Global Sensitivity Analysis of Fault Location Algorithms

by

Hoong Boon Ooi

A thesis submitted to the School of Electrical and

Electronic Engineering of the University of Adelaide in

partial fulfillment of the requirements for the degree of

Master of Engineering Science

In

Electrical Engineering

Adelaide, Australia

November, 2008

Contents

Abstract

Transmission lines of any voltage level are subject to faults. To speed up repairs and restoration of power, it is important to know where the fault is located. A fault location algorithm's result is influenced by a series of modeling equations, setting parameters and system factors reflected in voltage and current inputs. The factors mentioned are subject to sources of uncertainty including measurement and signal processing errors, setting errors and incomplete modeling of a system under fault conditions. These errors have affected the accuracy of the distance to fault calculation. Accurate fault location reduces operating costs by avoiding lengthy and expensive patrols. Accurate fault location speeds up repairs and restoration of lines, ultimately reducing revenue loss caused by outages. In this thesis, we have reviewed the fault location algorithms and also how the uncertainty affects the results of fault location.

Sensitivity analysis is able to analyze how the variation in the output of the fault location algorithms can be allocated to the variation of uncertain factors. In this research, we have used global sensitivity analysis to determine the most contributed uncertain factors and also the interaction of the uncertain factors. We have chosen Analysis of Variance (ANOVA) decomposition as our global sensitivity analysis. ANOVA decomposition shows us the insight of the fault location, such as relations between uncertain factors of the fault location.

Quasi regression technique has also been used to approximate a function. In this research, the transmission line fault location system is fitted into the ANOVA decomposition using quasi regression. From the approximate function, we are able to get the variance of the sensitivity of fault location to uncertain factors using Monte Carlo method. In this research, we have designed novel methodology to test the fault location algorithms and compare the fault location algorithms. In practice, such analysis not only helps in selecting the optimal locator for a specific application, it also helps in the calibration process.

Statement of Originality

I hereby declare that this is an original thesis and is entirely my own work under the guidance and advice of my supervisor Dr. Rastko Zivanovic. This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the Adelaide University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

Hoong Boon Ooi

Nov 2008

Acknowledgements

To begin with, I would like to thank my family, especially my Dad and Mom for supporting me financially and morally. They have encouraged me whenever I have tough time working on this thesis. You could not have given any more than you have done to me. Your love, support and trust is the invaluable wealth that I have. It has carried me for the years I lived through, and will continue to be the source of my encouragement and strength for my whole life.

I am indebted to many people who helped during the planning and writing of this thesis. I am grateful to my supervisor, Dr. Rastko Zivanovic for his broad knowledge and skills. You have shown me what patience means throughout the whole time of my study of this degree. Thank you for your million and one acts of service.

I would also like to thank many colleagues who had generously helped me out in this thesis and also giving me so many valuable comments. Especially, I would like to thank Ms. Debra James who explained to me a lot in mathematics and statistics which helped me so much in this degree. You have also encouraged me not to give up easily during the time I worked on the thesis.

In conclusion, I am grateful to the people that had helped me through the whole time of my study.

List of Tables

List of Figures

Nomenclatures

SA	Sensitivity Analysis
FLA	Fault Location Algorithm
DFL	Distance-to-Fault Locator
GSA	Global Sensitivity Analysis
SRC	Standardized Regression Coefficients
PDF	Probability Density Function
FP	Factor Prioritization
FF	Factor Fixing
VC	Variance Cutting
FM	Factors Mapping
ANOVA	Analysis of Variance
MC	Monte Carlo
QMC	Quasi Monte Carlo
ISE	Integrated Square Error
MSE	Mean Squared Error
LOF	Lack of Fit

Chapter 1.

Introduction

Unexpected faults can occur in power system transmission lines and obstruct the correct operation of the system. As these incidents are unforeseeable, electric utilities use protection systems to prevent these faults from being propagated as well as to avoid anomalous consequences in the line operation. With these targets in mind, the protection systems are provided with devices which isolate the faulty line if any voltage or current disturbance is detected. However, the service must be urgently restored as the transmission line in which a fault occurs cannot be kept isolated indefinitely. As designing a totally reliable system is not possible, for both technical and economic reasons, it has been necessary to develop a number of technologies aimed at locating faults in transmission lines and making the network operate correctly.

New technological developments in such fields as telecommunications, tele-control and electronics applied to data acquisition and electric protection devices are having an increasing influence on fault location procedures. It is also possible to receive information on voltage values, current values, operation of protection devices and electric equipment, etc. from substations located at line terminals. This makes it interesting to develop data processing algorithms which may be implemented using microprocessors at the control centre of the electric utility, at remote terminals located in substations and at the protection devices themselves. This enables a fault to be located without requiring a visual inspection of the line.

To speed up repairs and restoration of power, it is important to know where the fault is located. A fault location algorithm's result is influenced by a series of modeling equations, setting parameters and system factors reflected in voltage and current inputs. The factors are subject to sources of uncertainty including measurement and signal processing errors, setting errors and incomplete modeling of a system under fault conditions. These errors have affected the accuracy of the distance to fault calculation. Accurate fault location reduces operating costs by avoiding lengthy and expensive patrols. Accurate fault location speeds up repairs and restoration of lines, ultimately reducing revenue loss caused by outages.

In general, the techniques can be classified as follows [8]:

-microprocessor devices with different input signals and principle of operation:

- Impedance techniques: one-terminal, two-terminal (or multi-terminal),
- Traveling wave techniques,

-short circuit analysis software,

-customer calls,

-line inspection,

-lightning detection system,

-terminal and tracer methods for cables.

Hence, the objective of my research is to assess the uncertainties associated with the influencing factors by applying sensitivity analysis on fault location algorithm. From the sensitivity analysis, we are able to determine which uncertain factor that mostly contributes to the fault distance estimation. Such analysis can help in the calibration process. In practice, the electricity utilities are able to compare all different fault locator devices using sensitivity analysis, and select the robust, less sensitive fault locator in specific fault cases.

Sensitivity analysis (SA) is popular in financial applications, risk analysis, signal processing, neural networks and any area where models are developed [17]. Sensitivity analysis can also be used in model-based policy assessment studies. Sensitivity analysis is the study of how the variation (uncertainty) in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of a model. In more general terms uncertainty and sensitivity analyses investigate the robustness of a study.

There are two types of sensitivity analysis, local and global. In the first approach, the local response of the fault location output, obtained by varying input factors one at the time, is investigated while holding the others fixed to some fixed value [17]. This involves calculation of partial derivatives. The second global sensitivity analysis, analyzed the whole set of potential input factors and aim to give an overall indication of the way that the outcome varies, and in particular how the output varies in response to the input variations within the range of parameter uncertainty [17].

In this research project, global sensitivity analysis (SA) is applied in evaluation of the fault location algorithms. Global SA is able to determine not only the factor that mostly contributes to the distance to the fault result but also interactions between the factors. Analysis of Variance (ANOVA) decomposition has been chosen as a tool for global sensitivity analysis of the fault location [26]. The ANOVA allows us to quantify the notion that some variables and interactions are much more important than others. It is a sampling based process which the fault simulator and fault location algorithm are executed repeatedly for the values sampled of the input factors. ANOVA decompositions are actually a form of structure which decomposes the model function [16] which in our case, the function consists of fault simulator and fault location algorithm. Chapter 3 describes how we use the ANOVA decomposition in our research.

In order to fit our systems into the ANOVA decomposition, we choose Quasi Regression technique and Quasi Monte Carlo methods [28]. Quasi-regression is a technique for constructing approximation of function, based on observations sampled from the domain of the function to be approximated. In quasi-regression, the function is expanded in an orthogonal basis with an infinite number of coefficients [29, 30]. If we knew the coefficients, we are able to describe the extent to which the function depends on several of its input factors. We use Monte Carlo methods to estimate these coefficients. Monte Carlo methods are a class of computational algorithms that rely on repeated random sampling [33]. Monte Carlo methods are often used when simulating physical and mathematical systems. Because of their reliance on repeated computation and random or pseudo-random numbers, Monte Carlo methods are most suited to calculation by a computer. Monte Carlo methods tend to be used when it is infeasible or impossible to compute an exact result with a deterministic algorithm. In our research, we have chosen quasi-Monte Carlo method because it is more efficient at numerical integration compared to Monte Carlo method [19]. In chapter 4, we discussed the quasi regression and Monte Carlo methods in detail.

1.1. Description of the Fault Location Sensitivity Analysis Procedure

Figure 1.1 shown below illustrates the methodology of the sensitivity analysis. In the first stage, the samples of system factors (fault location, fault resistance, source impedance, etc) are generated with sampling strategy [33]. The points that have been generated are used to run the power system fault simulator to produce voltages and currents (V_s, I_s in Figure 1.1) at fault locator installation position.

These voltages and currents are then sent to the fault locator device. The setting factors are also sampled and send to the fault locator. In the last stage, the fault locator calculates the fault location value for each point in the factor space.

Function between the collected fault location values and the corresponding system factors is formulated using ANOVA decomposition [26]. The ANOVA is able to determine which system factors contribute the most and also determine the interaction between the system factors that influence the location result.

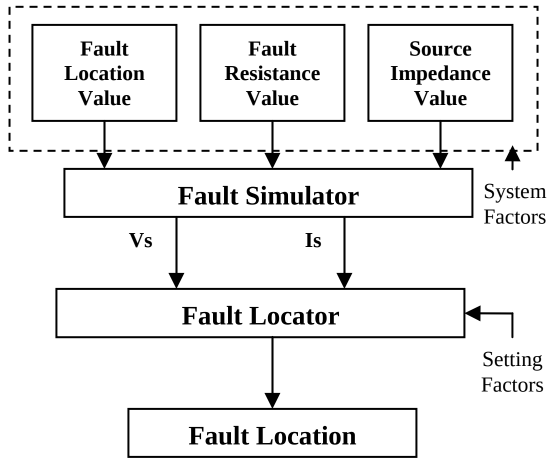


Figure 1.1: Methodology of the Sensitivity Analysis

There are research results that use similar statistical algorithms to analyze the distance protection algorithm [24, 25]. Their objectives are to increase the efficiency of the protection operation and to obtain more accurate fault location. There are no references that up to my knowledge uses global SA techniques and ANOVA in fault location testing and calibration.

Chapter 2:

Review of Fault Location Algorithms

Fault location techniques are used in power systems for accurate pinpointing of the fault location. Benefits of accurate fault location are considered as follows [1]:

-fast repair to restore power system,

-improves system availability and performance as well as reduces operating costs,

-saves time and expense of crew searching in bad weather and tough terrain,

-aids crew in disturbance diagnostics by:

- Identifying temporary faults,
- Detecting weak spots.

Variety of fault location techniques is used to achieve the aim and the above-specified benefits. In our research, we used impedance based techniques: one-terminal and two-terminal. Fault location algorithms are based on current and voltage data from one substation only, as well as algorithms, which make use of the information from the substations at both ends of the line. The single ended approach to transmission line fault location is important, as it is less expensive than the double-ended approach (no communication link required between the ends of the transmission line) and more reliable (the ability to operate requires only that the local end equipment is in operation). The one-terminal data algorithms determine the impedance and, as a result, distance to the fault. However, several factors affect the accuracy of the distance to the fault calculation. One of the main factors result from the combined effect of the load, fault resistance and equivalent impedances of the power systems connected to the ends of transmission line. The value of the fault resistance may be high, especially for ground faults and, accordingly, the accuracy of fault location may be insufficient.

The operation experience of one terminal data based fault locators shows that in the majority of cases ($\approx 80\%$) the accuracy of fault location may be considered as satisfactory [9]. The fault location error may comprise 1-2% of the monitored line length. In relatively infrequent cases ($\approx 10\%$) the error may reach 5% and even more. Fault location error is the percentage error in fault location estimate based on the total line length. The possibility of appearance of considerable errors weakens the confidence in the results of the measurements and causes the necessity to search fault location on the long line segments.

In real practice of power systems operation the length of the segment subjected to the inspection takes into account the possibility of considerable errors even in those cases when it is not likely that the error will appear. Various methods were developed during the recent years to detect the location of fault on a transmission line [8]. In this research, we focus on certain impedance-based fault location methods and provide fault location results from simulated faults.

2.1. Faults on Transmission Line

Short circuits occur in three-phase power systems as follows, in order of frequency of occurrence: single line-to-ground, line-to-line, double line-to-ground, and balanced three-phase faults. Other types of faults include one-conductor-open and two-conductors-open, which can occur when conductors break or when one or two phases of a circuit breaker inadvertently open. This section describes the symmetrical and unsymmetrical faults in three-phase power system. We have implemented a simulation program that simulate the faults in power system.

Symmetrical Fault

In power engineering, the symmetrical fault is a fault which affects each of the three-phases equally. Although the three-phase short circuit occurs the least, it was considered first, because of its simplicity. When a balanced three-phase fault occurs in a balanced three-phase system, there is only positive-sequence fault current; the zero-, positive-, and negative-sequence networks are completely uncoupled.

Unsymmetrical Faults

When an unsymmetrical fault occurs in an otherwise balanced system, the sequence networks are interconnected only at the fault location [4]. As such, the computation of fault currents is greatly simplified by the use of sequence networks. All faults have two components of fault currents; an ac or symmetrical component, including subtransient, transient, and steady-state currents; and a dc component.

System Representation

A three-phase power system is represented by its sequence networks. I make the following assumptions [3]:

1. The power system operates under balanced steady-state conditions before the fault occurs. Thus the zero-, positive-, and negative- sequence networks are uncoupled before the fault occurs. During unsymmetrical faults they are interconnected only at the fault location.
2. Prefault load current is neglected. Because of this, the positive- sequence internal voltages of all machines are equal to the prefault voltage V_F . Therefore, the prefault voltage at each bus in the positive-sequence network equals V_F .
3. Transformer winding resistances and shunt admittances are neglected.
4. Transmission-line series resistances and shunt admittances are neglected.
5. Synchronous machine armature resistance, saliency, and saturation are neglected.
6. All nonrotating impedance loads are neglected.

7. Induction motors are either neglected (especially for motors rated 50hp or less) or represented in the same manner as synchronous machines.

Note that these assumptions are made for simplicity, and in practice should not be made for all cases.

2.1.1. Three Phase Short Circuit

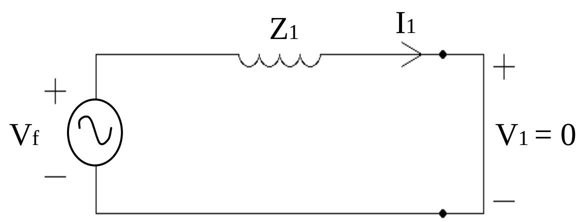


Figure 2.1: Positive Sequence Networks Representing the Three Phase Short Circuit

The fault currents are balanced in three-phase faults and have only a positive-sequence component. Therefore we work only with the positive-sequence network when calculating three-phase fault currents.

The positive-sequence fault current:

$$I_1 = \frac{V_F}{Z_1} \tag{2.1}$$

Also, the zero-sequence current and negative-sequence current are both zero. Therefore, the subtransient fault currents in each phase are,

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{bmatrix} \begin{bmatrix} 0 \\ I_1 \\ 0 \end{bmatrix} \tag{2.2}$$

Where $a = \exp(j\frac{2\pi}{3})$.

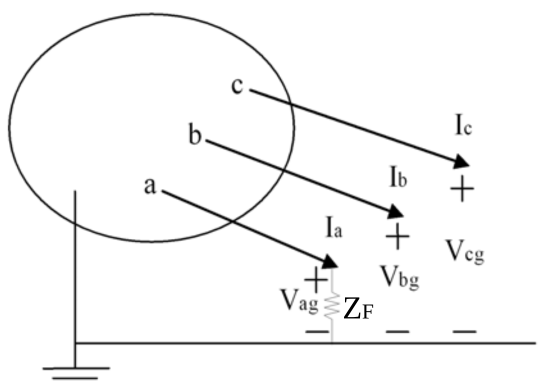
The sequence components of the line-to-ground voltages at the fault terminals are,

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0 \\ V_F \\ 0 \end{bmatrix} - \begin{bmatrix} Z_0 & 0 & 0 \\ 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \end{bmatrix} \begin{bmatrix} 0 \\ I_1 \\ 0 \end{bmatrix} \tag{2.3}$$

The sequence fault voltages are $V_0 = V_1 = V_2 = 0$, which must be true since $V_{ag} = V_{bg} = V_{cg} = 0$.

2.1.2. Single Line-to-Ground Fault

Consider a single line-to-ground fault from phase *a* to ground at the general three-phase bus shown in Figure 2.2(a). For generality, we include a fault impedance Z_f. In the case of a bolted fault, $Z_f = 0$, whereas for an arcing fault, Z_f is the arc impedance [6, 7]. In the case of a transmission-line insulator flashover, Z_f includes the total fault impedance between the line and ground, including the impedances of the arc and the transmission tower, as well as the tower footing if there are no neutral wires.



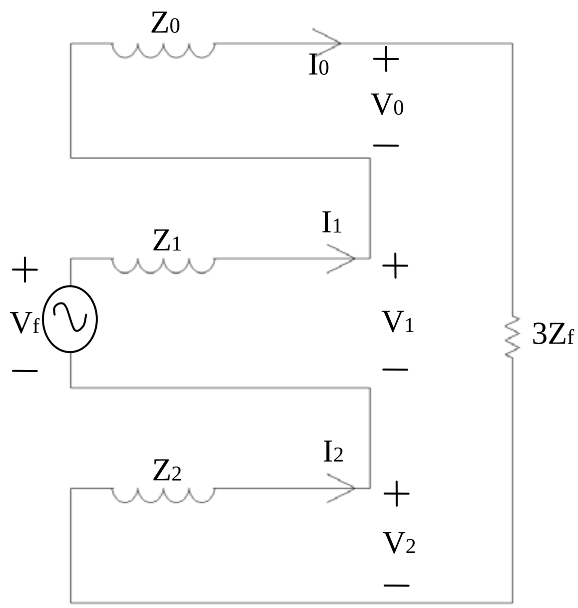
Fault conditions in phase domain:

$V_{ag} = Z_f I_a$

$I_b = I_c = 0$

Figure 2.2(a): Single Line-to-Ground Fault

The relations to be derived here apply only to a single line-to-ground fault on phase *a*. However, since any of the three phases can be arbitrarily labeled phase *a*, we do not consider single line-to-ground faults on other phases [3].



