ROBUST STATISTICAL METHODS FOR MEASUREMENT CALIBRATION IN LARGE ELECTRIC POWER SYSTEMS

Alireza Ghassemian

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

L. M. Mili, Chair
C. W. Coakley
A. A. Beex
Y. Liu
R. P. Broadwater

October, 1997
Blacksburg, Virginia

Keywords: Robust statistic, Measurement Calibration

ROBUST STATISTICAL METHODS FOR MEASUREMENT CALIBRATION IN LARGE ELECTRIC POWER SYSTEMS

by

Alireza Ghassemian
L. M. Mili, Chairman
Electrical Engineering

(ABSTRACT)

The objective of the Remote Measurements Calibration (RMC) method is to minimize systematic errors through an appropriate scaling procedure. A new method for RMC has been developed. This method solves the problems of observability, multiplicity of solutions, and ambiguity of reference points associated with the method proposed by Adibi *et al*. [6-9]. The new algorithm uses the simulated annealing technique together with the matroid method to identify and minimize the number of RTUs (Remote Terminal Units) required to observe the system. After field calibration, these RTUs provide measurements that are used to estimate the whole state of the system. These estimates are then returned as a reference for remotely calibrating the remaining RTUs. The calibration coefficients are estimated by means of highly robust estimator, namely the Least Median of Squares (LMS) estimator. The calibration method is applicable to large systems by means of network tearing and dynamic programming. The number of field calibrations can be decreased further whenever multiple voltage measurements at the same buses are available. The procedure requires that the measurement biases are estimated from recorded metered values when buses, or lines, or transformers are disconnected. It also

requires the application of a robust comparative voltage calibration method. To this end, a modified Friedman test has been developed and its robustness characteristics investigated.

ACKNOWLEDGMENTS

I would like to thank Dr. L. M. Mili for his years of tutelage and financial support during my graduate studies. His patience, hard work, and dedication has been an excelent example for me as a graduate student. I would also like to express my appreciation to Dr. C. W. Coakley for all his help. I wish him the best of luck in his career as a faculty member. I would also like to express my thanks to Dr. A. A. Beex, Dr. Y. Liu, and Dr. R. P. Broadwater for serving as members of my Ph.D. advisory committee.

Finally, I would like to thank all my family and friends for doing nothing but everything, to make it possible for me to finish this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Necessity of Measurement Calibration

An electrical power system is a combination of lines and transformers that connect the output of generators to the consumers. Because only a small amount of electrical energy can be stored, production should meet the continuous random fluctuation of demand at every instant of time. In addition, the system has to minimize the cost of delivered energy while maintaining the system in a normal and secure operating state. A balance between economy and security can be achieved through a state estimation function, which is responsible for providing a complete and reliable database from the collected measurements. State estimation accuracy highly relies on the measurement accuracy.

In a power system, practical experience has shown that discordant measurements, which are referred to as "outliers", appear at each state estimation run [1]. Several sources of outliers have been reported in the literature [2-7]. Major sources are

i) Model approximation, including measurement and network modeling errors;

ii) Rounding errors, which result from analog-to-digital conversion and scaling, and hence affect all the metered values;

iii) Gross measurement errors, which are due to device failures, incorrect wiring, non- instantaneous meter scan (time-skew), and bad calibration. The latter is due to inappropriate scaling procedures that do not properly compensate the systematic errors induced by the instruments in the measurement chains;

iv) Parameter errors, which contaminate the values of the π-equivalent circuit elements; and

v) Topology errors, which are generated by change in the status of the system.

