation error different lambda values (Fitting and Elimination: After 35% of DG penetration)

lambda	Train Error	Validation Error
0.000000	0.371147	0.906628
0.001000	0.370829	0.891978
0.003000	0.371599	0.902556
0.010000	0.373126	0.838795
0.030000	0.376195	0.806631
0.100000	0.385932	0.713293
0.300000	0.399902	0.628447
1.000000	0.413649	0.583857
3.000000	0.421015	0.572833
10.000000	0.425358	0.561992
11.000000	0.425616	0.560477
15.000000	0.426429	0.554570
20.000000	0.427199	0.547649
25.000000	0.427855	0.541283
50.000000	0.430750	0.516388
100.000000	0.436054	0.487405

3. *Relay operating time prediction after 65% of DG penetration*

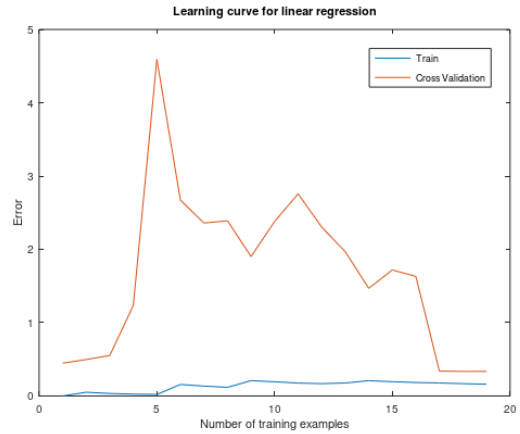


Figure F-50: Feature elimination and fitted model linear regression learning curves (relay operating time after 65% DG penetration)

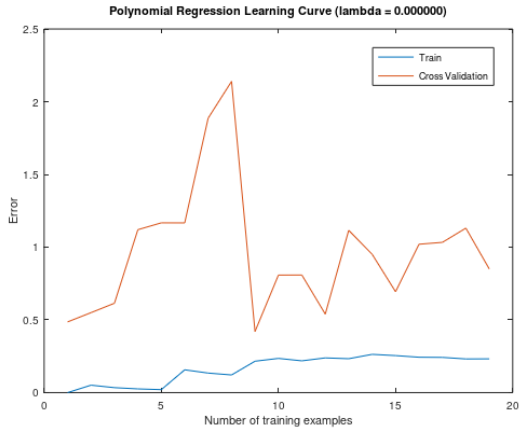


Figure F-51: Feature elimination and fitted model polynomial regression learning curves (relay operating time after 65% DG penetration)

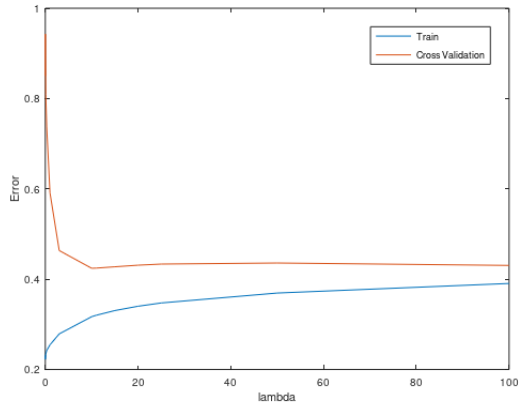


Figure F-52: Feature elimination and fitted model regularization learning curves (relay operating time after 65% DG penetration)

Table F-43: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated linear regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.447494
2	0.051189	0.496993
3	0.034126	0.551745
4	0.025595	1.236218
5	0.020476	4.593069
6	0.156791	2.671650
7	0.134393	2.358745
8	0.117594	2.390208
9	0.208839	1.901780
10	0.193769	2.381014
11	0.176033	2.757071
12	0.168074	2.304316
13	0.176776	1.968076
14	0.208753	1.468791
15	0.194677	1.719114
16	0.183605	1.630812
17	0.177386	0.338706
18	0.167846	0.335060
19	0.160421	0.335584

Table F-44: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated polynomial regression model- after 65% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.486069
2	0.051189	0.549817
3	0.034126	0.614474
4	0.025595	1.120213
5	0.020476	1.167222
6	0.156791	1.166939
7	0.134393	1.886864
8	0.121405	2.139863
9	0.215877	0.420259
10	0.235255	0.808117
11	0.218496	0.808586
12	0.238731	0.539979
13	0.233103	1.115833
14	0.263110	0.951290
15	0.254605	0.693676
16	0.243319	1.020165
17	0.242164	1.033155
18	0.231544	1.130882
19	0.232233	0.850177

Table F-45: Training error and cross-validation error for different lambda values on the fitted and feature elimination model (regularization parameter: after 65% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.232233	0.850177
0.001000	0.230879	0.879010
0.003000	0.222199	0.887689
0.010000	0.223479	0.942814
0.030000	0.226171	0.895321
0.100000	0.233980	0.809197
0.300000	0.241882	0.737168
1.000000	0.254455	0.591582
3.000000	0.278767	0.463383
10.000000	0.316903	0.424047
11.000000	0.320088	0.424481
15.000000	0.330458	0.427403
20.000000	0.340003	0.430949
25.000000	0.347302	0.433426
50.000000	0.369242	0.435780
100.000000	0.390447	0.430440

4. *Relay operating time prediction after 75% of DG penetration*

Table F-46: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated linear regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.448214
2	0.000000	0.819413
3	0.000000	2.079842
4	0.000000	3.374283
5	0.220734	1.946919
6	0.197081	1.789269
7	0.168927	3.671109
8	0.147811	1.080984
9	0.131394	0.449275
10	0.119153	0.445166
11	0.378554	0.421747
12	0.352141	0.424297
13	0.325064	0.423334
14	0.308463	0.357834
15	0.307765	0.346242
16	0.288653	0.345306
17	0.289084	0.341663
18	0.273742	0.341699
19	0.287279	0.411469