Fault conditions in sequence domain:

$I_0 = I_1 = I_2$

$(V_0 + V_1 + V_2) = 3Z_f I_1$

Figure 2.2(b): Interconnected Sequence Networks

From Figure 2.2(a):

Fault conditions in phase domain / Single line-to-ground fault:

$$I_b = I_c = 0 \tag{2.4}$$

$$V_{ag} = Z_f I_a \tag{2.5}$$

We now transform (2.4) and (2.5) to the sequence domain.

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{bmatrix}\begin{bmatrix} I_a \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} I_a \\ I_a \\ I_a \end{bmatrix} \tag{2.6}$$

where $a = 1\angle 120° = \frac{-1}{2} + j\frac{\sqrt{3}}{2}$

Also, using equation below:

$$(V_0 + V_1 + V_2) = Z_f (I_0 + I_1 + I_2) \tag{2.7}$$

From (2.6) and (2.7):

Fault conditions in sequence domain, Single line-to-ground fault:

$$I_0 = I_1 = I_2 \tag{2.8}$$

$$(V_0 + V_1 + V_2) = (3Z_f)I_1 \tag{2.9}$$

Equations (2.8) and (2.9) can be satisfied by interconnecting the sequence networks in series at the fault terminals through the impedance $(3Z_f)$ as shown in Figure 2.2(b). From this figure, the sequence components of the fault currents are:

$$I_0 = I_1 = I_2 = \frac{V_f}{Z_0 + Z_1 + Z_2 + (3Z_f)} \tag{2.10}$$

Transforming (2.10) to the phase domain,

$$I_a = I_0 + I_1 + I_2 = 3I_1 = \frac{3V_f}{Z_0 + Z_1 + Z_2 + (3Z_f)} \tag{2.11}$$

Note also,

$$I_b = (I_0 + a^2 I_1 + aI_2) = (1 + a^2 + a)I_1 = 0 \tag{2.12}$$

$$I_c = (I_0 + aI_1 + a^2 I_2) = (1 + a + a^2)I_1 = 0 \tag{2.13}$$

The sequence components of the line-to-ground voltages at the fault are determined from the equation below,

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0 \\ V_f \\ 0 \end{bmatrix} - \begin{bmatrix} Z_0 & 0 & 0 \\ 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} \tag{2.14}$$

The line-to-ground voltages at the fault can then be obtained by transforming the sequence voltages to the phase domain.

2.1.3. Line-to-Line Fault

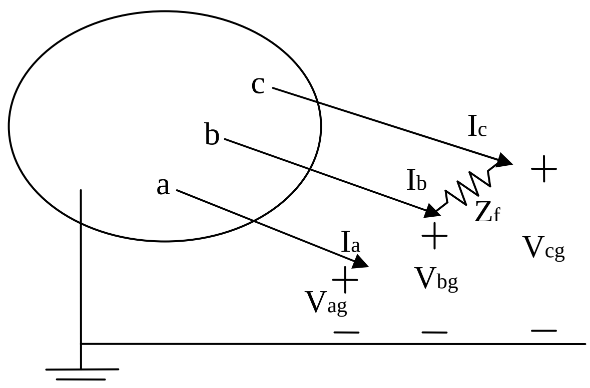
Consider a line-to-line fault from phase *b* to *c*, shown in Figure 2.3(a). Again, we include a fault impedance Z_f for generality. From Figure 2.3(a):

Fault conditions in phase domain
Line-to-line fault

$$I_a = 0 \tag{2.15}$$

$$I_c = -I_b \tag{2.16}$$

$$V_{bg} - V_{cg} = Z_f I_b \tag{2.17}$$



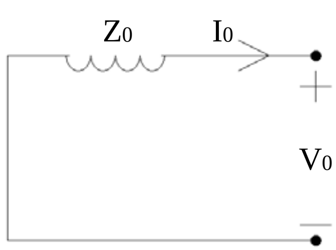
Fault conditions in phase domain:

$I_a = 0$

$I_c = -I_b$

$V_{bg} - V_{cg} = Z_f I_b$

Figure2.3(a) Line-to-Line Fault



Fault conditions in sequence domain:

$I_0 = 0$

$I_2 = -I_1$

$(V_1 - V_2) = Z_f I_1$

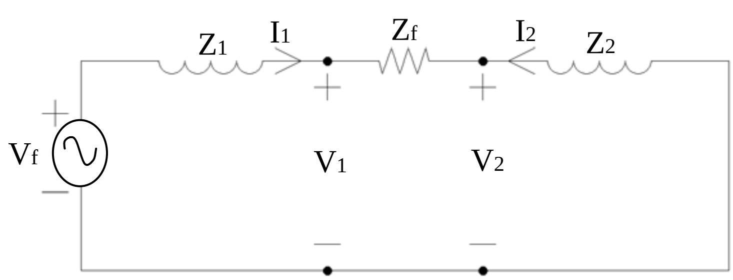


Figure2.3(b) Interconnected Sequence Networks

We transform (2.15) – (2.17) to the sequence domain.

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{bmatrix}\begin{bmatrix} 0 \\ I_b \\ -I_b \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{3}(a-a^2)I_b \\ \frac{1}{3}(a^2-a)I_b \end{bmatrix} \tag{2.18}$$

Using equation below,

$$(V_0 + a^2V_1 + aV_2) - (V_0 + aV_1 + a^2V_2) = Z_f(I_0 + a^2I_1 + aI_2) \tag{2.19}$$

Noting from (2.18) that $I_0 = 0$ and $I_2 = -I_1$, (2.19) simplifies to

$$(a^2-a)V_1 - (a^2-a)V_2 = Z_f(a^2-a)I_1$$

or

$$V_1 - V_2 = Z_f I_1 \tag{2.20}$$

Therefore, from (2.18) and (2.20):

Fault conditions in sequence domain

Line-to-line fault

$$I_0 = 0 \tag{2.21}$$

$$I_2 = -I_1 \tag{2.22}$$

$$(V_1 - V_2) = Z_f I_1 \tag{2.23}$$

Equation (2.21) – (2.23) are satisfied by connecting the positive- and negative-sequence networks in parallel at the fault terminals through the fault impedance Z_f, as shown in Figure 2.3(b). From this figure, the fault currents are:

$$I_1 = -I_2 = \frac{V_f}{(Z_1 + Z_2 + Z_f)} \qquad I_0 = 0 \tag{2.24}$$

Transforming (2.24) to the phase domain and using the identity $(a^2 - a) = -j\sqrt{3}$, the fault current in phase *b* is

$$I_b = I_0 + a^2I_1 + aI_2 = (a^2-a)I_1$$

$$= -j\sqrt{3}I_1 = \frac{-j\sqrt{3}V_f}{(Z_1 + Z_2 + Z_f)} \tag{2.25}$$

Note also,

$$I_a = I_0 + I_1 + I_2 = 0 \tag{2.26}$$

and

$$I_c = I_0 + aI_1 + a^2 I_2 = (a - a^2) I_1 = -I_b \tag{2.27}$$

which verify the fault conditions given by (2.15) and (2.16).

2.1.4. Double Line-to-Ground Fault

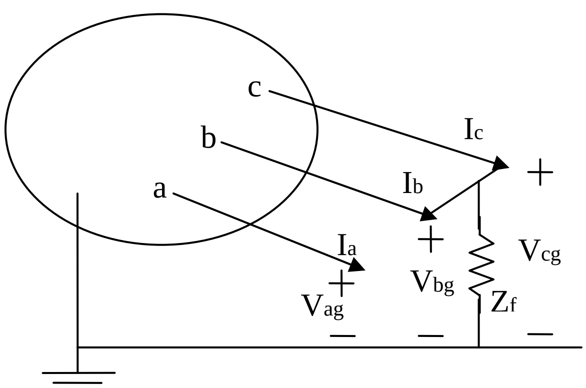
A double line-to-ground fault from phase *b* to phase *c* through fault impedance Z_f to ground is shown in Figure 2.4(a). From this figure:

Fault conditions in phase domain
Double Line-to-ground fault

$$I_a = 0 \tag{2.28}$$

$$V_{cg} = V_{bg} \tag{2.29}$$

$$V_{bg} = Z_f (I_b + I_c) \tag{2.30}$$

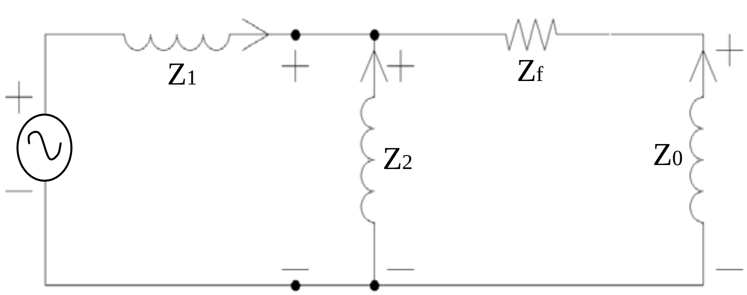


Fault conditions in phase domain:

$I_a = 0$

$V_{bg} = V_{cg} = Z_f (I_b + I_c)$

Figure 2.4(a) Double Line-to-Ground Fault



Fault conditions in sequence domain:

$I_0 + I_1 + I_2 = 0$

$V_0 - V_1 = (3Z_f) I_0$

$V_1 = V_2$

Figure 2.4(b) Interconnected Sequence Networks

Transforming (2.28) to the sequence domain,

$$I_0 + I_1 + I_2 = 0 \tag{2.31}$$

Also,

$$(V_0 + aV_1 + a^2V_2) = (V_0 + a^2V_1 + aV_2)$$

Simplifying:

$$(a^2 - a)V_2 = (a^2 - a)V_1$$

or

$$V_2 = V_1 \tag{2.32}$$

Now,

$$(V_0 + a^2V_1 + aV_2) = Z_f(I_0 + a^2I_1 + aI_2 + I_0 + aI_1 + a^2I_2) \tag{2.33}$$

Using (2.32) and the identity $a^2 + a = -1$ in (2.33),

$$(V_0 - V_1) = Z_f(2I_0 - I_1 - I_2) \tag{2.34}$$

From (2.31), $I_0 = -(I_1 + I_2)$; therefore, (2.34) becomes

$$V_0 - V_1 = (3Z_f)I_0 \tag{2.35}$$

From (2.31), (2.32) and (2.35), we summarize:

Fault conditions in sequence domain
Double line-to-ground fault

$$I_0 + I_1 + I_2 = 0 \tag{2.36}$$

$$V_2 = V_1 \tag{2.37}$$

$$V_0 - V_1 = (3Z_f)I_0 \tag{2.38}$$

Equation (2.36)-(2.38) are satisfied by connecting the zero-, positive-, and negative-sequence networks in parallel at the fault terminal; additionally, $(3Z_f)$ is included in series with the zero-sequence network. The positive-sequence fault current is,

$$I_1 = \frac{V_f}{Z_1 + [Z_2 // (Z_0 + 3Z_f)]} = \frac{V_f}{Z_1 + \left[\dfrac{Z_2(Z_0 + 3Z_f)}{Z_2 + Z_0 + 3Z_f}\right]} \tag{2.39}$$

Using current division, the negative- and zero-sequence fault currents are

$$I_2 = (-I_1)\left(\frac{Z_0 + 3Z_f}{Z_0 + 3Z_f + Z_2}\right) \tag{2.40}$$

$$I_0 = (-I_1)\left(\frac{Z_2}{Z_0 + 3Z_f + Z_2}\right) \tag{2.41}$$

These sequence fault currents can be transformed to the phase domain via,

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} \tag{2.42}$$

Also, the sequence components of the line-to-ground voltages at the fault are given by,

$$\begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0 \\ V_f \\ 0 \end{bmatrix} - \begin{bmatrix} Z_0 & 0 & 0 \\ 0 & Z_1 & 0 \\ 0 & 0 & Z_2 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} \tag{2.43}$$

2.2. One-Terminal Fault Location Methods

2.2.1. Impedance-Based Fault Location Methods and Requirements

Impedance-based methods require the following approach [8]:

1. Measure the voltage and current signals.
2. Extract the fundamental components.
3. Determine the fault type.
4. Apply impedance algorithm.

One-ended impedance methods of fault location are a standard feature in most numerical relays. One-ended impedance methods use a simple algorithm, and communication channels and remote data are not required (except when a channel is required to bring the fault location estimate to an operator).

One-ended impedance-based fault locators calculate the fault location from the apparent impedance seen by looking into the line from one end. To locate all fault types, the phase-to-ground voltages and currents in each phase must be measured. (If only line-to-line voltages are available, it is possible to locate phase-to-phase faults; if the zero-sequence source impedance, Z_0, is known, we can estimate the location for phase-to-ground faults).

If the fault resistance is assumed to be zero, we can use one of the impedance calculations in Table 2.1 to estimate the fault location.

Fault Type	Positive-Sequence Impedance Equation ($xZ_{1L} =$)
A-ground	$V_a/(I_a + 3kI_0)$
B-ground	$V_b/(I_b + 3kI_0)$
C-ground	$V_c/(I_c + 3kI_0)$
a-b or a-b-g	V_{ab}/I_{ab}
b-c or b-c-g	V_{bc}/I_{bc}
c-a or c-a-g	V_{ca}/I_{ca}
a-b-c	V_{ab}/I_{ab} or V_{bc}/I_{bc} or V_{ca}/I_{ca}

Table 2.1: Simple Impedance Equations

$k = (Z_{0L} - Z_{1L})/3Z_{1L}$,
Z_{0L} = Zero-sequence line impedance,
Z_{1L} = Positive-sequence line impedance,
x = Per unit distance to fault
I_0 = Zero-sequence current

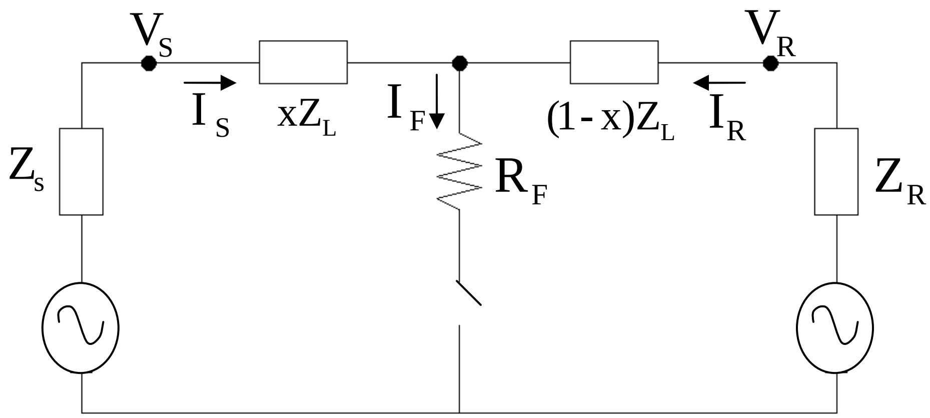


Figure 2.5: Circuit Representation of Line Fault

2.2.1.1. Reactance Method

Reactance Method works reasonably well for homogeneous systems when the fault does not involve significant resistance and load current [9]. Large errors are introduced to the fault location estimate by remote-end current feed, load impedance, power transmission angle, and different impedance angles of line and power system source impedances.

From Figure 2.6, the voltage drop from the S end of the line is:

$$Z_1 = Z_L$$

$$V_s = xZ_1 I_S + R_F I_F \tag{2.44}$$

For an A-phase to ground fault, $V_S = V_{a\text{-}g}$ and $I_s = I_a + 3kI_0$.

The goal is to minimize the effect of the $R_F I_F$ term.

The reactance method divides all terms by $\mathrm{I_S}$ (measured at the fault locator)

To do this, save the imaginary part, and solve for *x*,

$$\mathrm{Im}(V_S / I_S) = \mathrm{Im}(xZ_1) + \mathrm{Im}\left(\frac{R_f I_f}{I_S}\right)$$

If $\frac{I_f}{I_S}$ = Real Number,

$$\text{Im}(V_S / I_S) = \text{Im}(xZ_1)$$

$$x = \frac{\text{Im}\left(\frac{V_S}{I_S}\right)}{X_1} \tag{2.45}$$

$$\text{Error is 0 if } \angle I_S = \angle I_F \text{ or } R_F = 0$$

Although a reactance-based algorithm provides sufficient accuracy for many protective relaying applications, it is not accurate enough, in general, for use as a fault-locating algorithm.

2.2.1.2. Takagi Method

Takagi 1

Takagi Method is a method which uses prefault current as well as fault current and voltage [10]. The approach is to find a quantity which, when multiplied by the fault-voltage term, yields a purely real result. Then, when the imaginary components are selected and compared, an estimate of the distance to the fault is obtained, which is independent of the fault resistance.

The key is to find a quantity which is locally observable, and which is proportional to the complex conjugate of the fault voltage [11]. Consider conditions at the fault. The fault current is the sum of the currents from bus S and bus R. It may also be expressed as the sum of the superposition currents from buses S and R, which would flow if a superposition source of - V_{F0} is placed at the fault location, and if all other sources are set to zero. The superposition current component from bus S is the difference between the fault and prefault currents from bus S. That is,

$$\Delta \underline{I}_S = I_f - I_{pf} \tag{2.46}$$

where the subscripts *f* and *pf* denote fault and prefault conditions, respectively. $\Delta \underline{I}_s$ is a part of the fault current, we would expect $\Delta \underline{I}_s$ and I_F to be in close phase relationship, especially when most of the fault current comes from the S side of the system. I define the ratio of the fault current to the S-side superposition current component as:

$$A = \frac{I_F}{\Delta \underline{I}_S} \quad \text{and} \quad T = \angle A \tag{2.47}$$

When $\Delta \underline{I}_S$ is known, distance to a fault is calculated using [8]:

$$x = \frac{\text{Im}\left[V_S \Delta \underline{I}_S * \exp(-jT)\right]}{\text{Im}[Z_1 I_S \Delta \underline{I}_S * \exp(-jT)]} \tag{2.48}$$

Takagi then sets *T* equal to zero, giving the final result:

$$x = \frac{\text{Im}(V_S \Delta \underline{I}_S)}{\text{Im}(Z_1 I_S \Delta \underline{I}_S *)} \tag{2.49}$$

As will be seen, this method offers substantial improvements over Reactance Method, in most cases, by accounting for load flow.