In general, model approximations give rise to temporary small or medium size outliers, whereas device failures and bad calibrations cause permanent large size outliers. Figure 1 is a graphical representation of sequence of events that carries an analog field measurements to the control center. An instrument providing a measurement chain is a collection of voltage and current transformers that are connected to an analog-to-digital converter through voltage, current, and power transducers. After conversion, the binary measurement signals experience several stages of processing before reaching the control center through communication devices, each of which provides an opportunity for the introduction of an error. These errors are basically of four types. First type errors result from the non-linearity of the transformers. Second type errors are random in nature and relate to instrument precision. Third type errors are systematic and are caused by the drift and deterioration of instruments with time, temperature and environment. The fourth type includes those intermittent errors induced by the communication system.

The combined effects of these errors with non-linearity translate into outliers in the measurements, the so-called "bad data." The most effected measurements are the voltage measurements that are obtained from the capacitive-coupled voltage transformer (CVTs) [6]. A CVT-based voltage measurement is likely to be several percentage points off because of the aging and temperature dependencies of the capacitors. Typical errors in voltage magnitude of the order of 5-7% may be expected. Even when the CVTs are calibrated or newly installed, this variation is to be expected as the ambient conditions change.

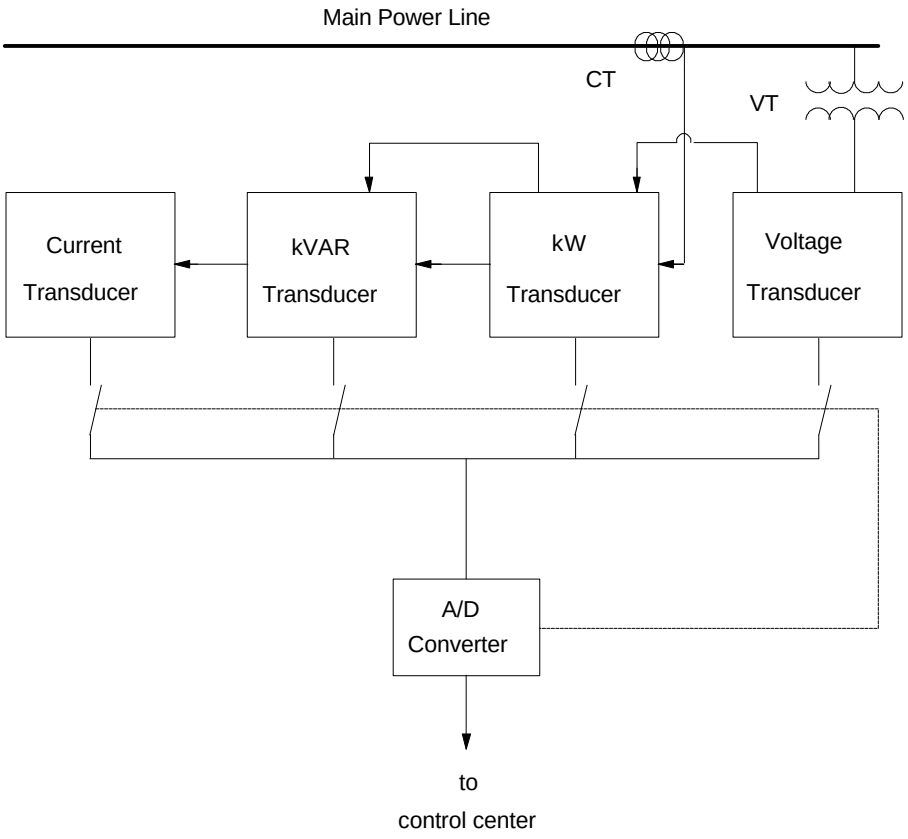

Figure 1. Power Line Measurements and Telemetry Schematic (taken from [6])

This magnitude error could translate into a very significant error in the estimate of the reactive power flow over a high voltage line. The metered values of power injections and power flows are also affected, for they rely on current transformers that operate over the full range of the saturation curve. Now, it is true that bad data can be handled by robust state estimation procedures; however, their number should not exceed half of the number of the redundant measurements. Indeed, when the bad data outnumbers the good, one can no longer distinguish between them, which induces the breakdown of any estimator. Therefore, a prerequisite to the state estimation function is that the majority of the measurements must be good data points, that is, suitably calibrated so that their bias remains small over the full scale range of the instruments.