Table F-47: Training error and cross-validation error for a given number of training examples (fitted and feature eliminated polynomial regression model- after 75% DG penetration)

# Training Examples	Train Error	Cross Validation Error
1	0.000000	0.432863
2	0.000000	1.635643
3	0.000000	13.044278
4	0.000000	98.448485
5	0.220734	30.302333
6	0.197081	27.878791
7	0.168927	10.673922
8	0.147825	6.877157
9	0.139987	2.371489
10	0.133934	2.076499
11	0.390815	1.637600
12	0.363672	1.436204
13	0.335454	1.790310
14	0.397412	0.558262
15	0.389061	0.632261
16	0.378601	0.522246
17	0.359925	0.521106
18	0.345223	0.544932
19	0.329684	0.505862

Table F-48: Training error and cross-validation error for different lambda values on the fitted and feature eliminated model (regularization parameter: after 75% DG penetration)

lambda	Train Error	Validation Error
0.000000	0.329684	0.505862
0.001000	0.328909	0.522582
0.003000	0.330706	0.495600
0.010000	0.328601	0.524669
0.030000	0.330962	0.495755
0.100000	0.333628	0.483210
0.300000	0.337456	0.501338
1.000000	0.344960	0.538475
3.000000	0.351655	0.537769
10.000000	0.361565	0.498155
11.000000	0.362644	0.495010
15.000000	0.366473	0.485737
20.000000	0.370370	0.478697
25.000000	0.373517	0.474331
50.000000	0.382960	0.466020
100.000000	0.390373	0.462625

3. Logistic regression as a protection miscoordination classifier

This section holds model parameters (theta values) for four combinations of the network protection parameters (features) as inputs for a model that determines which class of protection miscoordination is likely to occur.

i. Loss of coordination

See attachment 3 Code_scripts.zip: *ex2_regPred1, ex2_regPred, ex2_regPred2 and ex2_regPred3.*

1. *All features as inputs*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.002794
 0.038210
 0.026980
 -0.030721
 -0.015798

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.686244
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.055537
 0.600391
 0.581214
 0.546484
 0.640212

Program paused. Press enter to continue.
Train Accuracy: 95.652174
Validation Accuracy: 88.888889
```

Figure F-53: Loss of coordination prediction model parameters with all features as inputs

2. *Elimination features*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 0.052798
 0.043388
 -0.026831
 -0.002167
 0.017904

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.413207
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 0.178478
 0.616024
 0.530274
 0.649451
 0.600140

Program paused. Press enter to continue.
Train Accuracy: 95.652174
Validation Accuracy: 88.888889
```

Figure F-54: Loss of coordination prediction model parameters with relay positions eliminated

3. *Fitting (adding polynomials)*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.003631
 0.014449
 0.012714
 0.070938
 0.075888

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 5.553225
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.187070
 0.600381
 0.585233
 0.742620
 0.738494

Program paused. Press enter to continue.
Train Accuracy: 91.304348
Validation Accuracy: 100.000000
```

Figure F-55: Loss of coordination prediction model parameters with second-degree polynomials added

4. *Fitting and eliminating features*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 0.000000
 -0.000000
 -0.000000
 -0.000000
 -0.000000

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 4.563976
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.137008
 0.587328
 0.592803
 0.610737
 0.605276

Program paused. Press enter to continue.
Train Accuracy: 100.000000
Validation Accuracy: 77.777778
```

Figure F-56: Loss of coordination prediction model parameters with second-degree polynomials added and relay positions eliminated

ii. Loss of grading

See attachment 3 Code_scripts.zip: *ex2_regPred1_g, ex2_regPred_g, ex2_regPred2_g and ex2_regPred3_g.*

1. *All features as inputs*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.174852
 0.188005
 0.346554
 0.214891
 0.257933

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.094289
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.219110
 0.697292
 0.862672
 0.827572
 0.916704

Program paused. Press enter to continue.
Train Accuracy: 78.260870
Validation Accuracy: 22.222222
```

Figure F-57: Loss of grading prediction model parameters with all features as inputs

2. *Eliminating features*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 0.155942
 0.304715
 0.214380
 0.264574
 -0.192474

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.016628
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 0.277394
 0.857410
 0.804871
 0.938597
 0.391834

Program paused. Press enter to continue.
Train Accuracy: 73.913043
Validation Accuracy: 55.555556
```

Figure F-58: Loss of grading prediction model parameters with relay positions eliminated

3. *Fitting (adding polynomials)*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.273774
 0.313615
 0.300562
 0.318390
 0.321489

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 5.718383
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.367979
 0.929470
 0.924838
 0.993810
 0.991728

Program paused. Press enter to continue.
Train Accuracy: 69.565217
Validation Accuracy: 33.333333
```

Figure F-59: Loss of grading prediction model parameters with second-degree polynomials added

4. *Fitting and eliminating features*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.257490
 0.233085
 0.210101
 0.354159
 0.358044

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 5.007677
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.307710
 0.749341
 0.741820
 0.927626
 0.926209

Program paused. Press enter to continue.
Train Accuracy: 78.260870
Validation Accuracy: 22.222222
```

Figure F-60: Loss of grading prediction model parameters with second-degree polynomials added and relay positions eliminated

iii. De-sensitization

See attachment 3 Code_scripts.zip: *ex2_regPred1_d, ex2_regPred_d, ex2_regPred2_d and ex2_regPred3_d.*

1. *All features as inputs*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.116173
 0.150655
 0.185768
 -0.001643
 0.073603

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.045121
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.148409
 0.603556
 0.613784
 0.523019
 0.649353

Program paused. Press enter to continue.
Train Accuracy: 91.304348
Validation Accuracy: 33.333333
```

Figure F-61: De-sensitization prediction model parameters with all features as inputs

2. *Eliminating features*

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 0.283776
 0.289487
 0.052053
 0.191366
 0.041839

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 3.693215
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 0.452375
 0.887545
 0.627156
 0.864766
 0.624511

Program paused. Press enter to continue.
Train Accuracy: 86.956522
Validation Accuracy: 44.444444
```