The angle *T* can be determined from the circuit model. The angle *T* is zero if all impedances share the same angle. The angle *T* is constant if Z_1 and Z_R have equal angles. The accuracy of the Takagi algorithm is very dependent on the system conditions that affect the angle *T*. That is, the error is very sensitive to changes in *T*. Accuracy can be guaranteed only if it is known that the system is nearly homogeneous.

Takagi 2

Takagi Method 2 is a modified version of Takagi Method 1 [8]. It requires difference of positive sequence fault and pre-fault currents called the positive sequence superposition current.

$$\Delta \underline{I}_S = I_{f1} - I_{pf} \tag{2.50}$$

where the subscripts *f*1 and *pf* denote the positive sequence fault current and prefault conditions, respectively.
The distance to fault calculation is the same as Takagi 1.

Takagi 3

Takagi Method 3 uses zero-sequence current for ground faults instead of the superposition current [8]. Therefore, this method requires no pre-fault data.

$$x = \frac{\mathrm{Im}\{V_s I_{s0}^*\}}{\mathrm{Im}\{Z_1 I_s I_{s0}^*\}} \tag{2.51}$$

Takagi 3 also allows for angle correction. If the user knows the system source impedances, the zero-sequence current can be adjusted by angle *T* to improve the fault location estimate for a given line [8].

2.2.1.3. ABB Method

ABB Method implemented in ABB REL 5xx series of relays [12]. The distance-to-fault locator (DFL) in the REL 5xx line protection terminal is an essential complement to the distance protection function. The used calculation algorithm takes into consideration the effect of load currents, double end infeed and additional fault resistance.

The accuracy of the distance-to-fault measurement depends, to a certain extent, on the accuracy of the system parameters as entered into REL 5xx (e.g., source impedances at both ends of the protected line). If some parameters have actually changed in a significant manner relative to the set values, new values can be entered, locally or remotely, and a recalculation of the distance to the fault, using particular disturbance values, can be performed. This way a more accurate location of the fault can be determined and faster service can take place.

The influence of the zero sequence mutual impedance on the distance-to-fault calculation in case of faults on double circuit lines is compensated for by transferring the residual current from the healthy line to a separate input current transformer in the REL 5xx on a faulty line. This current is used only for purposes of a fault location function, but not for purposes of a distance protection function.

For transmission lines with voltage sources at both line ends, the effect of double end infeed and additional fault resistance must be taken into consideration when calculating the distance to the fault from the currents and voltages at one line end [13]. If this is not done, the accuracy of the calculated figure will vary with the load flow and the amount of additional fault resistance.

The calculation algorithm used in the distance-to-fault locator in REL 5xx line protection relay includes the effect of double-end infeed and additional fault resistance.

From Figure 2.5, the voltage drop from the S end of the line is:

$$V_s = xZ_L I_S + R_F I_F \tag{2.52}$$

The fault current is expressed in measurable quantities by:

$$I_F = \frac{I_{FS}}{D_S} \tag{2.53}$$

where:

I_{FS} is the change in current at the point of measurement, terminal S

D_S is a fault current distribution factor, i.e. the ratio between the fault current at line end S and the total fault current

For a single line,

$$D_S = \frac{(1-x)Z_L + Z_R}{Z_s + Z_L + Z_R} \tag{2.54}$$

In case of phase short circuits, the change in the line currents is used directly while for earth faults, the better defined positive sequence quantities of the network are used.

From the theory of symmetrical components for the single phase to the earth-faults it follows that:

$$I_F = 3I_0 \tag{2.55}$$

$$I_0 = I_{1S} + I_{1R} = I_{2S} + I_{2R} \tag{2.56}$$

$$I_{phase} = I_0 + I_1 + I_2 \tag{2.57}$$

Since I_F represents a change due to the short circuit, it follows that

$$I_{1S} + I_{1R} = \frac{I_{1S}}{D_S} \tag{2.58}$$

and

$$I_{2S} + I_{2R} = \frac{I_{2S}}{D_S} \tag{2.59}$$

If the positive and negative sequence impedances are equal, i.e.

$$I_F = 3I_0 = \frac{3}{2}\frac{I_{phase} - I_{0S}}{D_S} \tag{2.60}$$

The expressions for V_S, I_S and I_{FS} for different types of faults are given in table:

Fault Type	V_S	I_S	I_{FS}
A-ground	V_{AS}	$I_{AS} + K_N I_{SG}$	$\frac{3}{2} \times \Delta(I_{AS} - I_{0S})$
B-ground	V_{BS}	$I_{BS} + K_N I_{SG}$	$\frac{3}{2} \times \Delta(I_{BS} - I_{0S})$
C-ground	V_{CS}	$I_{CS} + K_N I_{SG}$	$\frac{3}{2} \times \Delta(I_{CS} - I_{0S})$
a-b-c or a-b or a-b-g	$V_{AS} - V_{BS}$	$I_{AS} - I_{BS}$	ΔI_{A-BS}
b-c or b-c-g	$V_{BS} - V_{CS}$	$I_{BS} - I_{CS}$	ΔI_{B-CS}
c-a or c-a-g	$V_{CS} - V_{AS}$	$I_{CS} - I_{AS}$	ΔI_{C-AS}

Table 2.2: Expressions for V_S, I_S and I_{FS}

Where the complex quantity K_N for zero sequence compensation is equal to:

$$K_N = \frac{Z_{0L} - Z_{1L}}{3Z_{1L}} \tag{2.61}$$

Δ is the change in current, i.e. the current after the fault less the current before the fault. In the following, the positive sequence impedance for Z_S, Z_R and Z_L is inserted into the equations, since this is the value used in the algorithm.

For double lines, the fault equation will be:

$$V_S = I_S \cdot xZ_{1L} + \frac{I_{FS}}{D_S} \cdot R_F + I_{0P} \cdot Z_{0M} \tag{2.62}$$

where:

I_{0P} is a zero sequence current of the parallel line.

Z_{0M} is a mutual zero sequence impedance

D_S is the distribution factor of the parallel line, which is

$$D_S = \frac{(1-x)\cdot(Z_{1S} + Z_L + Z_{1R}) + Z_{1R}}{2Z_{1s} + Z_{1L} + 2Z_{1R}} \tag{2.63}$$

The compensation factor K_N for the double line becomes:

$$K_N = \frac{Z_{0L} - Z_{1L}}{3Z_{1L}} + \frac{Z_{0M}}{3Z_{1L}} \cdot \frac{I_{0P}}{I_{0S}} \tag{2.64}$$

From these equations, it can be seen that if $Z_{0M} = 0$, the general fault location equation for single line is obtained. Only the distribution factor differs in these two cases.

Since the distribution factor D_S according to equation (2.54) or (2.63) is a function of *x*, the general equation (2.62) can be written in the form:

$$x^2 - xK_1 + K_2 - K_3 R_F = 0$$

where:

$$K_1 = \frac{V_S}{I_S Z_{1L}} + \frac{Z_{1R}}{Z_{1L} + Z_{ADD}} + 1 \tag{2.65}$$

$$K_2 = \frac{V_S}{I_S Z_{1L}} \cdot \left(\frac{Z_{1R}}{Z_{1L} + Z_{ADD}} + 1 \right) \tag{2.66}$$

$$K_3 = \frac{I_{FS}}{I_S Z_{1L}} \cdot \left(\frac{Z_{1S} + Z_{1R}}{Z_{1L} + Z_{ADD}} + 1 \right) \tag{2.67}$$

and:

$Z_{ADD} = Z_{1S} + Z_{1R}$ for parallel lines

I_S, I_{FS} and V_S are given in the above table

K_N is calculated automatically according to equation (2.2.1.3.13)

$Z_{1S}, Z_{1R}, Z_{1L}, Z_{0L}$ and Z_{0M} are setting parameters

For a single line, $Z_{0M} = 0$ and $Z_{ADD} = 0$. Hence, equation (2.64) is applicable to both single and parallel lines.

Equation (2.64) can be divided into a real part and an imaginary one:

$$x^2 - x \cdot \mathrm{Re}(K_1) + \mathrm{Re}(K_2) - R_F \cdot \mathrm{Re}(K_3) = 0 \tag{2.68}$$

$$-x \cdot \mathrm{Im}(K_1) + \mathrm{Im}(K_2) - R_F \cdot \mathrm{Im}(K_3) = 0 \tag{2.69}$$

If the imaginary part of K_3 is not zero or close to zero, R_F is solved according to equation (2.69), and inserted into equation (2.68). According to equation (2.68), the relative distance to the fault is solved as the root of a quadratic equation.

Equation (2.69) gives as a solution two different values for the relative distance to the fault. A simplified load compensated algorithm that gives an unequivocal figure for the relative distance to the fault, is used to establish the value that should be selected. If the load compensated algorithms according to the above do not give a reliable solution, a less accurate, non-compensated impedance model is used to calculate the relative distance to the fault.

The accuracy of the distance-to-fault calculation, using the non-compensated impedance model, is influenced by the pre-fault load current. Therefore, this method is used only if the load compensated models do not function and it indicates whether the non-compensated model was used when calculating the distance to the fault.

2.3. Two-Terminal Fault Location Method

Two-terminal methods can be more accurate than one-terminal method but data must be captured from both ends before an algorithm can be applied [14]. Many of the existing algorithms require the transfer of large amounts of data, alignment of the data sets, and iterative solutions to calculate the distance to the fault point. This makes their application limited to processing the data offline and adds considerable amount of time in the fault location process. In addition, some of the existing two-terminal methods cannot adequately handle mutual coupling and tapped loads with zero-sequence current infeeds, and are not applicable to more than two- terminal lines.

2.3.1. Two-Ended Negative-Sequence Impedance Method

It uses negative-sequence quantities from all line terminals for the location of unbalanced faults [15]. By using negative-sequence quantities, we negate the effect of prefault load and fault resistance, zero-sequence mutual impedance, and zero-sequence infeed from transmission line taps. Precise fault type selection is not necessary. Data alignment is not required because the algorithm employed at each line end uses the following quantities from the remote terminal (which do not require phase alignment).

- Magnitude of negative-sequence current, I_2
- Calculated negative-sequence source impedance, $Z_2\angle\theta_2°$

An observation from figure 2.6 below is that the negative-sequence fault voltage (V_{2F}) is the same when viewed from all ends of the protected line.

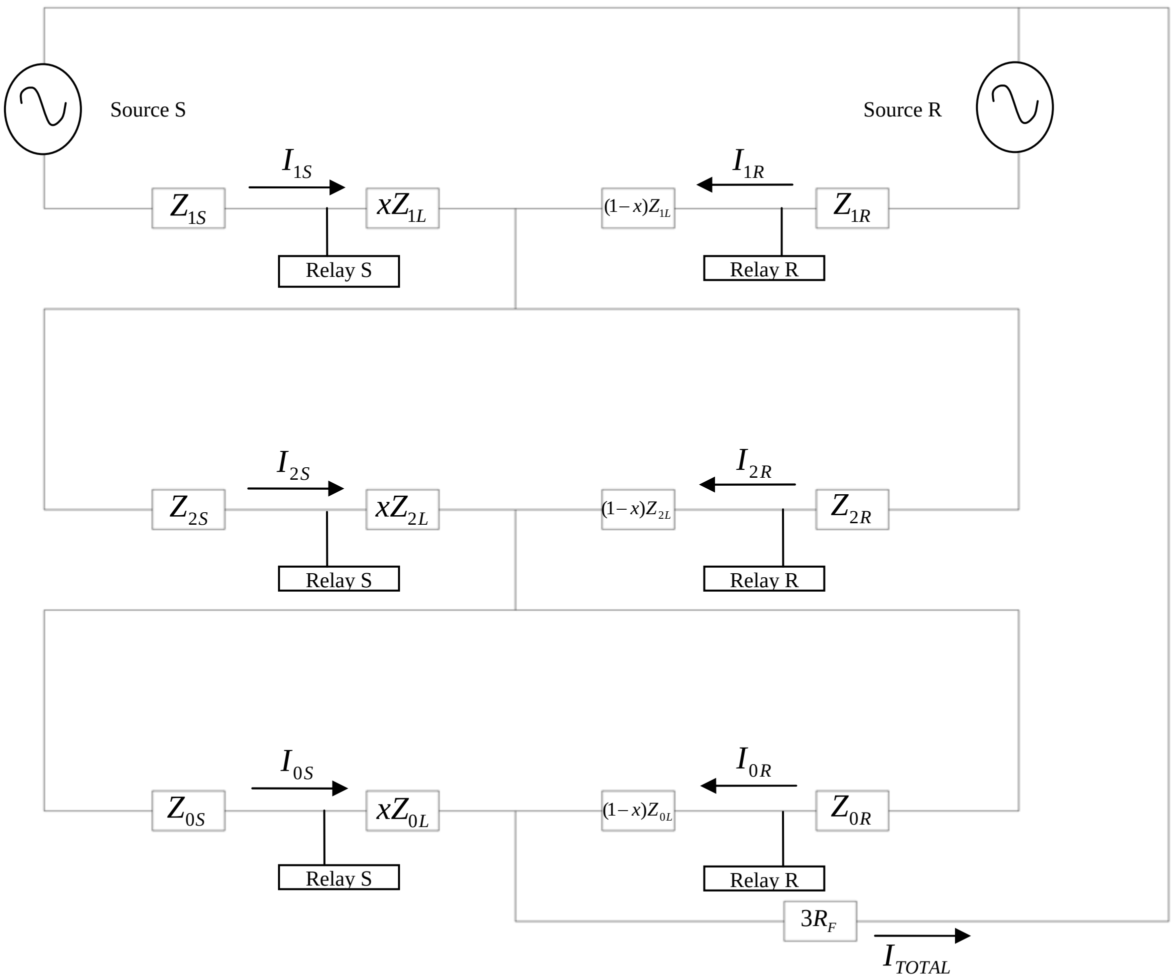


Figure 2.6: Connection of Sequence Networks for a Single Line-to-Ground Fault at x

At Relay S:

$$V_{2F} = -I_{2S}(Z_{2S} + xZ_{2L}) \tag{2.69}$$

At Relay R:

$$V_{2F} = -I_{2R}(Z_{2R} + (1-x)Z_{2L}) \tag{2.70}$$

Eliminate V_{2F} from Equation 8 and 9 and rearrange the resulting expression as follows:

$$I_{2R} = I_{2S}\frac{(Z_{2S} + xZ_{2L})}{(Z_{2R} + (1-x)Z_{2L})} \tag{2.71}$$

To avoid alignment of Relay S and R data sets, take the magnitude of both sides of Equation above as follows:

$$|I_{2R}| = \left| I_{2S} \frac{(Z_{2S} + xZ_{2L})}{(Z_{2R} + (1-x)Z_{2L})} \right| \tag{2.72}$$

Equation above is then simplified to Equation below.

$$|I_{2R}| = \left| \frac{(I_{2S} Z_{2S}) + x(I_{2S} xZ_{2L})}{(Z_{2R} + Z_{2L}) - xZ_{2L}} \right| \tag{2.73}$$

To further simplify Equation above, define the following variables:

$$I_{2S} Z_{2S} = a + jb \tag{2.74}$$

$$I_{2S} xZ_{2L} = c + jd \tag{2.75}$$

$$Z_{2R} + Z_{2L} = e + jf \tag{2.76}$$

$$Z_{2L} = g + jh \tag{2.77}$$

Substituting these variables into Equation above, we obtained:

$$|I_{2R}| = \left| \frac{(a + jb) + x(c + jd)}{(e + jf) - x(g + jh)} \right| \tag{2.78}$$

Taking the square of both terms of Equation above, expanding and rearranging terms results in a quadratic equation of the form:

$$A \cdot x^2 + B \cdot x + C = 0 \tag{2.79}$$

Equation above is solved for *x*. The coefficients of Equation above are given below.

$$A = |I_{2R}|^2 \cdot (g^2 + h^2) - (c^2 + d^2) \tag{2.80}$$

$$B = -2|I_{2R}|^2 \cdot (e \cdot g + f \cdot h) - 2(a \cdot c + b \cdot d) \tag{2.81}$$

$$C = |I_{2R}|^2 \cdot (e^2 + f^2) - (a^2 + b^2) \tag{2.82}$$

Chapter 3.

Uncertainty & Sensitivity Analysis

3.1. Introduction

Sensitivity analysis (SA) is the study of how the variation (uncertainty) in the output of a mathematical model can be apportioned, qualitatively or quantitatively, to different sources of variation in the input of a model [16]. In more general terms uncertainty and sensitivity analyses investigate the robustness of the fault location algorithms. While uncertainty analysis studies the overall uncertainty in the results of the algorithms, sensitivity analysis tries to identify which uncertain factors weights more on the fault location results. In sensitivity analysis we look at the effect of varying the uncertain factors of the algorithms on the fault location results. In both disciplines we strive to obtain information from the system with a minimum of physical or numerical experiments. In uncertainty and sensitivity analysis there is a crucial trade off between how scrupulous an analyst is in exploring the input assumptions and how wide the resulting inference may be [17].

The understanding of how the algorithm behaves in response to changes in its uncertain factors is of fundamental importance to ensure a correct use of the algorithm. In our research, the fault location algorithm is defined by a series of equations, uncertain factors, parameters, and variables aimed to characterize the process of finding fault location on the line.