The objective of measurement calibration is to minimize the systematic errors through an appropriate scaling procedure. The current practice is to calibrate the instruments on site. This operation requires the transportation of field calibration equipment and the connection/disconnection of terminals. Due to the large number of measurements and the time required to check each measurement, field calibrations are performed much less often than is desirable.

1.2 Review of Previous Work

To overcome these difficulties, Adibi *et al.* [6-9] suggest the use of the Remote Measurement Calibration (RMC) as a more efficient technique. The proposed method takes advantage of the availability, at the substation level and at the control center, of a large number of repetitive readings of the same measurement points over a wide range of their full scale. Accordingly, it seeks to estimate the calibration parameters (off-set, gain at the origin, and non-linearity coefficients) of the measurements situated around selected nodes on the basis of Ohm's and Kirchhoff's laws, which the industry refers to as Voltage (KV), Ampere, and Volt-Ampere (MVA) equality and bus summing. The KV equality expresses that the voltages on the branches connected to the same bus are equal, whereas the Ampere equality expresses that the primary current of a transformer equals the product

of the secondary current and the transformer off-nominal turn ratio. The MVA equality, on the other hand, expresses that the magnitude of the complex power flowing through each branch equals the product of the corresponding current and voltage. Similarly, bus summing states that the summation of the real and reactive powers around a bus is equal to zero.

The RMC method calibrates one station at a time, thus avoiding the problems associated with the time skewing of data. An important feature of Adibi's approach is that it does not make use of any reference to calibration. However, the average of the voltages, obtained at each snapshot of the system and calculated after the rejection of the outliers, are implicitly taken as references, and subsequently used for estimating the calibration coefficients of the remaining measurement points. Some of the drawbacks of this method are

i) The average combined with some outlier rejection rules as an estimator of location is prone to the masking effect of multiple non-calibrated metered values [5]. Indeed, the bias of this estimator may be carried beyond all bounds by the action of multiple conforming outliers;

ii) The least square estimator that is suggested to estimate the calibration coefficients is not robust. Here, robust regression methods have to be used because bad data, including bad leverage points, are likely to occur;

iii) The regression model based on the MVA equalities is either non-observable (singular Jacobian matrix) or has multiple solutions.

These weaknesses call for further research and development of a better calibration method. To achieve this goal, a new approach has been developed. It will be outlined next.

1.3 The Proposed Approach

The objective is to develop a Robust Remote Measurement Calibration (RRMC) method that will be implemented at the control center of a system observed by a redundant collection of Remote Terminal Units (RTUs). The method assumes that the system is balanced, has known parameters, and has known topology. Two complementary RRMC methods have been developed. One method involves a minimal field calibration to calibrate all the measurements of the system and the other one is based on the estimation of the meter bias values when the branches are disconnected.

The minimal field calibration procedure consists of the following three steps. The first step finds the minimal set of RTUs subject to system observability. They will be field calibrated. Field calibration ensures that the metered values of these RTUs have small systematic errors, although they may still carry random errors, which may be large. They will be taken as reference points. Their number typically ranges from 10% to 20% of the total number of the units for an extra-high voltage network provided with a large percentage of zero injections (perfect values). Various techniques can be used to find the minimum number of RTUs subject to system observability. Because the objective function has many local minima, we resorted to a random search technique to find the global minimum. The minimal RTU selection is achieved by means of a simulated annealing [20-25] procedure, coupled with a bisecting search method and a topological observability algorithm. Due to the large number of run cases that need to be carried out, this technique is relatively slow. However, simulation results showed that the computing times can be drastically reduced by properly tuning the parameters involved in the method. To further speed up the algorithm and to make it applicable to large systems, system decomposition has been used. It is based on Bellman's Optimality Principle (Dynamic Programming) [42-48] along with some logical decisions. The objective is to tear the system into a given number of smaller subsystems. These subsystems are found by minimizing the number of cut branches and re-connecting subsystems to each other while keeping the number of nodes in each of them under a pre-specified limit. Therefore, simulated annealing is run on smaller sub-systems, and the RTU selection for all the individual sub-systems are