Figure F-62: De-sensitization prediction model parameters with relay positions eliminated

3. Fitting (adding polynomials)

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.156626
 0.261786
 0.262686
 0.177077
 0.176501

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 5.227594
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.331240
 0.856848
 0.849850
 0.834500
 0.827468

Program paused. Press enter to continue.
Train Accuracy: 82.608696
Validation Accuracy: 66.666667
```

Figure F-63: De-sensitization prediction model parameters with second-degree polynomials added

4. Fitting and eliminating features

```
Cost at initial theta (zeros): 0.693147
Expected cost (approx): 0.693
Gradient at initial theta (zeros) - first five values only:
 -0.185531
 0.239737
 0.219458
 0.244802
 0.247470

Program paused. Press enter to continue.

Cost at test theta (with lambda = 10): 5.579067
Expected cost (approx): 3.16
Gradient at test theta - first five values only:
 -0.294850
 0.874417
 0.861874
 0.875408
 0.868823

Program paused. Press enter to continue.
Train Accuracy: 73.913043
Validation Accuracy: 66.666667
```

Figure F-64: De-sensitization prediction model parameters with second-degree polynomials added and relay positions eliminated

G. Appendix G – Model application

This appendix holds the results from the best regression models (section 6.5) used on the MV distribution networks (section 4.5) to determine relay/recloser operating times before and after three stages of DG penetration. These results from the best regression model were used as inputs to the most accurate logistic regression model to classify which set of network protection features and DG penetration level results in which protection coordination problem.

1. Model application output data

Table G-1 shows the values of the predicted relay operating times and protection miscoordination groups after three stages of DG penetration.

Table G-1: Predicted relay operating times after three DG penetration stages and protection miscoordination groups for the protective devices on the MV distribution networks

Features					Actual relay operating time (s) before DG penetration	Relay operating times (s) after DG penetration of			Loss coordination	De-sensitization	Loss grading
Relay/ recloser	TMS	Fault level (A)	Pick-up current setting (A)	Calculated operating time (s)		35% of ADMD	65% of ADMD	75% of ADMD			
423	0.01	231.00	124.70	0.11	18.45	18.45	18.45	7.36	0	0	0
318	0.03	231.00	138.56	0.41	5.58	5.58	5.58	0.15	0	0	0
200	0.09	306.00	157.27	0.57	9.09	9.09	9.09	2.49	0	0	0
101	0.11	479.00	196.10	0.86	10.46	10.46	10.46	10.46	0	0	0
CB	0.13	909.00	299.20	0.80	6.14	6.14	6.14	6.14	0	0	0
Trf	0.38	368.50	65.61	1.50	18.90	27.81	28.89	9.57	0	0	0
219/1	0.05	437.00	141.54	0.31	6.23	6.23	6.23	5.95	0	0	0
179/1	0.05	529.00	176.46	0.32	5.90	5.90	5.90	5.69	0	0	0
87/1	0.05	1031.00	269.28	0.26	4.02	4.02	4.02	7.58	0	0	0
148	0.01	519.00	125.52	0.05	7.85	7.85	7.85	9.42	0	0	0
114	0.07	519.00	139.47	0.35	9.42	9.42	9.42	6.61	0	0	0
67	0.10	630.00	154.97	0.60	9.84	9.84	9.84	2.77	0	0	0
1	0.24	1204.00	191.31	0.89	3.36	9.90	4.97	10.90	0	0	0
CB	0.37	2710.00	212.57	1.00	0.99	30.50	31.95	25.17	0	0	0
Trf	0.54	677.50	59.05	1.50	13.68	52.77	70.53	31.72	0	0	1
111/1	0.05	555.00	139.47	0.25	8.86	8.86	8.86	7.79	0	0	0
65T/1	0.05	860.00	154.97	0.20	5.82	5.82	5.82	9.37	0	0	0
65/176	0.05	264.00	91.51	0.33	16.16	16.16	16.16	5.31	0	0	0
65/127	0.12	328.00	112.97	0.81	15.55	15.55	15.55	15.55	0	0	0