The uncertain factors are subject to many sources of uncertainty including errors of measurement, absence of information and poor or partial understanding of the driving forces and mechanisms [21]. This uncertainty imposes a limit on our confidence in the response or the output of the algorithm. Uncertainty and sensitivity analysis offer valid tools for characterizing the uncertainty associated with the fault location algorithm. Uncertainty analysis quantifies the uncertainty in the outcome of our algorithm. Sensitivity analysis has the complementary role of ordering by importance the strength and relevance of the uncertain factors in determining the variation in the fault location results. In the fault location algorithm that involves many uncertain variables, sensitivity analysis is an essential ingredient of algorithm building and quality assurance.

There are several possible procedures to perform uncertainty and sensitivity analysis. The most common sensitivity analysis is sampling-based [17]. A sampling-based computation of sensitivity is one in which the algorithm is executed repeatedly for combinations of values sampled from the distribution (assumed known) of the uncertain factors. Sampling based methods can also be used to decompose the variance of the output of the algorithm [21]. From the variance we are able to find out the most contributing uncertain factor and also the contribution of the uncertain factors' interactions. The steps to perform sampling-based sensitivity analysis can be listed below [17]:

1. Specify the target function which in our case is the fault location algorithm and select the uncertain factors of interest.
2. Assign a probability density function to the selected factors.
3. Generate a matrix of the uncertain factors with that distribution(s) through an appropriate design. In this research, we have used quasi-random sequence to generate the samples of all the uncertain factors. In the next chapter, we review the quasi-random sequence and also the advantages of using it.
4. Evaluate and compute the distribution of the fault location algorithm outputs.

5. Select a method for assessing the influence or relative importance of each uncertain factor on the output of fault location algorithm. We have chosen ANOVA decomposition as the method which will be discussed in the next chapter.

3.2. Global Sensitivity Analysis

In this research, we have used global sensitivity analysis because we want to analyze the whole set of influential factors and give us an overall indication of the way that the fault location results varies. Most SA methods found in the literature and in practical applications are local SA. A global sensitivity analysis can display a non-linear response, and thus it can be incorrect to globally extrapolate the local results [18]. The analysis requires a quantitative assessment of the uncertainty around some best estimate value for Y (uncertainty analysis). Monte Carlo methods and variety of sampling strategies are implemented in global sensitivity analysis [19]. Furthermore, regression techniques are introduced. In this case, standardized regression coefficients (SRC) are used [17]. The regression algorithm with model input and output values are implemented in regression techniques. The output of the algorithm is a regression meta-model and the regression coefficient provides information about the sensitivity measure for the model factors. The sign of a regression coefficient is a measure of the effect of input factors on Y (output). Furthermore, the absolute value of the regression coefficient is used to order the factors by importance [17].

When discussing sensitivity with respect to factors, we shall interpret the term "factor" in a very broad sense [17]: "a factor is anything that can be changed prior to the execution of the model, possibly from a prior or posterior, continuous or discrete distribution." Factors can be "triggers", used to select one versus another model structure, one mesh size versus another or altogether different conceptualizations of the system. The input factors space is introduced to analyze nonlinear model with global sensitivity analysis.

3.3. Variance-based Sensitivity Measures

3.3.1. Properties of the variance based methods

An ideal sensitivity analysis method should cope with the influence of scale and shape. The influence of the input should incorporate the effect of the range of input variation and the form of its probability density function (pdf). It matters whether the pdf of an input factor is uniform or normal, and what are the distribution parameters. Variance based methods meet this demand [16].

A good method should allow for multidimensional averaging, contrary, for example, to what is done in computing partial derivatives, where the effect of the variation of a factor is taken when all others are kept constant at the central (nominal) value [17]. A sensitivity measure should be model independent. The method should work regardless of the additivity or linearity of the test model. A global sensitivity measure must be able to appreciate interaction effect, especially important for non-linear, non-additive models [21, 22]. The property is evident with the variance based measures.

An ideal measure should be able to treat grouped factors as if they were single factors. This property of synthesis is useful for the agility of the interpretation of the results. One would not want to be confronted with an SA made of dense tables of input-output sensitivity indices. Variance based methods are capable of grouping the factors.

In this research, we measure the uncertainty of the fault location using variance (precision error measure). The fault location variance is contributed by a total series of uncertain factors that we select. For example, the importance measure of a factor x_1 is equal to part of the fault location variance that is contributed by uncertainty of the factor x_1.

3.4. Introduction of ANOVA

Analysis of Variance (ANOVA) decomposition has been chosen as a tool for global sensitivity analysis of the fault location algorithm [23, 26]. The ANOVA for square integrable function is becoming a widely used tool for the exploratory analysis of functions. $[0,1]^d$ is the d-dimensional hypercube where the function is defined. The ANOVA allows us to quantify the notion that some variables and interactions are much more important than others. The result is a form of global sensitivity analysis, distinct from local methods based on partial derivatives. Within the ANOVA formulation, we may answer questions about variable importance. The ANOVA of a function in the domain $[0,1]^d$ involves $2^d - 1$ effect [28]. For moderately large d it becomes difficult to estimate them all.

The ANOVA decomposition is a representation of the form

$$f(x) = \sum_u f_u(x) \qquad (3.1)$$

where the sum is over subsets $u \subseteq \{1,2,...,d\}$ and $f_u(x)$ is a function on $[0,1]^d$ that depends on x only through x^j where $j \in u$. $f(x)$ represents the function of a fault location algorithm with input factors x. In the application, we are interested to represent (3.1) equation as the following expansion.

$$f(x_1,...,x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i) + \sum_{i=1}^{d} \sum_{j=i+1}^{d} f_{ij}(x_i, x_j) + ... + f_{1,2,...,d}(x_{1,...,}x_d) \qquad (3.2)$$

To make the decomposition unique, f_0 must be a constant and the line integrals of every summand in (3.2) over any of its own variables must be equal to zero. In addition all the summands are orthogonal. For example, consider the function $f(x) = x_1 x_2 x_3$ on $[0,1]^3$. It can be written

$$\begin{aligned} f(x) = {} & \frac{1}{8} + \frac{1}{4}\left(x_1 - \frac{1}{2}\right) + \frac{1}{4}\left(x_2 - \frac{1}{2}\right) + \frac{1}{4}\left(x_3 - \frac{1}{2}\right) \\ & + \frac{1}{2}\left(x_1 - \frac{1}{2}\right)\left(x_2 - \frac{1}{2}\right) + \frac{1}{2}\left(x_1 - \frac{1}{2}\right)\left(x_3 - \frac{1}{2}\right) + \frac{1}{2}\left(x_2 - \frac{1}{2}\right)\left(x_3 - \frac{1}{2}\right) \\ & + \left(x_1 - \frac{1}{2}\right)\left(x_2 - \frac{1}{2}\right)\left(x_3 - \frac{1}{2}\right) \end{aligned} \qquad (3.3)$$

There are 8 functions in (3.3), one for each subset of {1,2,3}. Each of these functions integrates to zero over the range of any of the variables in it.

Since we are dealing with multi-dimensional problems, special care must be taken in selecting the computational method to perform ANOVA decomposition. Computing ANOVA decomposition is a sampling based process in which the power system simulator and a fault locator are both executed repeatedly for combinations of values sampled from the assumed distribution of the input factors. The main goal here is to achieve desired accuracy in computing the ANOVA model with a minimal number of samples and corresponding simulator and locator executions. The obvious practical benefit is reduction of the time required to perform the analysis.

3.5. Sensitivity and Variable Importance

The variance of $f(x)$ is

$$\sigma^2(f(x)) = \sum_{u \neq 0} \sigma_u^2 \tag{3.4}$$

Where σ_u^2 is the variance of $f_u(x)$.

Sensitivity measure for the factor x_1

First take the average over all factors except x_1 as follows:

$$E[f(x) \mid x_1], \tag{3.5}$$

and then calculate variance over x_1

$$\sigma_1^2 = \sigma_{x_1}^2(E[f(x) \mid x_1]). \tag{3.6}$$

Sensitivity measure for x_1 is

$$S_{x_1} = \frac{\sigma_1^2}{\sigma^2(f(x))} \tag{3.7}$$

A sum of variances σ_i (individual effects) smaller than the total variance means that additional effects resulting from interactions between the influencing factors exists.

For the ANOVA decomposition written in form (3.2), the variance of $f(x)$ has a very useful structure (equivalent to (3.4)):

$$\sigma^2(f(x)) = \sum_{i=1}^{d} \sigma_i^{\ 2} + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \sigma_{ij}^{\ 2} + \ldots + \sigma_{1,2,\ldots,d}^{\ \ 2} \qquad (3.8)$$

The first decomposition expression, $\sum_{i=1}^{d} \sigma_i^{\ 2}$ represents the total variance of all the input factors. The second expression, $\sum_{i=1}^{d} \sum_{j=i+1}^{d} \sigma_{ij}^{\ 2}$ represents the total variance of all interaction between two input factors. The last expression represents the total variance of all interaction between all input factors.

The sensitivity measure of the main effects and interactions are defined as the ratio below:

One factor effect:

$$S_{x_i} = \frac{\sigma_i^2}{\sigma^2(f(x))} \qquad (3.9)$$

Two factors interaction:

$$S_{x_i x_j} = \frac{\sigma_{ij}^2}{\sigma^2(f(x))} \qquad (3.10)$$

Three factors interaction:

$$S_{x_i x_j x_k} = \frac{\sigma_{ijk}^2}{\sigma^2(f(x))} \quad \ldots \qquad (3.11)$$

And etc.

Chapter 4.

Sensitivity Analysis using Quasi Regression

4.1. Monte Carlo

Monte Carlo (MC) methods are the most widely used means for uncertainty analysis [34]. These methods involve random sampling from the distribution of inputs and successive model runs until a statistically significant distribution of outputs is obtained [35]. They can be used to solve problems with physical probabilistic structures, such as uncertainty propagation in models or solution of stochastic equations, or can be used to solve non-probabilistic problems, such as finding the area under a curve. Monte Carlo methods are also used in the solution of problems that can be modeled by the sequence of a set of random steps that eventually converge to a desired solution. Since these methods require a large number of samples (or model runs), their applicability is sometimes limited to simple models. In case of computationally intensive models, the time and resources required by these methods could be prohibitively expensive. In our research, Monte Carlo method is used in quasi-regression to estimate coefficients.

The Monte Carlo method normally solves the problem with large dimension *d*. There are two types of methods: Monte-Carlo and Quasi Monte-Carlo [36]. In Monte-Carlo methods, the points are chosen randomly. The procedure converges almost surely and there is a probabilistic error bound [39]. In Quasi Monte-Carlo methods, the points come from deterministic multidimensional sequences with very low irregularities of distribution. In particular it is common for the Quasi Monte-Carlo method to produce much more accurate answers than the Monte Carlo method does in a short period of time [35]. These two methods mainly have the same working techniques; the only difference is the sampling method. Quasi-random sequences are more evenly scattered throughout the region. Quasi Monte-Carlo can improve the accuracy of solving the integral. In order to see this, let the integral of one function $f(x)$ being evaluated using simulation. The idea is to use random points for the numerical evaluation of an integral. This is equivalent to determining the area under the function, see Figure 4.1 below.

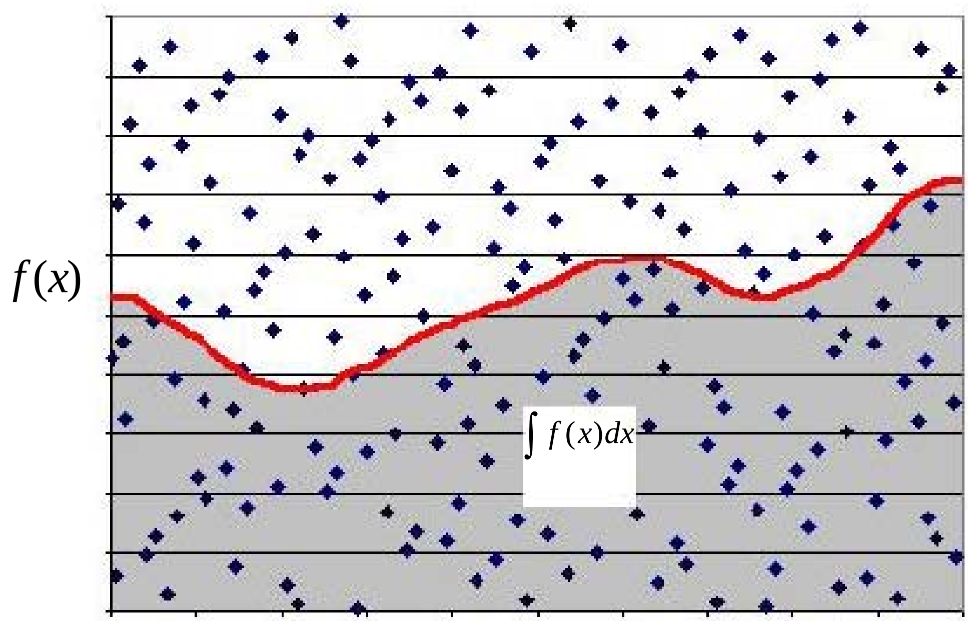


Figure 4.1: The Monte Carlo Integral

The integral of the function f(x) is approximately the total area times the fraction of points that fall under the curve of f(x). Naturally this method for evaluation of an integral (using "random" points) is competitive only for the multi-dimensional case and/or complicated functions. Note that the integral evaluation is better if the points are uniformly scattered in the entire area or, for the multi-dimensional case, in the *hypercube volume*.

An example is shown below.
If we wish to evaluate the integral,

$$I = \int_b^a f(x)dx \qquad (4.1)$$

We put a bounding box around the function $f(x)$, then the integral of $f(x)$ can be understood to be the fraction of the bounding box that is also within $f(x)$. So if we choose a point at random uniformly within the bounding box, the probability that the point is within $f(x)$ is given by the fraction of the area that $f(x)$ occupies. The integration scheme is then to take a large number of random points within the box and count the numbers that are within $f(x)$ to get the area.

$$I \approx \frac{n^*}{n} V \qquad (4.2)$$

Where, n^* is the number of points within $f(x)$, n is the number of points generated, and V is the volume of the bounding box.
A more efficient approach,

$$I = \int_a^b g(x) f(x) dx = \frac{1}{V} \int_a^b g(x) f(x) V dx \qquad (4.3)$$

Where
$g(x) = 1$ If x is in the domain
$g(x) = 0$ Otherwise

4.2. Quasi-Monte Carlo

The Quasi-Monte Carlo (QMC) method is used for this research, as this method is faster and more accurate compared to the Monte Carlo method. The quasi-Monte Carlo method is also called the *low discrepancy sequences method* [34]. In order to improve the convergence rate for the integral, one looks for quasi-random sequences with the lowest possible discrepancy, as for example Halton sequences, Faure sequences, Sobol' sequences etc [34]. The Sobol' sequences have been chosen for our research as the Sobol has more advantages than the others. It distributes the points uniformly as $N \to \infty$. It has good distribution for fairly small initial sets. In addition, the Sobol is a very fast computational algorithm [39].

4.3. Sobol Sequence

The Sobol sequence has the same base for all dimensions and proceeds a reordering of the vector elements within each dimension [33]. The Sobol sequence is simpler (and faster) than the Faure sequence in the aspect that Sobol sequence uses base 2 for all dimensions. So, there is some computational time advantage due the shorter cycle length.

However the simplicity of Sobol sequence compared with the Faure sequence, ends at this point because the reordering task is more complex. Sobol reordering is based on a set of "direction numbers", $\{V_i\}$. The V_i numbers are given by the equation $V_i = \frac{m_i}{2^i}$ where the m_i are odd positive integers less than 2^i, and V_i are chosen so that they satisfy a recurrence relation using the coefficients of a primitive polynomial in the Galois filed [33].

In other words, Sobol sequence uses the coefficients of irreducible primitive polynomials of modulo 2 for its complex reordering algorithm.

The points plotted below are the first 100 and 1000 elements in the Sobol sequence.

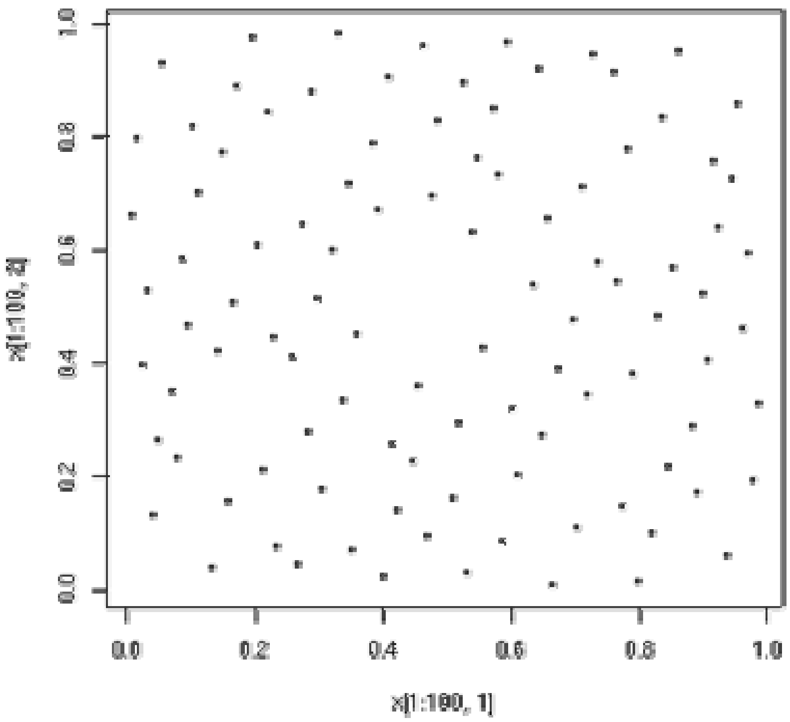


Figure 4.2: The first 100 points in a low-discrepancy sequence of the Sobol’ type

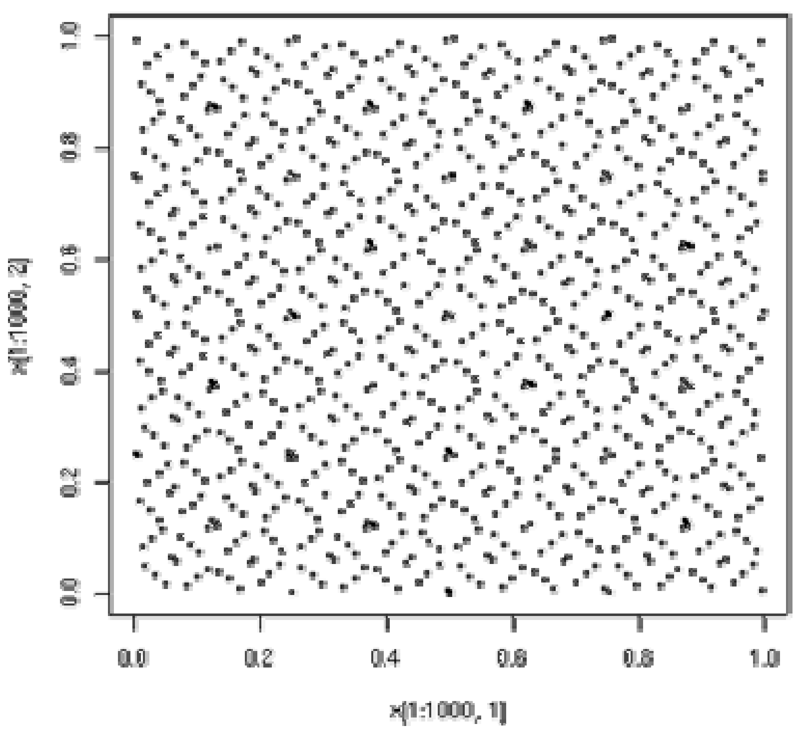


Figure 4.3: The first 1000 points in a low-discrepancy sequence of the Sobol’ type

4.4. Monte Carlo vs. Quasi-Monte Carlo Integration. Test Function.

Integration of a torus in three-dimensional space has been achieved using the Monte Carlo and Quasi Monte Carlo.