combined to form the final RTU set. By processing a non-redundant set of calibrated measurements that are selected again under the constraint of system observability, the second step calculates, by means of an iterative algorithm, the state of the system along with all the power flows and power injections. Finally, from historical data recorded for each non-field calibrated measurement, the third step estimates a quadratic calibration curve by minimizing a robust criterion, namely the median of the squared residuals. This estimator, known as the *Least Median of Squares* (LMS) [2-5, 39] estimator, is proven to be able to handle a large fraction of bad data.

As for voltage calibration, a new method called *Comparative Voltage Calibration* (CVC) has been developed. This method assumes that the measurement biases are constant for the processed measurements. It consists of two steps. First, it identifies, by means of a robust statistical test developed by us, pairs of voltage measurements whose metered values are close to each other over a large variation of system operating conditions. Then, it uses these calibrated pairs as a reference to estimate the calibration coefficients of the remaining voltage measurements. The LMS estimator is used again for this task. The statistical test that we developed is a new version of the Friedman Rank Sum test [31, 33-34, 49-51], which is shown to be robust. The proposed modification consists of adding Gaussian random noise to the metered values so that the test checks whether these values are within a certain tolerance. Also, the expression of the resistance to rejection [52-55] of this test has been derived and tested.

An alternative method to field calibration is to estimate at the control center the biases of the measurements from metered values recorded when the associated bus or branch is disconnected. The biases being known, the model proposed by Adibi [6-9] becomes observable, which allows us to estimate the linear and non-linear coefficients of the measurements located in both sides of a transformer. This further decreases the number of RTUs needed to be field calibrated.

1.4 Outline of the Dissertation

The objective of this dissertation is to present a reliable approach for RTU calibrations. Chapter 2 reviews and assesses the Remote Measurement Calibration method introduced by Adibi *et al.* [6-9]. Chapter 3 describes a Robust Remote Measurement Calibration (RRMC) algorithm for minimal field calibration. The new algorithm takes advantage of concepts such as simulated annealing, topological observability analysis and robust regression techniques. The developed algorithm is tested on different systems and the results are presented. Chapter 4 illustrates an algorithm for system decomposition using Dynamic Programming [42-48]. Advantages of this algorithm over an already existing one is studied and the results are analyzied. Chapter 5 describes an algorithm for Comparative Voltage Calibration; it goes over few statistical Multiple Comparison tests and explores the resistance to rejection of the Friedman Rank Sum test [33, 34, 49-55]. Chapter 6 proposes a complementary method to RRMC field calibration for measurements around the transformers in the system. The proposed method is tested by means of simulated data.

CHAPTER 2

STATE-OF-THE-ART

2.1 Current State of Knowledge

The RMC Method proposed by Adibi *et al.* [6-9] is a technique that makes use of repetitive readings of the same measurement at the control center. Its main function is to estimate measurement calibration coefficients on the basis of KV, Ampere, and MVA equalities, and Bus Summing. The Ampere equality and Bus Summing neglect the current and power losses in the magnetizing branch of the transformers. Moreover, the measurement calibration is executed one station at a time so as to avoid time-skew.

2.2 Description of the RMC Method

The RMC method is a parameter identification of the model relating the calibrated to the metered voltages, currents and powers. This model is given by;

$$V_c = a_1 + b_1 V_m \,, \tag{1}$$

$$I_c = a_2 + b_2 I_m + c_2 I_m^2 \,, \tag{2}$$

$$P_c = a_3 + b_3 P_m + c_3 P_m^2 \,, \tag{3}$$

$$Q_c = a_4 + b_4 Q_m + c_4 Q_m^2 \,, \tag{4}$$

where the parameters a_i, b_i and c_i are referred to as the off-set, gain, and non-linearity coefficients, respectively. Here a capital letter with subscript $\underline{c}$ stands for a calibrated quantity, whereas a capital letter with subscript $\underline{m}$ stands for a metered quantity.