65/67	0.15	330.00	125.52	1.10	15.81	15.81	15.81	15.81	0	0	0
65/1	0.26	888.00	154.97	1.02	8.98	12.96	8.50	9.40	0	0	0
48	0.17	901.00	172.18	1.31	7.55	7.55	7.55	7.55	0	0	0
65/39/58	0.06	379.00	82.36	0.27	14.43	14.43	14.43	8.33	0	0	0
65/39/56	0.12	379.00	91.51	0.57	14.67	14.67	14.67	5.79	0	0	0

65/39/54	0.17	393.00	101.67	0.86	14.69	14.69	14.69	2.91	0	0	0
65/39/2	0.21	398.00	112.97	1.15	14.87	4.76	14.87	14.87	0	0	0
65/39/1	0.26	582.00	125.52	1.18	11.72	12.79	8.26	4.74	0	0	0
65/2	0.19	587.00	139.47	1.47	15.05	0.29	15.05	15.05	0	0	0
65/64/3	0.06	473.00	125.52	0.31	12.34	12.34	12.34	7.00	0	0	0
65/164/1	0.06	277.00	91.51	0.38	16.02	16.02	16.02	4.98	0	0	0
65/128	0.09	277.00	101.67	0.68	16.39	16.39	16.39	16.39	0	0	0
65/138/36	0.08	266.00	82.36	0.47	16.48	16.48	16.48	4.32	0	0	0
65/138/1	0.06	266.00	91.51	0.77	17.07	17.07	17.07	17.07	0	0	0
65/39/39/1	0.08	435.00	101.67	0.38	13.40	13.40	13.40	7.29	0	0	0
65/39/53/1	0.08	398.00	91.51	0.38	14.13	14.13	14.13	7.43	0	0	0
297	0.07	388.00	82.36	0.31	4.72	4.72	4.72	8.25	0	0	0
277	0.16	413.00	91.51	0.75	5.56	5.56	5.56	5.86	0	0	0
174	0.12	413.00	101.67	1.05	10.52	10.52	10.52	10.52	0	0	0
123	0.15	622.00	112.97	0.74	8.48	8.48	8.48	4.57	0	0	0
75	0.13	829.00	125.52	0.82	7.45	7.45	7.45	7.45	0	0	0
65	0.16	1208.00	139.47	0.68	1.49	1.49	1.49	7.14	0	0	0
50	0.24	1335.00	154.97	0.77	0.60	10.45	5.63	15.28	0	0	0
2	0.18	1833.00	172.18	0.97	0.95	1.13	0.95	0.94	0	0	0
1	0.30	3510.00	191.31	0.71	0.50	19.59	16.97	25.25	0	0	0
CB	0.42	3510.00	212.57	1.01	1.02	37.28	42.27	30.74	0	0	0
Trf	0.80	787.50	59.05	2.10	12.19	84.96	149.06	50.92	0	1	1
2/2	0.05	3510.00	154.97	0.11	0.01	0.01	0.01	12.32	0	0	0
74/21	0.05	490.00	101.67	0.22	11.98	11.98	11.98	8.78	0	0	0
74/27	0.11	337.00	112.96	0.70	14.86	14.86	14.86	0.77	0	0	0
74/9	0.14	292.00	125.52	1.17	19.28	19.28	19.28	19.28	0	0	0
107/1	0.05	764.00	112.97	0.18	6.23	6.23	6.23	10.05	0	0	0
140/6	0.05	393.00	91.51	0.24	10.92	10.92	10.92	8.21	0	0	0
140/1	0.11	393.00	101.67	0.54	11.15	11.15	11.15	5.34	0	0	0
197/1	0.05	511.00	91.51	0.20	6.71	6.71	6.71	9.39	0	0	0
198/1	0.05	507.00	91.51	0.20	6.73	6.73	6.73	9.36	0	0	0
220/6/1	0.05	370.00	82.36	0.23	8.09	8.09	8.09	8.44	0	0	0
220/1	0.11	370.00	91.51	0.53	8.33	8.33	8.33	5.83	0	0	0
235/1	0.05	397.00	91.51	0.24	6.91	6.91	6.91	8.26	0	0	0
259/5/1	0.05	383.00	82.36	0.22	6.26	6.26	6.26	8.61	0	0	0
259/1	0.11	383.00	91.51	0.52	6.50	6.50	6.50	6.19	0	0	0
280/8/1	0.05	358.00	74.12	0.22	5.90	5.90	5.90	8.78	0	0	0
280/1	0.11	358.00	82.36	0.52	6.15	6.15	6.15	6.59	0	0	0
280/57/1	0.05	284.00	74.12	0.26	7.09	7.09	7.09	7.55	0	0	0
49/373	0.09	236.00	91.51	0.63	17.75	17.75	17.75	0.00	0	0	0
49/250	0.11	387.00	112.97	0.94	15.68	15.68	15.68	15.68	0	0	0
49/148	0.12	400.00	125.52	0.90	15.14	15.14	15.14	15.14	0	0	0
49/79	0.12	590.00	139.47	0.82	11.85	11.85	11.85	11.85	0	0	0

49/29/21	0.06	803.00	125.52	0.22	7.87	7.87	7.87	9.90	0	0	0
49/29/1	0.13	803.00	139.47	0.52	7.96	7.96	7.96	9.28	0	0	0
49/1	0.21	1052.00	154.97	0.75	6.15	5.22	6.15	11.65	0	0	0
49/63/11	0.06	623.00	139.47	0.28	10.43	10.43	10.43	8.16	0	0	0
49/110/1	0.06	621.00	125.52	0.26	10.67	10.67	10.67	8.74	0	0	0
49/140/1	0.06	487.00	125.52	0.31	12.78	12.78	12.78	7.22	0	0	0
49/147/1	0.06	464.00	125.52	0.32	13.15	13.15	13.15	6.85	0	0	0
49/166/1	0.06	509.00	112.97	0.27	12.61	12.61	12.61	8.21	0	0	0
49/175/20	0.06	421.00	101.67	0.29	14.17	14.17	14.17	7.67	0	0	0
49/175/1	0.11	421.00	112.97	0.59	14.39	14.39	14.39	4.52	0	0	0
49/195/7	0.06	404.00	101.67	0.30	14.44	14.44	14.44	7.39	0	0	0
49/195/1	0.11	404.00	112.97	0.60	14.67	14.67	14.67	4.00	0	0	0