The following is the function,

$$f(x, y, z) = 1 + \cos(\frac{\pi r^2}{a^2}) \qquad r \le r_0$$

$$f(x, y, z) = 0 \qquad r > r_0$$

Where

$$r^2 = (\sqrt{x^2 + y^2} - R_0)^2 + z^2$$

R_0 is the major radius of the torus. r is the minor radial.

With the parameter $R_0 = 0.6$, $r_0 = 0.3$, $a = 0.3$

The integration result is $I = \pi^2 0.6^3 \approx 1.066$

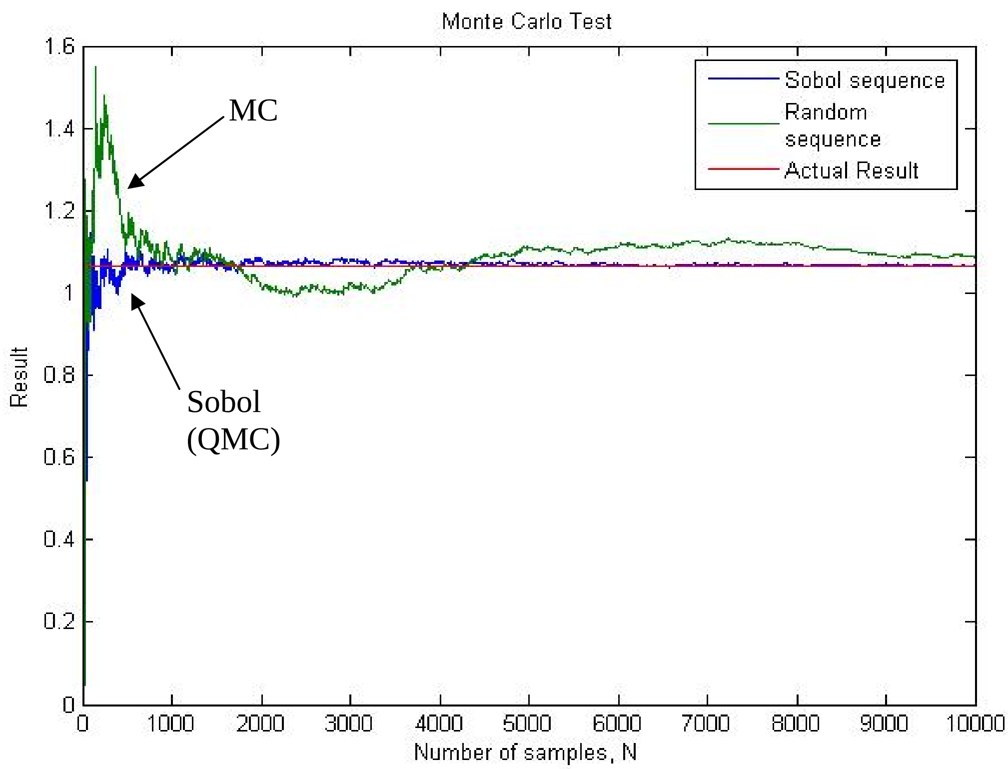


Figure 4.4: Monte Carlo vs. Quasi-Monte Carlo

Figure 4.4 compares the result of Monte Carlo, quasi-Monte Carlo and the actual result as a function of number of samples. From the figure, we can observe that speed of convergence favor quasi-Monte Carlo over Monte Carlo. Quasi-Monte Carlo has superior asymptotic accuracy compared to Monte Carlo. Hence, quasi-Monte Carlo is preferable for our research.

4.5. Quasi-Regression

Quasi regression is a method that is used to approximate a function on the hypercube in *d* (*number of variables in the particular function*) dimensions [28]. It is very useful for those methods that require a large number of function evaluations. In practice, we want to apply sensitivity analysis on the fault locator which consists of unknown complex computational models. Hence, we have chosen the quasi-regression in the research because it works well for black box functions. In the quasi-regression method, we approximate an unknown function (f) with another function (f^). Fast computation time and ANOVA decomposition are the advantages of quasi-regression.

From the definition of quasi-regression, we know that all integrals are done on $[0\ 1]^d$ interval (unit cube). Thus, all the basis functions used in the model are understood to be over the unit cube. All the univariate basis functions are low order orthogonal polynomials. For the function with *d* dimensions, its approximate function is equal to tensor products of univariate basis functions.

Jiang, T. and A.B.Owen [29] state that the approximate function f^ will be expanded in an orthonormal basis form with an infinite number of coefficients. If we knew these coefficients we could use them to analyze which uncertain factor contributes the most to the fault location results and also the interactions between the uncertain factors. In the article [30], the authors stated that we can approximate the coefficients using Monte Carlo methods. I am following this approach in my development.

4.5.1. One dimension

Consider a sequence $\{\phi_m(x) : m \geq 0\} = \phi_0(x), \phi_1(x), \phi_2(x), ...$ of functions defined over [0,1] and possessing the following properties:

- Orthonormal:

$$\int_0^1 \phi_m(x)\phi_l(x)dx = \delta_{ml}, \forall_m \quad (4.4)$$

- And contain a special element

$$\phi_0(x) = 1, \forall_x \in [0,1]. \quad (4.5)$$

where δ_{ml} is Kronecker symbol:

$$\delta_{ml} = 1, \text{ if } m = l, \quad (4.6)$$

$$\delta_{ml} = 0, \text{if } m \neq l, \quad (4.7)$$

Note, that from properties (4.4) and (4.5) follows that functions $\{\phi_m(x)\}$ are mean centered for positive indices:

$$\int_0^1 \phi_m(x)dx = 0, \forall_m \in \mathrm{N} \quad (4.8)$$

Univariate orthonormal basis functions, meeting above requirements can be constructed, for example, from orthogonal polynomials, trigonometric functions etc.

In our work we make use of basis function constructed from Legendre polynomials by appropriate rescaling of its argument (remind that Legendre polynomials are defined over the interval [-1,1]) and after that appropriate normalization. Omitting an intermediate derivation, let us give the final result. Thus, the orthonormal polynomials can be obtained with the help of the following recurrent formulae:

$$\phi_m(x) = \frac{\sqrt{2m+1}}{m}\left[\sqrt{2m-1}(2x-1)\phi_{m-1}(x) - \frac{(m-1)}{\sqrt{2m-3}}\phi_{m-2}(x)\right], \quad (4.9)$$

$$\phi_1(x) = \sqrt{3}(2x-1) \ , \ \phi_0(x) = 1, \quad (4.10)$$

Introducing for convenience an auxiliary variable $z = 2x - 1$ we can rewrite above expressions

$$\phi_m(z) = \frac{\sqrt{2m+1}}{m}\left[\sqrt{2m-1}z\phi_{m-1}(z) - \frac{(m-1)}{\sqrt{2m-3}}\phi_{m-2}(z)\right], \tag{4.11}$$

$$\phi_0(z) = 1, \quad \phi_1(z) = \sqrt{3}z \tag{4.12}$$

The explicit expressions for first five terms of the basis sequence are

$$\phi_0(z) = 1, \tag{4.13}$$

$$\phi_1(z) = \sqrt{3}z, \tag{4.14}$$

$$\phi_2(x) = \frac{\sqrt{5}}{2}(3z^2 - 1), \tag{4.15}$$

$$\phi_3(x) = \frac{\sqrt{7}}{2}z(5z^2 - 3), \tag{4.16}$$

$$\phi_4(x) = \frac{3}{8}(z^2(35z^2 - 30) + 3) \tag{4.17}$$

<u>Link with Monte Carlo integration</u>

The orthonormal basis $\{\phi_m(x)\}$ allows representing an arbitrary function in the following form

$$F(x) = \sum_{m=0}^{\infty} \beta_m \phi_m(x), \tag{4.18}$$

where coefficients β_m can be found from

$$\beta_m = \int_0^1 F(x)\phi_m(x)dx \tag{4.19}$$

Hence the problem of calculating the coefficients can be solved using numerical integration, which can be solved using among others standard Monte Carlo techniques (Monte Carlo and quasi-Monte Carlo integration). In practice, when the dimensionality of the problem is large it is difficult to find a competitor to Monte Carlo technique.

4.5.2. Multi-dimension

Let F be a real valued function defined over a unit hypercube I^d. We consider a problem of approximation of this function by a finite set of orthonormal functions.

4.5.2.1. Tensor product basis

We construct our basis function over I by taking tensor products univariate basis function. Let $r = (r_1, r_2, ..., r_d)$ be a vector of d non-negative integers, and denote d components of u by x^i for $i = 1,...,d$. To each such index vector r there corresponds a unique tensor product function

$$\psi_r(x) = \prod_{i=1}^{d} \phi_{ri}(x^i). \tag{4.20}$$

It is easy to prove that $\psi_{0,0,...,0}(x) = 1$ and

$$\int_I \psi_r(u)\psi_s(u)du = \delta_{rs} \tag{4.21}$$

Therefore, $\{\psi_r(x)\}$ form an orthonormal basis on I. In term of this basis we can write

$$F(x) = \sum_{r \in \bigcup} \beta_r \psi_r(x) \tag{4.22}$$

Where the sum is assumed over the whole infinite set. The coefficients of this expansion β_r are defined by

$$\beta_r = \int_I F(x)\psi_r(x)dx \tag{4.23}$$

4.5.2.2. Approximation by a finite set of basis functions

In practice we can estimate only finitely many coefficients. We truncate the infinite set to a finite set [28] $\mathrm{R} = \mathrm{R}_{B0,B1,B\infty} = \{\mathrm{r} : \|\mathrm{r}\|_0 \le B_0, \|\mathrm{r}\|_1 \le B_1, \|\mathrm{r}\|_\infty \le B_\infty\}, \mathrm{R}_{B_0,B_1,B_\infty}$, for problem-dependent values B_0, B_1 and B_∞. Here $\|\mathrm{r}\|_0, \|\mathrm{r}\|_1$ and $\|\mathrm{r}\|_\infty$ are different ways to measure the size of r defined as

$$\|\mathrm{r}\|_0 = \sum_{i=1}^{d} 1_{r_i > 0},\ \|\mathrm{r}\|_1 = \sum_{i=1}^{d} r_i \text{ and } \|\mathrm{r}\|_\infty = \max_{1 \le i \le d} r_i \tag{4.24}$$

and called rank, degree and order, respectively [28].

$F(x)$ can be written in the form:

$$F(x) = \sum_{\mathrm{r} \in \bigcup} \beta_\mathrm{r} \psi_\mathrm{r}(x) = \sum_{\mathrm{r} \in \mathrm{R}} \beta_\mathrm{r} \psi_\mathrm{r}(x) + \varepsilon(x), \tag{4.25}$$

where $\varepsilon(x)$ is deterministic truncation error:

$$\varepsilon(x) = \sum_{\mathrm{r} \notin \mathrm{R}} \beta_\mathrm{r} \psi_\mathrm{r}(x). \tag{4.26}$$

The above expression can be written in matrix notations. Assume that that the number of basis functions used in approximation is p. In mathematical language $p = \mathrm{Card}[\mathrm{R}_{B_0, B_1, B_\infty}]$ is the cardinality of the set $\mathrm{R}_{B_0, B_1, B_\infty}$. The finite set $\mathrm{R}_{B_0, B_1, B_\infty}$ can be mapped to a finite set Q_p of univariate indices $j = 1, \ldots, p$ with $\mathrm{Card}[\mathrm{Q}_p] = \mathrm{Card}[\mathrm{R}_{B_0, B_1, B_\infty}] = p$. Let

$$\Psi(\mathrm{x}) = \begin{pmatrix} \psi_1(\mathrm{x}) \\ \psi_2(\mathrm{x}) \\ \vdots \\ \psi_p(\mathrm{x}) \end{pmatrix} \text{ and } \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} \tag{4.27}$$

be a column vector of multivariate base functions and be a column vector of coefficients correspondingly. Then it can be written as,

$$F(x) = \Psi^T(\mathrm{x})\beta + \varepsilon(x), \tag{4.28}$$

where $\Psi^T(\mathrm{x})$ is the row vector.

The integrated square error (ISE),

$$\mathrm{ISE} = \int (F(x) - \Psi^T(x)\beta)^2 \, dx \tag{4.29}$$

qualifies error due to truncation of the infinite set $\bigcup$ to a finite set $\mathrm{R}_{B_0, B_1, B_\infty}$. It can be shown that the optimal β leading to the minimal ISE is,

$$\beta^* = \arg\min_\beta \int (F(x) - \Psi^T(x)\beta)^2 \, dx = \int F(x)\Psi(x) dx \tag{4.30}$$

Let

$$\hat{\beta}_{\mathrm{r}}^{(n)} = \frac{1}{n}\sum_{k=1}^{n} F(x_k)\psi_r(x_k), \quad r \in \mathrm{R}_{B_0,B_1,B_\infty} \tag{4.31}$$

be the quasi-regression estimate of β_{r} based on samples $(\mathrm{u}_1,\ldots,\mathrm{u}_n)$.

It can be rewritten in vector form as follows

$$\hat{\beta}^{(n)} = \frac{1}{n}\sum_{k=1}^{n} F(x_k)\psi(x_k). \tag{4.32}$$

The approximating function

$$\hat{F}^{(n)}(x) = \sum_{\mathrm{r}\in\mathrm{R}} \hat{\beta}_{\mathrm{r}}^{(n)}\psi_r(x) = \psi^T(x)\hat{\beta}^{(n)} \tag{4.33}$$

We mention that the coefficient $\beta_{\mathbf{r}}$ is equal to the expected value of $\mathrm{f}(x)\psi_r(x)$ where $\boldsymbol{x} \sim \mathrm{U}(0, 1)^d$. A.B.Owen [28] presents quasi-Monte Carlo sampling to estimate β_{r}. He defines

$$\tilde{\beta}_{\boldsymbol{r},n} = \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{x}_i)\psi_{\boldsymbol{r}}(\boldsymbol{x}_i) \tag{4.34}$$

$$\mathrm{SV}_{r,\mathrm{n}} = \frac{1}{\mathrm{n}-1}\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{f}(x_{\mathrm{i}})\psi_r(x_{\mathrm{i}}) - \tilde{\beta}_{r,\mathrm{n}})^2 \tag{4.35}$$

Both $\tilde{\beta}_{\boldsymbol{r},n}$ and $\mathrm{SV}_{r,\mathrm{n}}$ can be updated simultaneously via

$$\hat{\beta}_r^{(\mathrm{n})} = \hat{\beta}_r^{(\mathrm{n}-1)} + \frac{1}{\mathrm{n}}(f(x_{\mathrm{k}})\psi_r(x_{\mathrm{k}}) - \hat{\beta}_r^{(\mathrm{n}-1)}) \tag{4.36}$$

$$\mathrm{SV}_r^{(\mathrm{n})} = \mathrm{SV}_r^{(\mathrm{n}-1)} + \frac{\mathrm{n}-1}{\mathrm{n}}(f(x_{\mathrm{k}})\psi_r(x_{\mathrm{k}}) - \hat{\beta}_r^{(\mathrm{n}-1)})^2 \tag{4.37}$$

The significance of the above updating formulas is that they require only a single pass over the data, and are numerically stable [28, 29].

Sensitivity Measure via Functional ANOVA

From the quasi Monte Carlo integration, we can use the betas, β_r to calculate the variance of each component of the ANOVA decomposition [26]

$$\sigma_u^2(f) = \sum_{r \in Q_u} \beta_r^2$$

Hence, the sensitivity measures of the component,

$$S_u(f) = \sigma_u^2(f) / \sigma^2(f)$$

We can estimate the error of approximation, using the calculated function values, via the following formula [28]:

$$\mathrm{MSE}(n) = \frac{1}{M} \sum_{k=n-M+1}^{n} (F(u_k) - \Psi^T(u_k)\hat{\beta}^{(k-1)})^2 \qquad (4.38)$$

And the average ISE is estimated using the Mean Squared Error (MSE) over $M = \sqrt{2n}$ [4] or $M = n^{2/3}$ [29] recent observations.