The calibrated voltages, currents, and powers on branches connecting the same bus, say the ith bus, to another bus, say the jth or the kth bus, satisfy Ohm's and Kirchhoff's laws, yielding the following equalities:

--KV equality: $$V_{ij} = V_{ik} \ , \tag{5}$$

--Ampere equality: $$\sum_j I_{ij} = 0 \ , \tag{6}$$

--Bus Summing: $$\sum_j P_{ij} = 0 \ , \tag{7}$$

$$\sum_j Q_{ij} = 0 \ , \tag{8}$$

--MVA equality: $$\left(P_{ij}^2 + Q_{ij}^2 \right)^{\frac{1}{2}} - V_{ij} I_{ij} = 0 \ . \tag{9}$$

To estimate the calibration coefficients, two distinct RMC approaches have been suggested. The first approach makes use of all the foregoing equalities, and hence assumes the availability of current measurements [6-7]. However, generally currents are monitored only at the substation level, and this situation is unlikely to change in the future. To overcome this problem, a second approach that does not assume any current measurements has been proposed [8-9]. This approach is implemented through a two-step method. First, the voltages are calibrated on the basis of the KV equality. Then, the sum of the real powers and the reactive powers, both in the primary and the secondary sides of the transformers having a turn ratio a, are calibrated by making use of the Ampere equality,

$$a I_{cp} = I_{cs} \ , \tag{10}$$

and the MVA equalities,

$$\left(P_{cp}^2 + Q_{cp}^2\right)^{\frac{1}{2}} = V_{cp} I_{cp} \,, \tag{11}$$

$$\left(P_{cs}^2 + Q_{cs}^2\right)^{\frac{1}{2}} = V_{cs} I_{cs} \,, \tag{12}$$

where the subscripts p and s stand for primary and secondary, respectively. Taking the ratio between (11) and (12) yields

$$\frac{V_{cp}}{V_{cs}} = \mathrm{a}\, \frac{\left(P_{cp}^2 + Q_{cp}^2\right)^{\frac{1}{2}}}{\left(P_{cs}^2 + Q_{cs}^2\right)^{\frac{1}{2}}} \,. \tag{13}$$

Note that here the magnetizing current of the transformer is neglected.

The estimation is performed as follows. Given *m* calibrated voltage values of V_{cp} and V_{cs} obtained at *m* snapshots, one estimates the calibrated coefficients *(a, b, c)* of P_{cp} *and* Q_{cp}. This is a typical multiple regression problem. Adibi *et al.* [6-9] estimate the coefficients in the least squares sense by minimizing a quadratic function given by

$$\sum_{k=1}^{m} \left[\frac{V_{cp}}{V_{cs}} - \mathrm{a}\, \frac{\left(P_{cp}^2 + Q_{cp}^2\right)^{\frac{1}{2}}}{\left(P_{cs}^2 + Q_{cs}^2\right)^{\frac{1}{2}}} \right]^2 \,. \tag{14}$$

Let us illustrate the RMC procedure through an example. Consider a substation that consists of two auto-transformers connecting three high-voltage lines to four low-voltage lines as shown in Figure 2. The RMC algorithm comprises three main steps.

Step 1: *Calibration of the primary and secondary voltages.*

On the basis of the voltage equality (5), the calibration coefficients of the primary and secondary voltages are estimated by processing the *m* sample means (or the sample medians) of the voltage metered values obtained at *m* different snapshots . This is a simple regression in which the sample means or the sample medians are used as response variables. In this parameter identification, the voltage measurements that show up as outliers with large standardized residuals are rejected.