49/252/1	0.06	393.00	101.67	0.31	14.62	14.62	14.62	7.19	0	0	0
49/195/18	0.06	404.00	82.36	0.26	14.70	14.70	14.70	8.69	0	0	0
49/357/1	0.12	250.00	91.51	0.83	17.83	17.83	17.83	17.83	0	0	0
49/261	0.12	250.00	101.67	1.13	18.30	18.30	18.30	18.30	0	0	0
49/471/1	0.05	235.00	74.12	0.30	17.51	17.51	17.51	6.18	0	0	0
49/478/1	0.05	229.00	74.12	0.31	17.61	17.61	17.61	5.96	0	0	0
49/470	0.05	229.00	82.36	0.61	18.13	18.13	18.13	18.13	0	0	1
4	0.26	2877.00	154.97	0.60	0.36	11.82	7.35	22.69	0	0	0
3	0.37	2877.00	172.18	0.90	0.80	30.37	31.77	28.28	0	0	0
2	0.48	2877.00	191.31	1.20	1.43	46.05	56.98	33.08	0	0	0
CB	0.57	2877.00	212.57	1.50	2.24	58.83	81.84	36.97	0	0	0
Trf	1.04	719.25	59.05	2.83	13.76	104.32	232.83	65.35	0	1	1
86/1	0.05	719.00	125.52	0.20	11.15	11.15	11.15	9.49	0	0	0
111/98	0.05	307.00	112.97	0.35	13.26	13.26	13.26	4.68	0	0	0
200/1	0.05	390.00	112.97	0.28	8.17	8.17	8.17	6.85	0	0	0
304/3	0.05	270.00	91.51	0.32	6.13	6.13	6.13	5.53	0	0	0
314/1	0.05	261.00	91.51	0.33	5.87	5.87	5.87	5.19	0	0	0
202	0.05	261.00	112.97	0.63	10.80	10.80	10.80	10.80	0	0	0
61/7/1	0.06	895.00	112.97	0.20	0.04	0.04	0.04	10.60	0	0	0
61/40/1	0.06	618.00	112.97	0.24	10.41	10.41	10.41	9.23	0	0	0
61/2	0.13	618.00	125.52	0.54	13.09	13.09	13.09	7.71	0	0	0
111/29/1	0.06	452.00	112.97	0.30	10.94	10.94	10.94	7.44	0	0	0
111/1	0.11	452.00	125.52	0.60	11.15	11.15	11.15	4.08	0	0	0
5	0.13	565.00	139.47	0.81	13.98	13.98	13.98	13.98	0	0	0
146/12/1	0.06	519.00	101.67	0.25	8.59	8.59	8.59	8.89	0	0	0
146/1	0.12	519.00	112.97	0.55	8.77	8.77	8.77	6.96	0	0	0

113	0.10	572.00	125.52	0.82	9.82	9.82	9.82	9.82	0	0	0
356/1	0.01	265.00	91.51	0.07	3.66	3.66	3.66	8.88	0	0	0
387/1	0.01	236.00	91.51	0.07	2.83	2.83	2.83	8.62	0	0	0
339	0.01	236.00	101.67	0.37	5.30	5.30	5.30	5.30	0	0	0

2. Vertical analysis on the results from the model

The predicted operating time of each relay was compared to the operating time of the downstream and upstream relay. Table G-2 shows the points of interest from the vertical analysis and the predicted operating times.

Table G-2: Vertical analysis on predicted relay operating times after three stages of DG penetration and protection miscoordination groups for the protective devices on the MV distribution networks

Relay	TMS	Fault level (A)	Pick up current	Calculated operation time (s)	Actual relay operating time (s)before DG penetration	Relay operating times (s) after DG penetration of			Loss coordination	De-sensitization	Loss grading	Operating time ratio				Difference (de-sensitization at 35, 65 and 75)			Upstream minus current operating time at 35%, 65% and 75%			
						35% of ADMD	65% of ADMD	75% of ADMD														
423	0.01	231.00	124.70	0.11	18.45	18.45	18.45	7.36	0	0	0	0.30	0.30	0.30	0.02	0.00	0.00	-0.28	-12.87	-12.87	-12.87	-7.21
318	0.03	231.00	138.56	0.41	5.58	5.58	5.58	0.15	0	0	0	1.63	1.63	1.63	16.41	0.00	0.00	14.78	3.51	3.51	3.51	2.34
200	0.09	306.00	157.27	0.57	9.09	9.09	9.09	2.49	0	0	0	1.15	1.15	1.15	4.20	0.00	0.00	3.05	1.37	1.37	1.37	7.97
101	0.11	479.00	196.10	0.86	10.46	10.46	10.46	10.46	0	0	0	0.59	0.59	0.59	0.59	0.00	0.00	0.00	-4.31	-4.31	-4.31	-4.31
CB	0.13	909.00	299.20	0.80	6.14	6.14	6.14	6.14	0	0	0	3.08	4.53	4.70	1.56	1.45	0.17	-3.14	12.75	21.67	22.74	3.43
Trf	0.38	368.50	65.61	1.50	18.90	27.81	28.89	9.57	0	0	0											
219/1	0.05	437.00	141.54	0.31	6.23	6.23	6.23	5.95	0	0	0	1.46	1.46	1.46	0.42	0.00	0.00	-1.04	2.86	2.86	2.86	-3.46
179/1	0.05	529.00	176.46	0.32	5.90	5.90	5.90	5.69	0	0	0	1.77	1.77	1.77	1.84	0.00	0.00	0.07	4.56	4.56	4.56	4.77