The Lack of Fit (LOF) is

$$\mathrm{LOF} = \frac{\mathrm{ISE}}{\mathrm{Var}[F(u)]} \qquad (4.39)$$

and describes the fraction of variance of the function $F(u)$ not explained by the regression ANOVA model. The LOF value can be estimated in practice by the following formula:

$$\mathrm{LOF}(n) = \frac{\mathrm{MSE}(n)}{SV_n^2} \qquad (4.40)$$

where SV_n^2 is the sample variance of the function $F(u)$.

Chapter 5.

Uncertainty & Sensitivity Analysis Study of Fault Location Algorithm

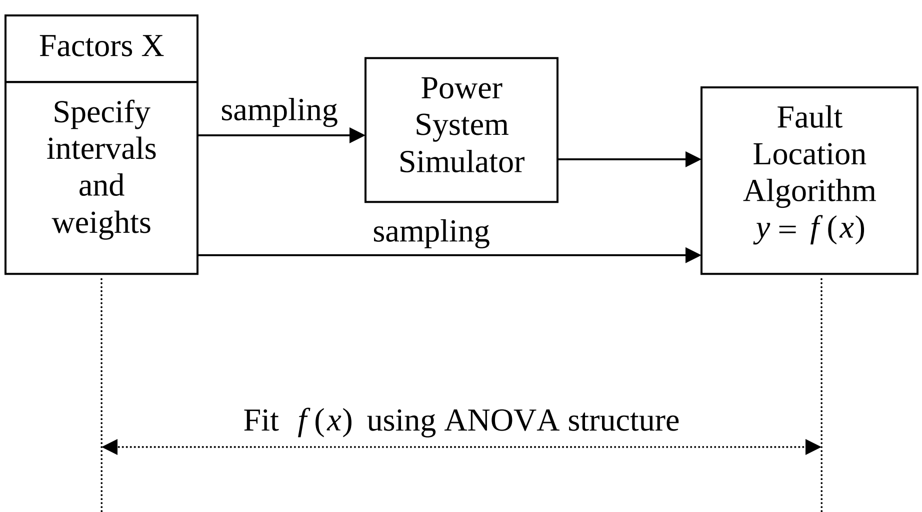


Figure 5.1: The Experimental Setup for the Sensitivity Analysis

Figure 5.1 illustrates the experimental setup for the sensitivity analysis on the fault locator. First, we identify the uncertain factors in the fault calculation and fault location algorithm. In this research, we have chosen a few of the system factors and setting factors which will be mention in the next section. All the factors will be scaled into specific intervals. The intervals of the factors will affect the result of the sensitivity analysis.

The sample points of system factors and setting factors is generated with Sobol's sampling strategy. The inputs of the Sobol [32] are dim and seed. Dim are the number of system and setting factors that I want to generate using i4_sobol function [33]. In our case, dim is equal to 9. Seed is for the initial sequence.

The system factors that have been generated are used to run power system simulator. The samples are then analyzed by the power system simulator, which produces the information that is required by the fault locator. The power system simulator simulates the different faults in the power system.

The results from the power system simulator and the setting factors are then sent to the fault location algorithm. The setting factors that have been generated are used as samples to be analyzed in the fault locator. In this research, I have tested 6 different fault location algorithms, which are Reactance method, Takagi 1 method, Takagi 2 method, Takagi 3 method,ABB method and Two-Ended Negative Sequence Impedance Method. The fault location algorithm is able to analyze and calculate the fault distance. I have compared the accuracy of the different algorithms which will be discussed later.

In the last stage, the fault location algorithm calculates the fault location value for each point in the factor space. The final fault location value and the system factors are analysed using ANOVA decomposition. The ANOVA is able to determine which system factors contribute the most and also determine the interaction between the system factors that influence the location result.

The goal of this study is to analyze in the systematic way how sensitive are the fault location algorithms to the selected factors and their interactions. The sensitivity measures are calculated using coefficients of the fitted approximating function (ANOVA decomposition) that is linking selected factors and their interactions to a fault location. The functions used in the tensor product basis where Legendre polynomials up to the order 5 but other orthogonal polynomials, wavelets or Fourier series can be used as well [5, 6].

5.1. Setup

Specification:

Uniform distribution for all factors in the intervals specified in the Table below.

Factors	**Interval**	**Description**
System Factors		
x_1	0 to 0.1 pu	Fault Resistance
x_2	-40 to 40 degrees	Load Flow Angle
Setting Factors		
x_3	-20 to 20 %	Percentage error of zero sequence setting
x_4	-10 to 10 %	Percentage error in line zero-sequence resistance Re{Z_0}
x_5	-10 to 10 %	Percentage error in line zero-sequence reactance Im{Z_0}
x_6	-10 to 10 %	Percentage error in s-source resistance Re{Z_{S0}}
x_7	-10 to 10 %	Percentage error in s-source reactance Im{Z_{S0}}
x_8	-10 to 10 %	Percentage error in r-source resistance Re{Z_{r0}}
x_9	-10 to 10 %	Percentage error in r-source resistance Im{Z_{r0}}
x_{10}	-4 to 4 %	Percentage error in receiving end current (I_{r2})
x_{11}	-4 to 4 %	Percentage error in sending end current (I_{S2})

Table 5.1: Intervals for all the Factors

5.2. Comparison between Takagi 3 and ABB

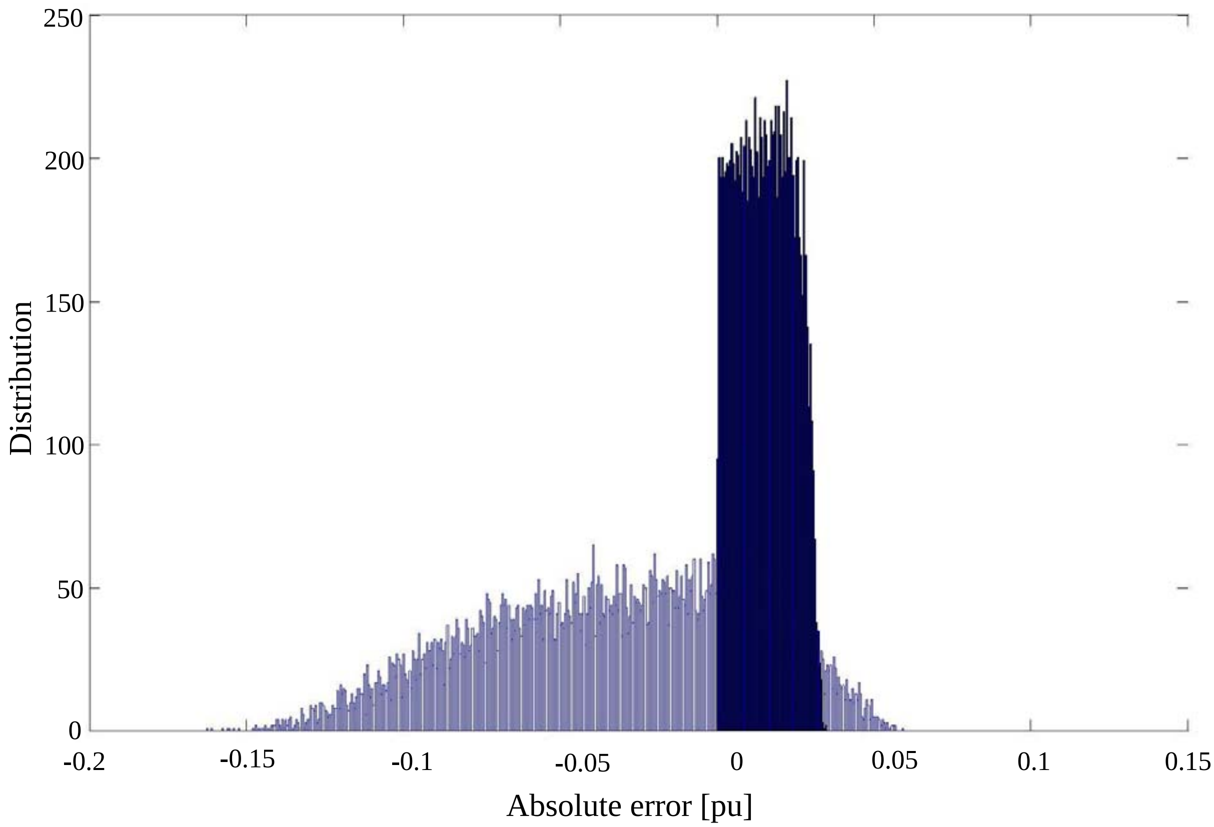


Figure 5.2: Comparison between Takagi 3 (transparent histogram) and ABB (blue histogram)

Figure 5.2 compares the accuracy of Takagi 3 and ABB methods. The distance to the fault that we are testing is 0.5 pu. We have used 12000 samples from the factor space to compare both of the algorithms. In this comparison we have chosen factors from x_1 to x_9. Figure 5.2 shows that Takagi 3 has larger variance compare to ABB method. Takagi 3 method has a bias of -0.0432 pu while ABB has a bias of 0.0286 pu. The ABB algorithm is more accurate than Takagi 3 method because ABB has a narrower variance and a smaller bias compared to Takagi 3. The reason is the ABB method is not sensitive to the source impedances and also zero sequence components. The sensitivity measures of the factors are listed in the Table 5.2.

Sensitivity Measures for the Takagi 3 Algorithm									
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	SM
1	0	0	0	0	0	0	0	0	0.022545
0	1	0	0	0	0	0	0	0	0.000309
0	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0.101792**
0	0	0	1	0	0	0	0	0	0.000004
0	0	0	0	1	0	0	0	0	0.000145
0	0	0	0	0	1	0	0	0	0.000092
0	0	0	0	0	0	1	0	0	0.000028
0	0	0	0	0	0	0	1	0	0.000002
0	0	0	0	0	0	0	0	1	0.000045

Sensitivity Measures for the ABB Algorithm									
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	SM
1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0.006905**
0	1	0	0	0	0	0	0	0	0.000148
0	0	1	0	0	0	0	0	0	0.000001
0	0	0	1	0	0	0	0	0	0.000006
0	0	0	0	1	0	0	0	0	0.000011

Table 5.2: Sensitivity Measures for Takagi 3 and ABB

The sensitivity measures in Table 5.2 can be used to quantify and understand influence of uncertain factors on fault location. Table 5.2 does not contain the interaction but only the significant factors are included. The sensitivity measures for single phase to ground are calculated at fault distance 0.5 pu. The sensitivities give us the following conclusions:

a) Takagi 3 algorithm is more sensitive to zero sequence setting (x_3).

b) ABB algorithm is not sensitive to zero sequence components such as x_3, x_4 and x_5.

c) ABB algorithm is not sensitive to source impedances at all such as x_6, x_7, x_8 and x_9. The sensitivity measures are zero.

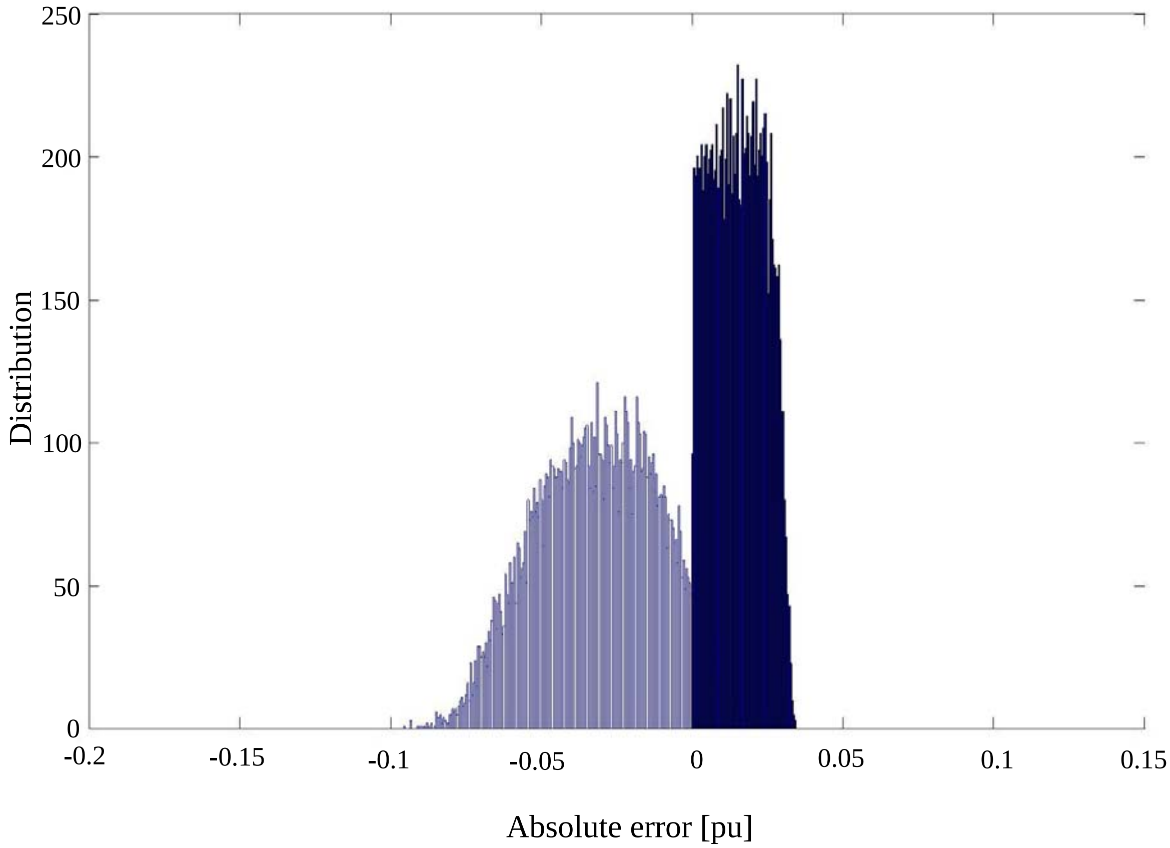


Figure 5.3: Comparison between Takagi 3 and ABB (x_3 interval decreased)

Figure 5.3 compares the accuracy of Takagi 3 and ABB methods with x_3 interval decreased. I have used the same system and setting factors. I have decreased the interval for the percentage error of zero sequence setting from [-20 to 20] to [-5 to 5]. We decrease the interval to see the effect on both of the algorithms. Accuracy improvement is more significant for Takagi 3 because it is sensitive to zero sequence setting. ABB method has the similar sensitivity measures as in the previous discussion. This is because ABB method is not sensitive to zero sequence setting.

Sensitivity Measures for the Takagi 3 Algorithm									
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	SM
1	0	0	0	0	0	0	0	0	0.025175
0	1	0	0	0	0	0	0	0	0.000252
0	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0.007054**
0	0	0	1	0	0	0	0	0	0.000005
0	0	0	0	1	0	0	0	0	0.000195
0	0	0	0	0	1	0	0	0	0.000113
0	0	0	0	0	0	1	0	0	0.000033
0	0	0	0	0	0	0	1	0	0.000003
0	0	0	0	0	0	0	0	1	0.000058

Sensitivity Measures for the ABB Algorithm									
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	SM
1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0.006941**
0	1	0	0	0	0	0	0	0	0.000235
0	0	1	0	0	0	0	0	0	0.000006
0	0	0	1	0	0	0	0	0	0.000002
0	0	0	0	1	0	0	0	0	0.000037

Table 5.3: Sensitivity Measures for Takagi 3 and ABB (x_3 interval decreased)

The sensitivity measures for single phase to ground are calculated at fault distance 0.5 pu. The sensitivities give us the following conclusions:

a) Takagi 3 algorithm is more sensitive to zero sequence setting (x_3) but the sensitivity measures reduces compare to Table 5.2.

b) ABB algorithm is not sensitive to zero sequence components such as x_3, x_4 and x_5.

c) ABB algorithm is not sensitive to source impedances at all such as x_6, x_7, x_8 and x_9. The sensitivity measures are zero.

d) Both of the methods are not sensitive to the impedances.

5.3. Comparison between Takagi 2 and Two-Ended Negative-Sequence Impedance Method

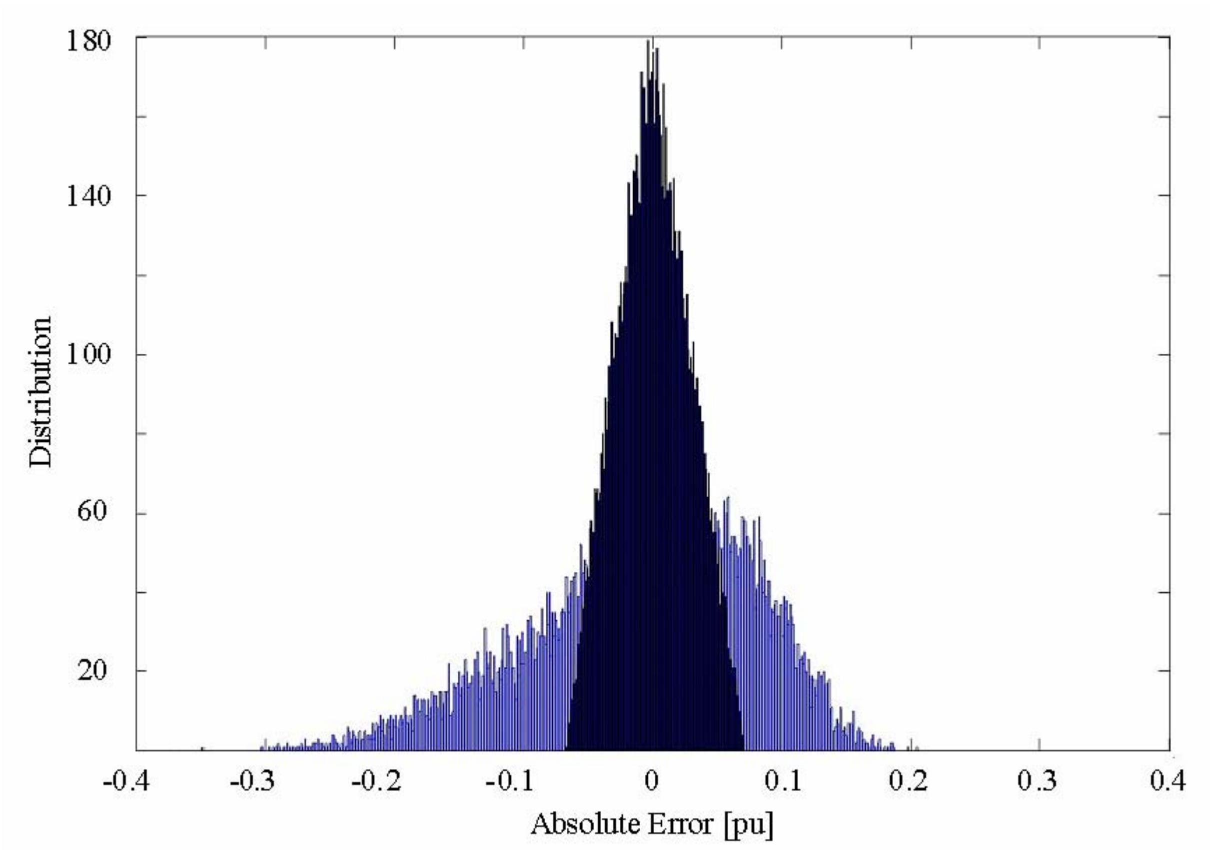


Figure 5.4: Comparison between Takagi 2 (transparent histogram) and Two-Ended Negative-Sequence Impedance Method (blue histogram)

Figure 5.4 compares the accuracy of Takagi 2 and two-ended negative-sequence impedance method. The distance to the fault that we are testing is 0.9 pu. In this comparison, we have included two more extra input factors which are negative sequence of receiving-end current and negative-sequence of sending-end current. We have included these two factors because the two-ended impedance method is not sensitive to any of the factors in Table 5.1 except the negative sequence currents. Figure 5.4 shows that Takagi 2 has a larger variance than the two-ended impedance method. In this comparison, the two-ended negative-sequence impedance method is more reliable.