Step 2: *Calibration of the sum of the real and reactive powers.*

On the basis of the Ampere equality (6), the calibration coefficients of the sums of the powers of the primary and the secondary sides of the transformers are estimated by minimizing the objective function (14).

Step 3: *Calibration of the powers on the lines and transformers.*

By making use of the Bus Summing (7),(8) and the calibrated sums of the powers obtained in step 2, the calibration of the real and reactive power measurements on the lines and transformers are estimated.

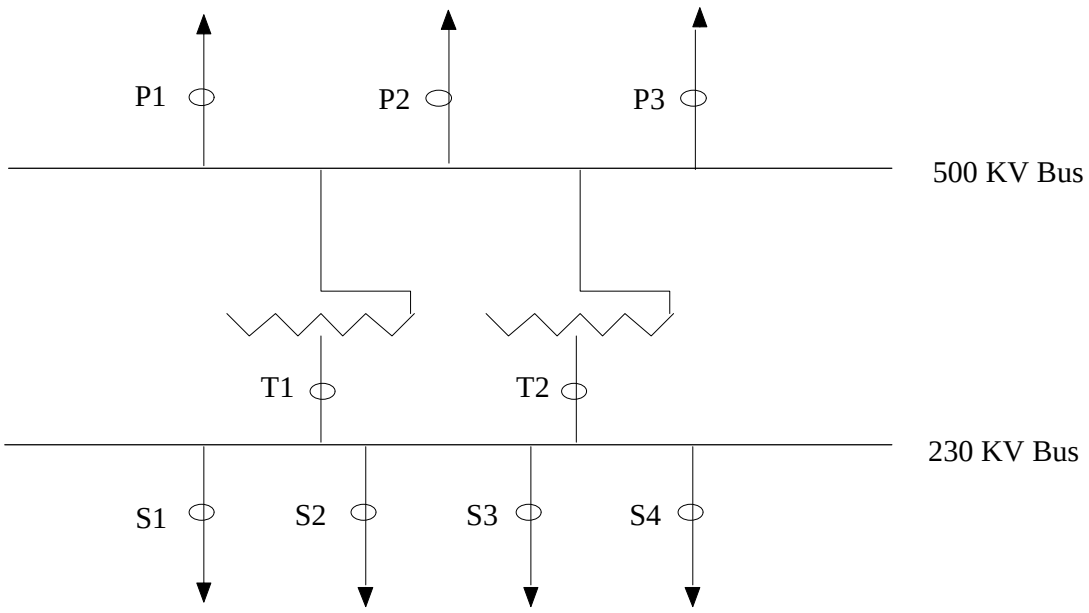

Figure 2. One-line Diagram of the Transmission Substation.

Now we show that this procedure has one major drawback, which is the non-satisfaction of the observability requirement of the model. We assume that Step 1 provides a calibrated voltage for both sides of the transformer. Equation (13) yields

$$\left(\frac{V_{cp}}{\mathrm{a}V_{cs}}\right)^2 = \frac{P_{cp}^2 + Q_{cp}^2}{P_{cs}^2 + Q_{cs}^2} \ . \tag{15}$$

Substituting (3) and (4) into (15) gives

$$Z = \frac{\left(a_{ip} + b_{ip}P_{mp} + c_{ip}P_{mp}^2\right)^2 + \left(a_{jp} + b_{jp}Q_{mp} + c_{jp}Q_{mp}^2\right)^2}{\left(a_{is} + b_{is}P_{ms} + c_{is}P_{ms}^2\right)^2 + \left(a_{js} + b_{js}Q_{ms} + c_{js}Q_{ms}^2\right)^2} \ , \tag{16}$$