87/1	0.05	1031.00	269.28	0.26	4.02	4.02	4.02	7.58	0	0	0	1.53	1.53	1.53	0.81	0.00	0.00	-0.72	2.13	2.13	2.13	-1.43
148	0.01	519.00	125.52	0.05	7.85	7.85	7.85	9.42	0	0	0	1.20	1.20	1.20	0.70	0.00	0.00	-0.50	1.57	1.57	1.57	-2.81
114	0.07	519.00	139.47	0.35	9.42	9.42	9.42	6.61	0	0	0	1.04	1.04	1.04	0.42	0.00	0.00	-0.63	0.42	0.42	0.42	-3.84
67	0.10	630.00	154.97	0.60	9.84	9.84	9.84	2.77	0	0	0	0.34	1.01	0.51	3.94	0.66	-0.50	3.43	-6.47	0.06	-4.87	8.13
1	0.24	1204.00	191.31	0.89	3.36	9.90	4.97	10.90	0	0	0	0.30	3.08	6.43	2.31	2.79	3.35	-4.12	-2.37	20.60	26.98	14.27
CB	0.37	2710.00	212.57	1.00	0.99	30.50	31.95	25.17	0	0	0	13.76	1.73	2.21	1.26	-12.03	0.48	-0.95	12.69	22.26	38.58	6.55
Trf	0.54	677.50	59.05	1.50	13.68	52.77	70.53	31.72	0	0	1											
111/1	0.05	555.00	139.47	0.25	8.86	8.86	8.86	7.79	0	0	0	0.58	0.58	0.58	0.16	0.00	0.00	-0.42	0.98	0.98	0.98	-5.02

65T/1	0.05	860.00	154.97	0.20	5.82	5.82	5.82	9.37	0	0	0										
65/176	0.05	264.00	91.51	0.33	16.16	16.16	16.16	5.31	0	0	0	0.96	0.96	0.96	2.93	0.00	0.00	1.96	-0.61	-0.61	-0.61	10.23
65/127	0.12	328.00	112.97	0.81	15.55	15.55	15.55	15.55	0	0	0	1.02	1.02	1.02	1.02	0.00	0.00	0.00	0.27	0.27	0.27	0.27
65/67	0.15	330.00	125.52	1.10	15.81	15.81	15.81	15.81	0	0	0	0.57	0.82	0.54	0.59	0.25	-0.28	0.06	-6.83	-2.85	-7.31	-6.41
65/1	0.26	888.00	154.97	1.02	8.98	12.96	8.50	9.40	0	0	0	0.84	0.58	0.89	0.80	-0.26	0.31	-0.08	-1.44	-5.42	-0.96	-1.86
48	0.17	901.00	172.18	1.31	7.55	7.55	7.55	7.55	0	0	0	0.45	1.31	0.66	1.44	0.87	-0.65	0.79	-4.18	2.35	-2.58	3.35
65/39/58	0.06	379.00	82.36	0.27	14.43	14.43	14.43	8.33	0	0	0	1.02	1.02	1.02	0.70	0.00	0.00	-0.32	0.24	0.24	0.24	-2.53
65/39/56	0.12	379.00	91.51	0.57	14.67	14.67	14.67	5.79	0	0	0	0.59	0.59	0.59	0.12	0.00	0.00	-0.47	0.01	0.01	0.01	-2.88
65/39/54	0.17	393.00	101.67	0.86	14.69	14.69	14.69	2.91	0	0	0	1.01	0.32	1.01	5.10	-0.69	0.69	4.09	0.18	-9.93	0.18	11.96
65/39/2	0.21	398.00	112.97	1.15	14.87	4.76	14.87	14.87	0	0	0	-0.79	-2.69	-0.56	-0.32	-1.90	2.13	0.24	-3.15	8.03	-6.61	-10.13
65/39/1	0.26	582.00	125.52	1.18	11.72	12.79	8.26	4.74	0	0	0	1.28	0.02	1.82	3.18	-1.26	1.80	1.35	3.33	-12.50	6.79	10.31
65/2	0.19	587.00	139.47	1.47	15.05	0.29	15.05	15.05	0	0	0	0.17	0.11	0.18	0.76	-0.05	0.06	0.58	-7.49	-11.47	-7.01	-2.26
65/64/3	0.06	473.00	125.52	0.31	12.34	12.34	12.34	7.00	0	0	0											
65/164/1	0.06	277.00	91.51	0.38	16.02	16.02	16.02	4.98	0	0	0											
65/128	0.09	277.00	101.67	0.68	16.39	16.39	16.39	16.39	0	0	0											
65/138/36	0.08	266.00	82.36	0.47	16.48	16.48	16.48	4.32	0	0	0											
65/138/1	0.06	266.00	91.51	0.77	17.07	17.07	17.07	17.07	0	0	0											
65/39/39/1	0.08	435.00	101.67	0.38	13.40	13.40	13.40	7.29	0	0	0	0.87	0.95	0.62	0.65	0.08	-0.34	0.03	-1.68	-0.61	-5.14	-2.55
65/39/53/1	0.08	398.00	91.51	0.38	14.13	14.13	14.13	7.43	0	0	0	0.83	0.91	0.58	0.64	0.08	-0.32	0.05	-2.40	-1.34	-5.87	-2.69
297	0.07	388.00	82.36	0.31	4.72	4.72	4.72	8.25	0	0	0	1.18	1.18	1.18	0.71	0.00	0.00	-0.47	0.84	0.84	0.84	-2.39
277	0.16	413.00	91.51	0.75	5.56	5.56	5.56	5.86	0	0	0	1.89	1.89	1.89	1.79	0.00	0.00	-0.10	4.97	4.97	4.97	4.66
174	0.12	413.00	101.67	1.05	10.52	10.52	10.52	10.52	0	0	0	0.81	0.81	0.81	0.43	0.00	0.00	-0.37	-2.04	-2.04	-2.04	-5.95
123	0.15	622.00	112.97	0.74	8.48	8.48	8.48	4.57	0	0	0	0.88	0.88	0.88	1.63	0.00	0.00	0.75	-1.03	-1.03	-1.03	2.88
75	0.13	829.00	125.52	0.82	7.45	7.45	7.45	7.45	0	0	0	0.20	0.20	0.20	0.96	0.00	0.00	0.76	-5.96	-5.96	-5.96	-0.31
65	0.16	1208.00	139.47	0.68	1.49	1.49	1.49	7.14	0	0	0	0.01	0.25	0.13	0.36	0.23	-0.11	0.23	-0.89	8.96	4.14	8.14