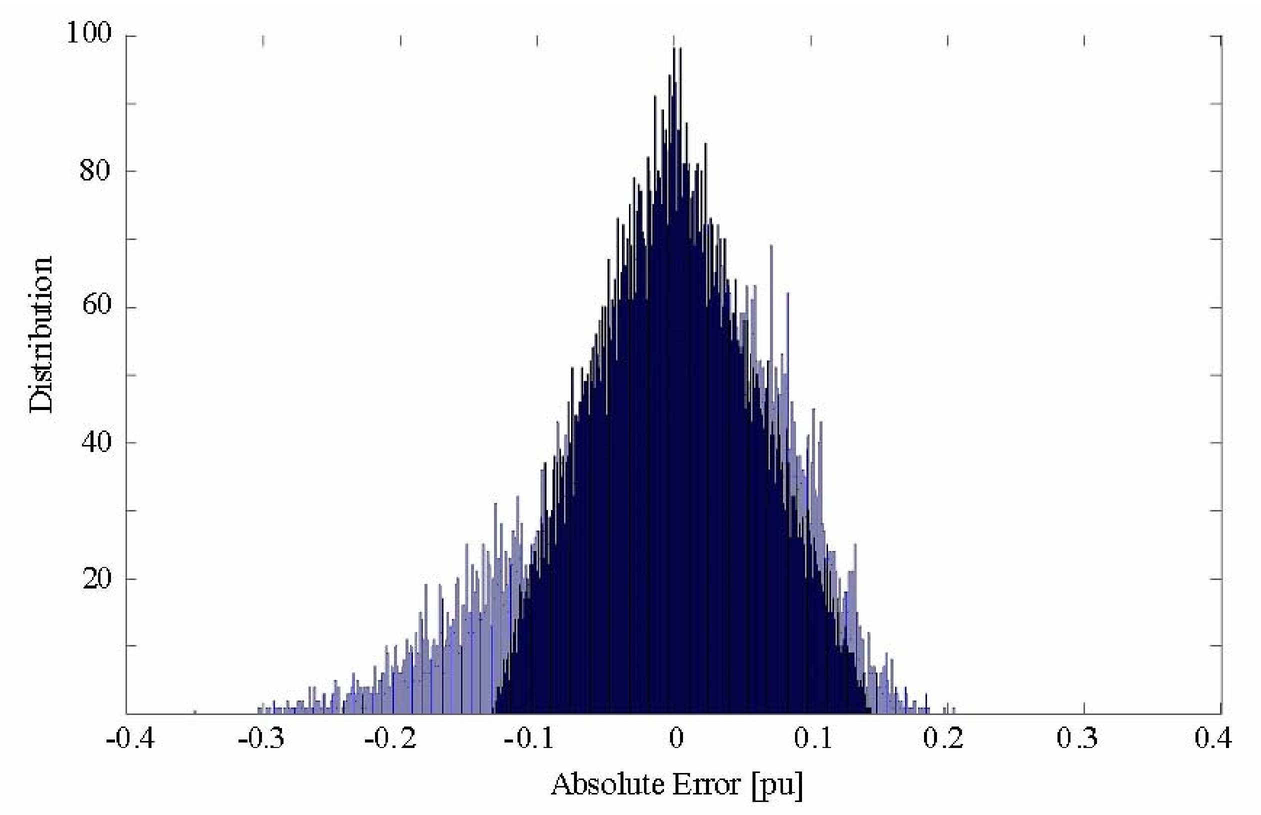


Figure 5.5: Comparison between Takagi 2 and Two-Ended Negative-Sequence Impedance Method (x_{10} & x_{11} interval increased)

Figure 5.5 compares the accuracy of Takagi 2 and two-ended negative-sequence impedance methods with x_{10} and x_{11} interval increased. I have used the same system and setting factors. I have increased the interval for the percentage error of negative sequence currents from [-4 to 4] to [-8to 8]. We increase the interval to see the effect on both of the algorithms. Takagi2 method has the similar distribution as the Figure 5.4. This is because Takagi2 method is not very sensitive to negative sequence currents. The two-ended impedance method has much larger variance compare to Figure 5.4. From both of the figures, we conclude that the two-ended negative-sequence impedance method is accurate when the percentage error of negative sequence currents is small. When the percentage error increases, Takagi 2 method is more reliable than the two-ended impedance method.

5.4. Lack of Fit (LOF)

The LOF estimate determines the approximation accuracy achieved in the latest update.

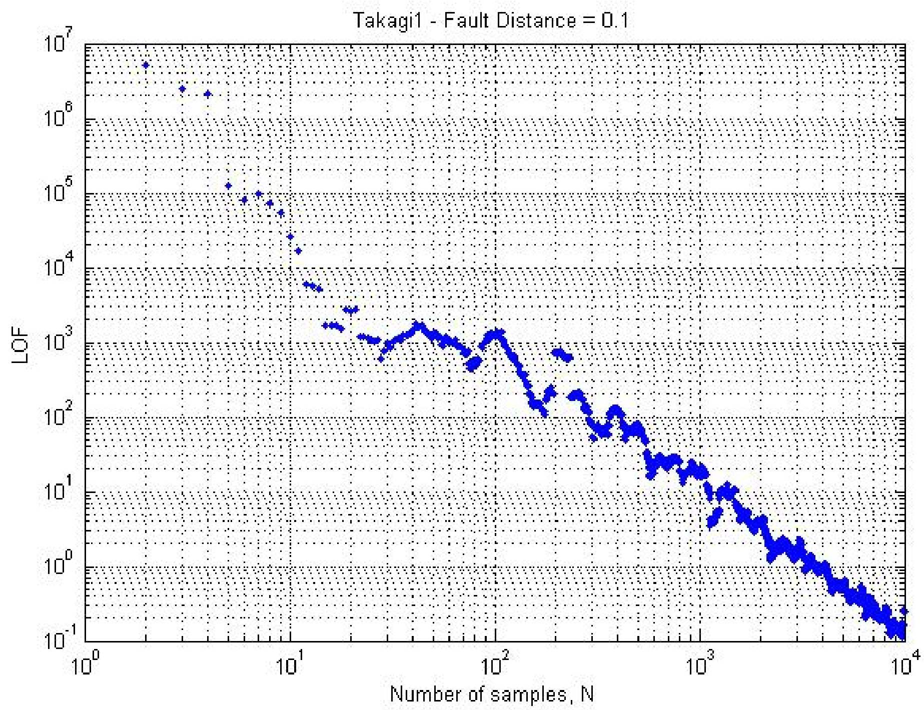


Figure 5.6: LOF for the Takagi 1 Method at Fault Distance 0.1 pu

Figure 5.6 shows LOF as a function of sample size for Takagi 1 method at fault distance 0.1pu. When the fault distance increases, LOF increases too. In order to decrease the LOF, sample size has to be increased but this will consume time. The LOF for Reactance, Takagi 2 and Takagi 3 methods are very similar.

5.5. Single Phase to Ground Short Circuit Case Study

We can observe in Figure 5.7 that as the fault distance increases, the sensitivity of fault resistance for Takagi's methods decrease exponentially. While the sensitivity measures on fault resistance for Reactance Method is not significant compared to Takagi's methods. Takagi's methods are more sensitive than Reactance Method to fault resistance factor.

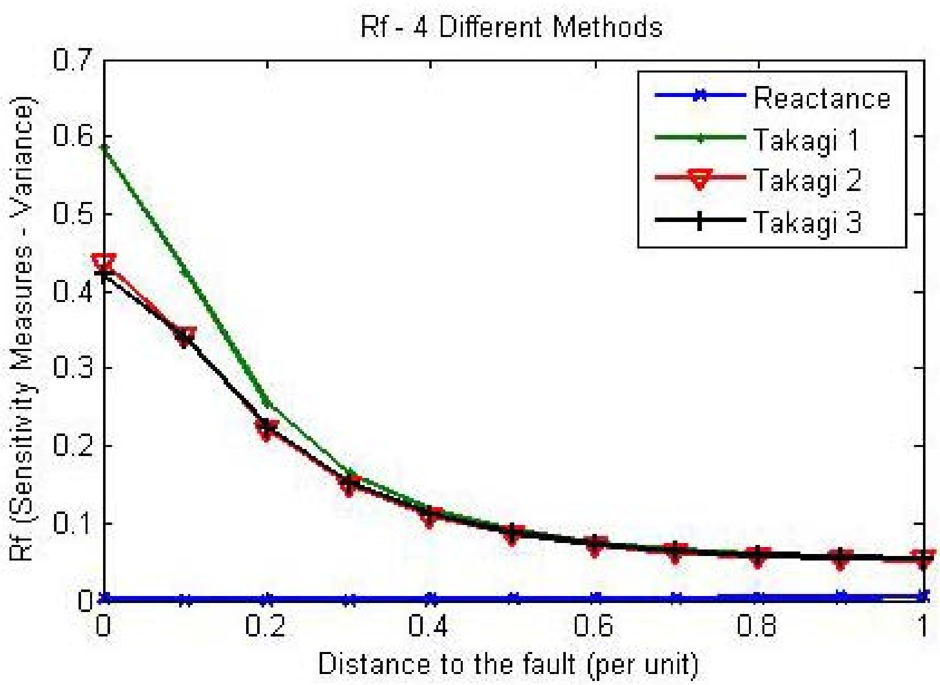


Figure 5.7: Fault Resistance (R_f) Sensitivity Measures for 4 Methods with Fault Distance varying from 0 to 1pu

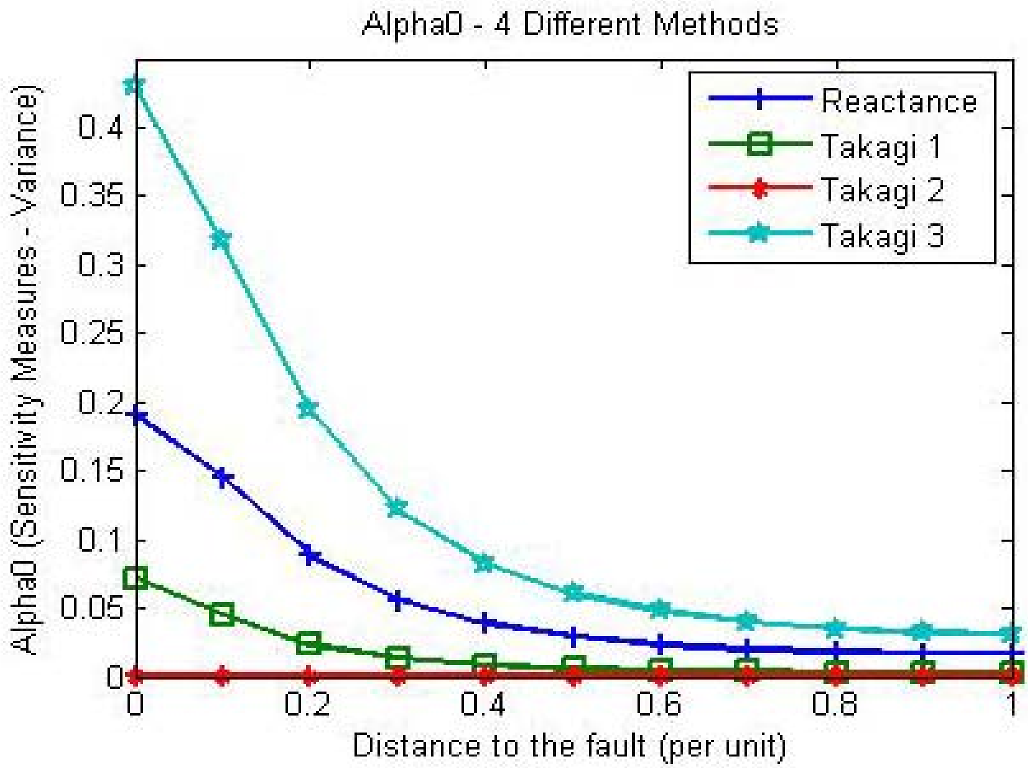


Figure 5.8: Zero Sequence Sending End Source Impedance Angle (Alpha0) Sensitivity Measures for 4 Methods with Fault Distance varying from 0 to 1pu

Figure 5.8 shows how the fault location result will be affected by the sensitivity to zero sequence sending end source impedance angle. The sensitivity measures for 4 methods shown in Figure 5.8 are decreasing exponentially when the fault distance increases. Takagi 3 method is the most sensitive to the zero sequence sending end source impedance angles because the method uses of zero sequence current. Takagi 2 is not sensitive because the method does not use any zero sequence information.

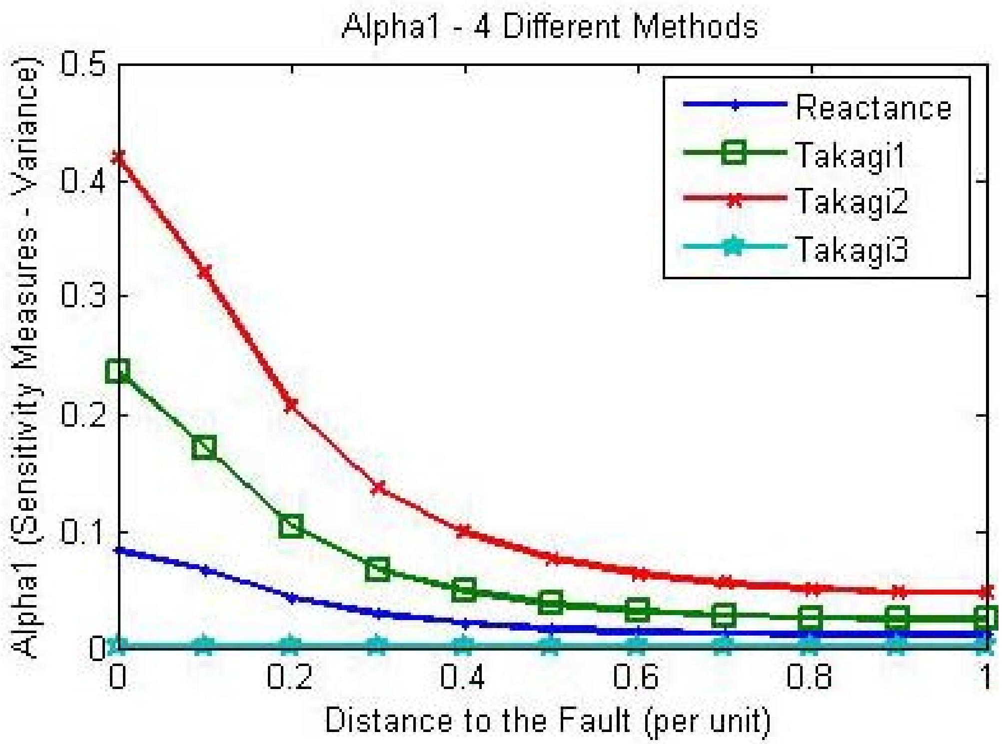


Figure 5.9: Positive Sequence Sending End Source Impedance Angle (Alpha1) Sensitivity Measures for 4 Methods with Fault Distance varying from 0 to 1pu

Figure 5.9 shows how the fault location result will be affected by the sensitivity to positive sequence sending end source impedance angle. The sensitivity measures for 4 methods shown in Figure 5.9 are decreasing exponentially when the fault distance increases. Takagi 2 method is the most sensitive to the positive sequence sending end source impedance angle because the method uses of positive sequence current. Takagi 3 is not sensitive because the method does not use any positive sequence information.