50	0.24	1335.00	154.97	0.77	0.60	10.45	5.63	15.28	0	0	0	12.60	0.72	1.34	0.49	-11.88	0.62	-0.85	6.95	-2.90	1.91	-7.73
2	0.18	1833.00	172.18	0.97	0.95	1.13	0.95	0.94	0	0	0	0.53	17.33	17.95	26.87	16.80	0.62	8.92	-0.44	18.46	16.03	24.31
1	0.30	3510.00	191.31	0.71	0.50	19.59	16.97	25.25	0	0	0	0.02	0.83	0.94	0.68	0.81	0.11	-0.26	0.52	17.69	25.30	5.49
CB	0.42	3510.00	212.57	1.01	1.02	37.28	42.27	30.74	0	0	0	0.26	1.85	3.24	1.11	1.58	1.39	-2.13	11.17	47.68	106.80	20.18
Trf	0.80	787.50	59.05	2.10	12.19	84.96	149.06	50.92	0	1	1					1.00	1.00					

2/2	0.05	3510.00	154.97	0.11	0.01	0.01	0.01	12.32	0	0	0	80.01	95.64	80.01	0.08	15.63	-15.63	-79.93	0.93	1.12	0.93	-11.38
74/21	0.05	490.00	101.67	0.22	11.98	11.98	11.98	8.78	0	0	0	1.24	1.24	1.24	0.09	0.00	0.00	-1.15	2.88	2.88	2.88	-8.01
74/27	0.11	337.00	112.96	0.70	14.86	14.86	14.86	0.77	0	0	0	1.30	1.30	1.30	24.95	0.00	0.00	23.65	4.42	4.42	4.42	18.50
74/9	0.14	292.00	125.52	1.17	19.28	19.28	19.28	19.28	0	0	0	0.15	0.15	0.15	0.15	0.00	0.00	0.00	-11.83	-11.83	-11.83	-11.83
107/1	0.05	764.00	112.97	0.18	6.23	6.23	6.23	10.05	0	0	0											
140/6	0.05	393.00	91.51	0.24	10.92	10.92	10.92	8.21	0	0	0	1.02	1.02	1.02	0.65	0.00	0.00	-0.37	0.23	0.23	0.23	-2.87
140/1	0.11	393.00	101.67	0.54	11.15	11.15	11.15	5.34	0	0	0	0.16	0.16	0.16	0.08	0.00	0.00	-0.07	-2.67	-2.67	-2.67	-0.77
197/1	0.05	511.00	91.51	0.20	6.71	6.71	6.71	9.39	0	0	0											
198/1	0.05	507.00	91.51	0.20	6.73	6.73	6.73	9.36	0	0	0											
220/6/1	0.05	370.00	82.36	0.23	8.09	8.09	8.09	8.44	0	0	0	0.15	0.15	0.15	0.10	0.00	0.00	-0.04	0.24	0.24	0.24	-2.61
220/1	0.11	370.00	91.51	0.53	8.33	8.33	8.33	5.83	0	0	0	0.18	0.18	0.18	0.18	0.00	0.00	0.00	2.19	2.19	2.19	4.69
235/1	0.05	397.00	91.51	0.24	6.91	6.91	6.91	8.26	0	0	0											
259/5/1	0.05	383.00	82.36	0.22	6.26	6.26	6.26	8.61	0	0	0	1.04	1.04	1.04	0.72	0.00	0.00	-0.32	0.24	0.24	0.24	-2.41
259/1	0.11	383.00	91.51	0.52	6.50	6.50	6.50	6.19	0	0	0	0.17	0.17	0.17	0.17	0.00	0.00	0.00	4.02	4.02	4.02	4.33
280/8/1	0.05	358.00	74.12	0.22	5.90	5.90	5.90	8.78	0	0	0	0.10	0.10	0.10	0.11	0.00	0.00	0.01	0.24	0.24	0.24	-2.19
280/1	0.11	358.00	82.36	0.52	6.15	6.15	6.15	6.59	0	0	0	1.60	1.60	1.60	1.49	0.00	0.00	-0.11	3.68	3.68	3.68	3.24
280/57/1	0.05	284.00	74.12	0.26	7.09	7.09	7.09	7.55	0	0	0											
49/373	0.09	236.00	91.51	0.63	17.75	17.75	17.75	0.00	0	0	0	0.24	0.24	0.24	0.24	0.00	0.00	0.00	-2.07	-2.07	-2.07	15.67
49/250	0.11	387.00	112.97	0.94	15.68	15.68	15.68	15.68	0	0	0	0.23	0.23	0.23	0.23	0.00	0.00	0.00	-0.54	-0.54	-0.54	-0.54
49/148	0.12	400.00	125.52	0.90	15.14	15.14	15.14	15.14	0	0	0	0.18	0.18	0.18	0.18	0.00	0.00	0.00	-3.29	-3.29	-3.29	-3.29
49/79	0.12	590.00	139.47	0.82	11.85	11.85	11.85	11.85	0	0	0	1.28	1.51	1.28	0.85	0.23	-0.23	-0.43	1.72	2.65	1.72	-1.74
49/29/21	0.06	803.00	125.52	0.22	7.87	7.87	7.87	9.90	0	0	0	0.12	0.12	0.12	0.13	0.00	0.00	0.02	0.08	0.08	0.08	-0.62
49/29/1	0.13	803.00	139.47	0.52	7.96	7.96	7.96	9.28	0	0	0	0.09	0.07	0.09	0.17	-0.01	0.01	0.08	-1.80	-2.74	-1.80	2.36
49/1	0.21	1052.00	154.97	0.75	6.15	5.22	6.15	11.65	0	0	0	0.01	0.02	0.01	0.01	0.00	0.00	0.00	-5.21	-4.09	-5.21	-10.71
49/63/11	0.06	623.00	139.47	0.28	10.43	10.43	10.43	8.16	0	0	0											
49/110/1	0.06	621.00	125.52	0.26	10.67	10.67	10.67	8.74	0	0	0											
49/140/1	0.06	487.00	125.52	0.31	12.78	12.78	12.78	7.22	0	0	0											
49/147/1	0.06	464.00	125.52	0.32	13.15	13.15	13.15	6.85	0	0	0											
49/166/1	0.06	509.00	112.97	0.27	12.61	12.61	12.61	8.21	0	0	0											