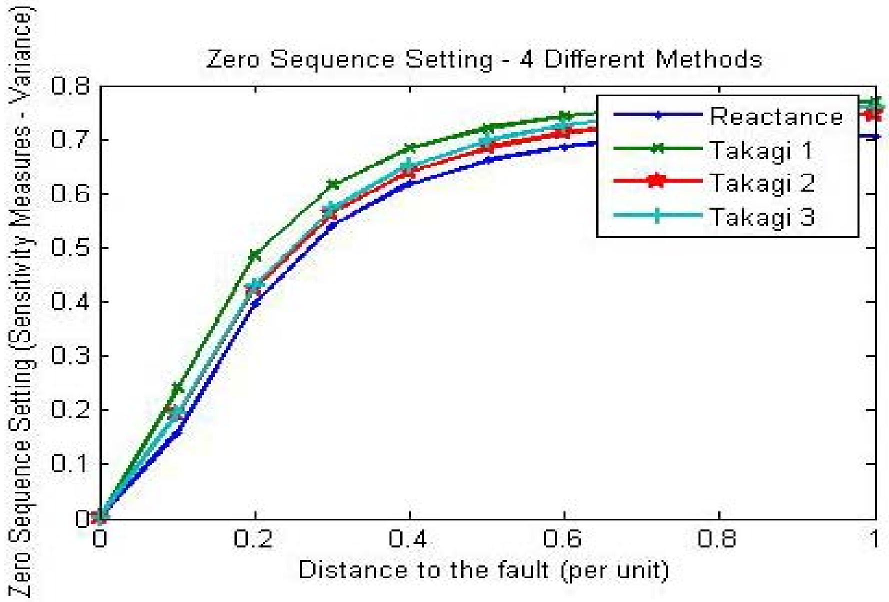


Figure 5.10: Percentage Error of Zero Sequence Setting Sensitivity Measures for 4 Methods with Fault Distance varying from 0 to 1pu

We can observe in Figure 5.10 that the sensitivity of percentage error of zero sequence setting for 4 different methods increases as the fault distance increases. Takagi 1 method is the most sensitive to the percentage error of zero sequence setting, and follows with Takagi 3, Takagi 2 and Reactance methods. The sensitivity measures for all 4 methods are similar.

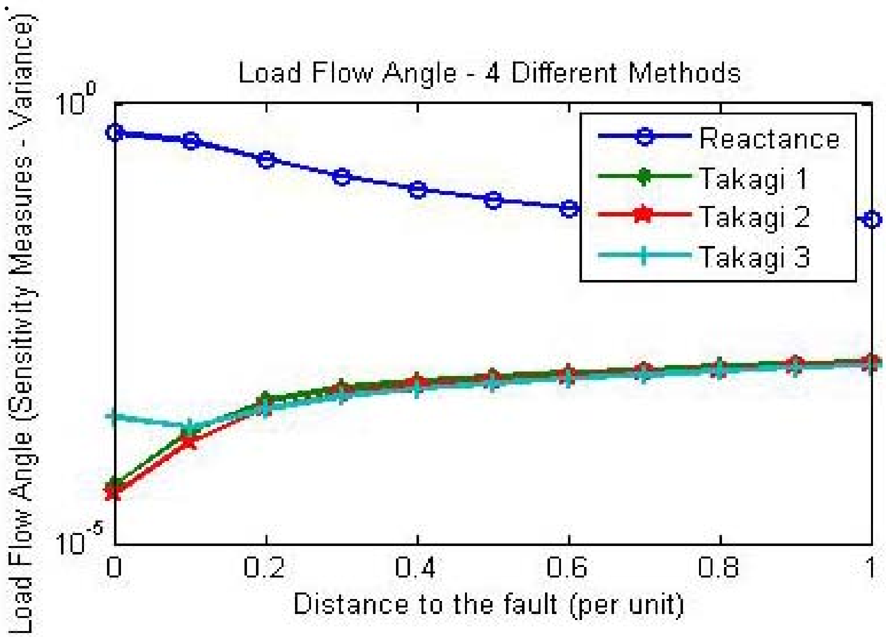


Figure 5.11: Load Flow Angle Sensitivity Measures for 4 Methods with Fault Distance varying from 0 to 1pu

Figure 5.11 shows the sensitivity measures on load flow angle for 4 methods. Reactance method is the most sensitive. The Reactance method sensitivity is decreasing gradually when the fault distance increases. It has a significant sensitivity value compared to Takagi's methods. Takagi's methods have similar sensitivity values and they increase gradually when the fault distance increases. The Reactance method is the most sensitive because it is not compensating for the pre-fault load flow current.

Main Effects	Reactance Method	Takagi Method 1	Takagi Method 2	Takagi Method 3
x1	0.001401	0.090125	0.085048	0.087147
x2	0.079933	0.000783	0.000714	0.000665
x3	0.016260	0.037654	0.076595	0.000316
x4	0.029110	0.006083	0.000564	0.060551
x5	0.661947	0.720518	0.684806	0.697948
Interaction				
x1x2	0.043223	0.001778	0.001717	0.001707
x1x3	0.006815	0.013019	0.028482	0.000733
x1x4	0.014126	0.005216	0.001304	0.032062
x1x5	0.004015	0.003207	0.003100	0.003006
x2x3	0.001259	0.001111	0.001051	0.001199
x2x4	0.001498	0.001097	0.000983	0.001356
x2x5	0.006364	0.001210	0.001166	0.001137
x3x4	0.002196	0.001969	0.001935	0.001797
x3x5	0.001443	0.001170	0.001289	0.001506
x1x2x4	0.012656	0.011240	0.010793	0.010701
x1x3x5	0.025243	0.022969	0.021829	0.022299
x2x3x5	0.023636	0.020318	0.020013	0.018361
x2x4x5	0.012704	0.010990	0.010949	0.009767
x2x3x4x5	0.014962	0.013518	0.013036	0.012713

Table 5.4: Sensitivity Measures (Fault Location is 0.5pu)

*x*1 – Fault Resistance,

*x*2 – Load Flow Angle,

*x*3 – Positive Sequence Sending-End Source
Impedance Angle,

*x*4 – Zero Sequence Sending-End Source Impedance Angle,

*x*5 – Percentage Error of Zero Sequence Setting

Based on the achieved ANOVA decomposition, assessing of the influence or relative importance of each factor on the fault location accuracy can be delivered. Sensitivity measures [8] which are normalized variances, for the fault location methods tested, are presented in Table 5.4. The sensitivity measures in Table 5.4 can be used to quantify and understand influence of uncertain factors on fault location. Table 5.4 does not contain all the interactions but only the significant interactions are included. The sensitivity measures for single phase to ground are calculated at fault distance 0.5pu. The sensitivities give us the following conclusions:

a) Reactance Method is more sensitive to load flow angle (x_2).
b) Takagi 1 Method is more sensitive to fault resistance (x_1) and percentage error of zero sequence setting (x_5).
c) Takagi 2 Method is more sensitive to positive sequence sending-end source impedance angle (x_3).
d) Takagi 3 Method is more sensitive to zero sequence sending-end source impedance angle (x_4).
e) The combination of fault resistance and positive sequence sending-end source impedance angle (x_1x_3) can influence considerably Takagi 2 method while the other methods are not sensitive.
f) The combination of fault resistance and zero sequence sending-end source impedance angle (x_1x_4) can influence considerably Takagi 3 method while the other methods are not sensitive.

g) Reactance method is more sensitive to the combined effect of 3 factors, $x_1x_2x_4$ (0.012654), $x_1x_3x_5$ (0.025243), $x_2x_3x_5$ (0.023636), $x_2x_4x_5$ (0.012704). The other methods are not sensitive.

5.5. Three Phase Short Circuit Case Study

In three phase short circuit case, I only analyze positive sequence network. Takagi 3 method is not suitable for this case because it uses zero sequence current. The sensitivity measures using Takagi 2 method is the same as Takagi 1 method because the phase A current in Takagi 1 is equal to the positive sequence current in Takagi 2. The sensitivity measures for three phase short circuit case are set at fault distance 0.5pu.

The sensitivities in Table 5.5 lead us to the following conclusion:

a) Reactance Method is more sensitive to load flow angle (x_2) and positive sequence sending-end source impedance angle (x_3).

b) Takagi 1 Method is more sensitive to fault resistance (x_1) and positive sequence sending-end source impedance angle (x_3).

c) The combination of fault resistance and load flow angle (x_1x_2) can influence considerably Reactance method while Takagi 1 method is not sensitive.

d) The combination of fault resistance and positive sequence sending-end source impedance angle (x_1x_3) can influence considerably Takagi 1 method while Reactance method is not sensitive.

Main Effects	Reactance Method	Takagi Method 1
x1	0.006820	0.330659
x2	0.351501	0.000256
x3	0.262641	0.326960
x4	0.000222	0.000275
x5	0.000139	0.000150
Interaction		
x1x2	0.120738	0.002735
x1x3	0.083725	0.104496
x1x4	0.001671	0.002386
x1x5	0.007038	0.009560
x1x3x5	0.030805	0.041855
x2x3x5	0.015035	0.019654
x2x4x5	0.018641	0.025347
x2x3x4x5	0.011496	0.015609

Table 5.5: Sensitivity Measures (Fault Location is 0.5pu)

Chapter 6.

Conclusion

6.1. Fault Location Algorithms

Overview of fault location techniques for power transmission lines is presented. Distinctive features of different fault location algorithms classified as the impedance techniques are discussed.

One-end fault location for traditional uncompensated transmission lines compensating for the infeed effect under resistive faults is presented. The issue of the location accuracy and its improvement is discussed. Measurements from impedance relays at the line terminals can be processed. The accuracy of the fault location algorithms has been tested and compared. The accuracy is mainly affected by the system and setting factors that I have chosen. In this research, I have tested one-ended fault location algorithm and negative-sequence two-ends impedance method. The two-end fault location algorithm can be more accurate but data must be captured from both ends before the algorithm can be applied.

6.2. Uncertainty & Sensitivity Analysis

Sensitivity Analysis is popular in financial applications, risk analysis, signal processing, neural networks and any area where models are developed. Sensitivity analysis can also be used in model-based policy assessment studies. Global sensitivity analysis, proposed in this thesis, analyzed the whole set of potential input factors and aim to give an overall indication of the way that the outcome varies, and in particular how the output varies in response to the input variations within the range of parameter uncertainty. Monte Carlo methods and variety of sampling strategies are implemented in global sensitivity analysis.

Quasi regression is a method that is used to approximate a function on the unit cube in *d* (*number of variables in the particular function*) dimensions. Fast computation time and ANOVA decomposition are the advantages. As statistical approaches to the approximate function, univariate basis functions, tensor products and orthogonal polynomials are introduced.

Analysis of Variance (ANOVA) decomposition has been chosen as a tool for global sensitivity analysis of the fault location. The ANOVA allows us to quantify the notion that some variables and interactions are much more important than others. Computing ANOVA decomposition is a sampling based process in which the power system simulator and a fault locator are both executed repeatedly for combinations of values sampled from the assumed distribution of the input factors.

I had compared the sensitivity measures of all the uncertain factors for the fault location algorithms. In chapter 4, I had discussed the sensitivity measures which can be used to quantify and understand the influence of uncertain factors on fault location. The interaction of the uncertain factors is discussed in this chapter. Based on the achieved ANOVA decomposition, assessing of the influence or relative importance of each factor on the fault location accuracy can be delivered.

6.3. Future Work

The uncertainty and sensitivity analysis study of fault location algorithms had shown the idea of such analysis can help in selecting the optimal locator for a specific application and also can help in the calibration process in practice. It is therefore to think promisingly that the global sensitivity analysis can be implemented for testing the practical fault locator in future.

Bibliography

[1] M.M.Saha, K.Wikstrom, J.Izykowski, E.Rosolowski. "Fault Location Techniques," in *Department of Electrical Engineering Wroclaw, Poland & Department TTD ABB Automation Products AB SE-721 59 Vasteras, Sweden.*

[2] Zivanovic, R. and H.B. Ooi (2007). "Global Sensitivity Analysis of Fault Location Algorithm," *AUPEC'07,* (Perth, Australia), 9-13, Dec. 2007.

[3] J.D. Glover, M.Sarma. *Power System Analysis and Design.* Boston: PWS-KENT Publishing Company, 2001

[4] R.C.Dorf, J.A.Svoboda. *Electric Circuits*. NY: John Wiley & Sons, 2001

[5] V.D. Toro. *Electric Power Systems*. NJ: Prentice-Hall Englewood Cliffs, 1992.

[6] S.A. Nasar, F.C. Trutt. *Electric Power Systems*. Florida: CRC Press, 1999.

[7] H.Saadat. *Power System Analysis*. Boston: McGraw-Hill, 1999.

[8] K.Zimmerman and D.Costello, "Impedance-Based Fault Location Experience," *SEL Technical Paper, available at http://www.selinc.com/techpprs/SEL_Zimmerman_ImpedBFL_6180_20041004.pdf*

[9] E.O.Schweitzer. "Evaluation and Development of Transmission Line Fault-Locating Techniques which Use Sinusoidal Steady-State Information," in *Ninth Annual Western Protective Relay Conference*, (Spokane, Washington), Oct. 1982.

[10] T.Takagi, at al. "Development of a New Type Fault Locator using the One-Terminal Voltage and Current Data," *IEEE Transactions on Power Apparatus and Systems, Vol. PAS-101, No.8, pp.2892-2898*, Aug. 1982.

[11] A.Wiszniewski. "Accurate Fault Impedance Locating Algorithm," in *IEE Proceedings, Part C, Vol.130, No.6, pp 311-315*, 1983

[12] L.Eriksson, M.M.Saha, G.D.Rockefeller. "An Accurate Fault Locator with Compensation for Apparent Reactance in The Fault Resistance Resulting from Remote-End Infeed," in *IEEE/PES 1984 Summer Meeting*, (Seattle, Washington), 1984.

[13] R.Poncelet. "Fault Locator, Complete Documentation," in *University Libre De Bruxelles*, Mar. 1995.

[14] I.Zanora, J.F.Minambres, A.J.Mazon, R.Alvarez-Isasi, J.Lazaro. "Fault Location on Two-Terminal Transmission Lines Based on Voltages," in *IEE Proceedings – Gener. Transm. Distrib., Vol.143, No.1, pp. 1-6*, Jan. 1996

[15] D.A.Tziouvaras, J.B. Roberts, G.Benmouyal. "New Multi-ended Fault Location Design for Two- or Three-Terminal Lines" in *Schweitzer Engineering Laboratories, Inc.* (USA).

[16] I.M.Sobol and S.Tarantola. "Global Sensitivity Analysis for Non-Linear Simulation Models," *International Journal of Non-Linear Mechanics*.

[17] A.Saltelli, S.Tarantola, F.Campolongo and M.Ratto. *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models.* John Wiley & Sons Ltd, 2005.

[18] W.Chen, R.Jin, A.Sudjianto. "Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design Under Uncertainty" *available at* *http://ideal.mech.northwestern.edu/pdf/JMD_Analytical.pdf*

[19] R.Zivanovic, H.B.Ooi. "Sensitivity Analysis of Transmission Line Fault Location," *IEEE Powertech 2007*, Paper No. 498, (Lausanne Switzerland), 1-5, July. 2007.

[20] O.Bonin. "Sensitivity Analysis and Uncertainty Analysis for Vector Geographical Applications," in *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.*

[21] K.Chan, A.Saltelli, S.Tarantola, "Sensitivity Analysis of Model Output: Variance-Based Methods Make the Difference," in *Proceedings of the 1997 Winter Simulation Conference*, 1997.

[22] W.Chen, R.Jin, A.Sudjianto. "Analytical Uncertainty Propagation in Simulation-Based Design under Uncertainty", AIAA-2004-4356, *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, (Albany, NY), August 30-September 1, 2004.

[23] I.M.Sobol. "Sensitivity analysis for nonlinear mathematical models," in *Mathematical Modeling and Computational Experiment*, Vol. 1, pp.407–414, 1993.

[24] A.Sauhats, A.Jonins, V.Chuvychin, M.Danilova. "Fault Location Algorithms for Power Transmission Lines Based on Monte-Carlo Method," in *IEEE Porto Power Tech Conference*, (Porto, Portugal), 2001.

[25] M.Bockarjova, A.Sauhats, G.Andersson. "Statistical Algorithm for Power Transmission Lines Distance Protection," in *9th International Conference on Probabilistic Methods Applied to Power Systems,* (Stockholm, Sweden), June, 2006.

[26] R.Liu, A.B.Owen. "Estimating mean dimensionality of ANOVA decompositions," *Technical report*, *Stanford University*, *Statistics Department*, 2005.

[27] J.J.Faraway. "Practical Regression and ANOVA using R," *www.stat.lsa.umich.edu*, 2002.

[28] An, J. and A.B. Owen. "Quasi Regression," *Journal of Complexity*, Vol. 17, pp.588-607, 2001.

[29] Jiang, T. and A.B.Owen. "Quasi-regression for visualization and interpretation of black box functions," *Technical Report, available at http://www.stat.stanford.edu/~owen/reports/qregvis.pdf*, 2002.

[30] T.Jiang, A.B. Owen. "Quasi-regression with shrinkage." *Technical report*, *Stanford University*, *Statistics Department*, 2001.

[31] A.Antoniadis. "Outils statistiques pour I' analyse de sensibilite : Analyse de la variance et quasi-regression," *Toulouse 2nd*, 2006.

[32] J.Burkardt. "Source Codes in Fortran90," *http://www.csit.fsu.edu/~burkardt/f_src/f_src.html*

[33] I.M.Sobol. *A primer for the Monte Carlo Method*, CRC Press, Inc, 1994.

[34] A.B.Owen. “Multidimensional variation for quasi-Monte Carlo,” *Technical report, Stanford University, Statistics Department*, 2004.

[35] H.Faure. “Monte Carlo and Quasi-Monte Carlo Methods for Numerical Integration,” *Institute de Mathematiques de Luminy, U.P.R. 9016 CNRS 163 avenue de Luminy, case 907, F-13288 Cedex 09*, (Marseille, France).

[36] G.Wei. “Monte Carlo and Quasi-Monte Carlo Methods,” *Department of Mathematics, Hong Kong Baptist University*, (Hong Kong).

[37] K.M.Hanson. “Quasi-Monte Carlo : halftoning in high dimensions.” *Proc. SPIE 5016, pp. 161-172,* 2003.

[38] R.Reuillon and D.Hill. “Faster Quasi Random Number Generator for Most Monte Carlo Simulations,” *Blaise Pascal University*, 2005.

[39] A.B.Owen. “Quasi-Monte Carlo Sampling,” *Technical report, Stanford University, Statistics Department*.

[40] C.Lemieux. “Randomized Quasi-Monte Carlo: A Tool for Improving the Efficiency of Simulations in Finance,” *Winter Simulation Conference*, 2004.