49/175/20	0.06	421.00	101.67	0.29	14.17	14.17	14.17	7.67	0	0	0											

49/175/1	0.11	421.00	112.97	0.59	14.39	14.39	14.39	4.52	0	0	0										
49/195/7	0.06	404.00	101.67	0.30	14.44	14.44	14.44	7.39	0	0	0										
49/195/1	0.11	404.00	112.97	0.60	14.67	14.67	14.67	4.00	0	0	0										
49/252/1	0.06	393.00	101.67	0.31	14.62	14.62	14.62	7.19	0	0	0										
49/195/18	0.06	404.00	82.36	0.26	14.70	14.70	14.70	8.69	0	0	0										
49/357/1	0.12	250.00	91.51	0.83	17.83	17.83	17.83	17.83	0	0	0										
49/261	0.12	250.00	101.67	1.13	18.30	18.30	18.30	18.30	0	0	0	0.86	0.86	0.86	0.86	0.00	0.00	0.00	-2.62	-2.62	-2.62	-2.62
49/471/1	0.05	235.00	74.12	0.30	17.51	17.51	17.51	6.18	0	0	0	1.04	1.04	1.04	2.93	0.00	0.00	1.90	0.62	0.62	0.62	11.95
49/478/1	0.05	229.00	74.12	0.31	17.61	17.61	17.61	5.96	0	0	0	0.21	0.21	0.21	0.21	0.00	0.00	0.00	0.52	0.52	0.52	12.17
49/470	0.05	229.00	82.36	0.61	18.13	18.13	18.13	18.13	0	0	1	0.98	0.98	0.98	0.00	0.00	0.00	-0.98	-0.38	-0.38	-0.38	-18.12
4	0.26	2877.00	154.97	0.60	0.36	11.82	7.35	22.69	0	0	0	0.01	0.35	0.36	0.32	0.34	0.02	-0.04	0.45	18.55	24.41	5.59
3	0.37	2877.00	172.18	0.90	0.80	30.37	31.77	28.28	0	0	0	1.78	1.52	1.79	1.17	-0.26	0.28	-0.62	0.63	15.68	25.21	4.80
2	0.48	2877.00	191.31	1.20	1.43	46.05	56.98	33.08	0	0	0	1.56	1.28	1.44	1.12	-0.29	0.16	-0.32	0.81	12.78	24.86	3.90
CB	0.57	2877.00	212.57	1.50	2.24	58.83	81.84	36.97	0	0	0	0.15	1.15	2.56	0.72	1.00	1.41	-1.84	11.52	45.49	150.99	28.38
Trf	1.04	719.25	59.05	2.83	13.76	104.32	232.83	65.35	0	1	1					1.00	1.00					
86/1	0.05	719.00	125.52	0.20	11.15	11.15	11.15	9.49	0	0	0	0.88	0.88	0.88	1.03	0.00	0.00	0.15	-1.33	-1.33	-1.33	0.33
111/98	0.05	307.00	112.97	0.35	13.26	13.26	13.26	4.68	0	0	0	0.12	0.12	0.12	0.04	0.00	0.00	-0.08	-2.11	-2.11	-2.11	-0.59
200/1	0.05	390.00	112.97	0.28	8.17	8.17	8.17	6.85	0	0	0	0.11	0.11	0.11	0.11	0.00	0.00	0.00	2.63	2.63	2.63	3.95
304/3	0.05	270.00	91.51	0.32	6.13	6.13	6.13	5.53	0	0	0											
314/1	0.05	261.00	91.51	0.33	5.87	5.87	5.87	5.19	0	0	0											
202	0.05	261.00	112.97	0.63	10.80	10.80	10.80	10.80	0	0	0	0.91	0.91	0.91	0.91	0.00	0.00	0.00	-0.98	-0.98	-0.98	-0.98
61/7/1	0.06	895.00	112.97	0.20	0.04	0.04	0.04	10.60	0	0	0	331.27	331.27	331.27	0.73	0.00	0.00	-330.54	13.05	13.05	13.05	-2.89
61/40/1	0.06	618.00	112.97	0.24	10.41	10.41	10.41	9.23	0	0	0	1.26	1.26	1.26	0.83	0.00	0.00	-0.42	2.67	2.67	2.67	-1.53
61/2	0.13	618.00	125.52	0.54	13.09	13.09	13.09	7.71	0	0	0	0.14	0.14	0.14	0.14	0.00	0.00	0.00	0.89	0.89	0.89	6.27
111/29/1	0.06	452.00	112.97	0.30	10.94	10.94	10.94	7.44	0	0	0	0.11	0.11	0.11	0.04	0.00	0.00	-0.07	0.21	0.21	0.21	-3.35
111/1	0.11	452.00	125.52	0.60	11.15	11.15	11.15	4.08	0	0	0	0.10	0.10	0.10	0.10	0.00	0.00	0.00	-1.33	-1.33	-1.33	5.74
5	0.13	565.00	139.47	0.81	13.98	13.98	13.98	13.98	0	0	0	0.03	0.85	0.53	1.62	0.82	-0.32	1.10	-13.62	-2.16	-6.63	8.71
146/12/1	0.06	519.00	101.67	0.25	8.59	8.59	8.59	8.89	0	0	0	1.02	1.02	1.02	0.78	0.00	0.00	-0.24	0.18	0.18	0.18	-1.93
146/1	0.12	519.00	112.97	0.55	8.77	8.77	8.77	6.96	0	0	0	1.12	1.12	1.12	1.41	0.00	0.00	0.29	1.05	1.05	1.05	2.86
113	0.10	572.00	125.52	0.82	9.82	9.82	9.82	9.82	0	0	0	1.42	1.42	1.42	1.42	0.00	0.00	0.00	4.15	4.15	4.15	4.15

356/1	0.01	265.00	91.51	0.07	3.66	3.66	3.66	8.88	0	0	0	1.45	1.45	1.45	0.60	0.00	0.00	-0.85	1.64	1.64	1.64	-3.58
387/1	0.01	236.00	91.51	0.07	2.83	2.83	2.83	8.62	0	0	0	1.87	1.87	1.87	0.62	0.00	0.00	-1.26	2.47	2.47	2.47	-3.32
339	0.01	236.00	101.67	0.37	5.30	5.30	5.30	5.30	0	0	0	2.04	2.04	2.04	2.04	0.00	0.00	0.00	5.50	5.50	5.50	5.50