where Z = $\left(\frac{V_{cp}}{\mathrm{a}V_{cs}}\right)^2$ can be regarded as a function of 12 parameters contained in $\underline{x} = \left[a_{ip}, b_{ip}, c_{ip}, ..., a_{js}, b_{js}, c_{js}\right]$,yielding $\underline{z} = \underline{h}(\underline{x})$. Suppose that m value of Z are calculated from m snapshots. Let Z_k be the value of Z obtained at the kth snapshot. Assume that we have m snapshots and let $\underline{z} = \left[Z_1, Z_2, ..., Z_m\right]^T = \underline{h}(\underline{x})$. A first order Taylor series expansion of $\underline{z}$ about an operating point yields

$$\Delta\underline{z} = \underline{\underline{H}}\Delta\underline{x} \ , \tag{17}$$

where $\underline{\underline{H}} = \frac{\partial\, \underline{h}(\underline{x})}{\partial\, \underline{x}}$ is the $(m \times 12)$ Jacobian matrix given by

$$\underline{\underline{H}} = \left[\frac{\partial Z}{\partial a_{ip}}, \frac{\partial Z}{\partial b_{ip}}, \frac{\partial Z}{\partial c_{ip}}, \frac{\partial Z}{\partial a_{jp}}, \frac{\partial Z}{\partial b_{jp}}, \frac{\partial Z}{\partial c_{jp}}, \frac{\partial Z}{\partial a_{is}}, \frac{\partial Z}{\partial b_{is}}, \frac{\partial Z}{\partial c_{is}}, \frac{\partial Z}{\partial a_{js}}, \frac{\partial Z}{\partial b_{js}}, \frac{\partial Z}{\partial c_{js}}\right], \tag{18}$$

and where $\Delta\underline{x}$ is a *12* dimensional vector defined by

$$\Delta\underline{x} = \left[\Delta a_{ip}, \Delta b_{ip}, \Delta c_{ip}, \Delta a_{jp}, \Delta b_{jp}, \Delta c_{jp}, \Delta a_{is}, \Delta b_{is}, \Delta c_{is}, \Delta a_{js}, \Delta b_{js}, \Delta c_{js}\right]^T. \quad (19)$$

Let us derive the expression of the partial derivatives of *Z*. From (17) we get

$$\frac{\partial Z}{\partial a_{ip}} = \frac{2\left(a_{ip} + b_{ip}P_{mp} + c_{ip}P_{mp}^2\right)}{B}, \quad (20)$$

$$\frac{\partial Z}{\partial b_{ip}} = \frac{2P_{mp}\left(a_{ip} + b_{ip}P_{mp} + c_{ip}P_{mp}^2\right)}{B}, \quad (21)$$

$$\frac{\partial Z}{\partial c_{ip}} = \frac{2P_{mp}^2\left(a_{ip} + b_{ip}P_{mp} + c_{ip}P_{mp}^2\right)}{B}, \quad (22)$$

$$\frac{\partial Z}{\partial a_{is}} = \frac{-2\left(a_{is} + b_{is}P_{ms} + c_{is}P_{ms}^2\right)A}{B^2}, \quad (23)$$

$$\frac{\partial Z}{\partial b_{is}} = \frac{-2P_{ms}\left(a_{is} + b_{is}P_{ms} + c_{is}P_{ms}^2\right)A}{B^2}, \quad (24)$$

$$\frac{\partial Z}{\partial c_{is}} = \frac{-2P_{ms}^2\left(a_{is} + b_{is}P_{ms} + c_{is}P_{ms}^2\right)A}{B^2}, \quad (25)$$

where

$$A = \left(a_{ip} + b_{ip}P_{mp} + c_{ip}P_{mp}^2\right)^2 + \left(a_{jp} + b_{jp}Q_{mp} + c_{jp}Q_{mp}^2\right)^2, \quad (26)$$

$$B = \left(a_{is} + b_{is} P_{ms} + c_{is} P_{ms}^2\right)^2 + \left(a_{js} + b_{js} Q_{ms} + c_{js} Q_{ms}^2\right)^2 . \quad (27)$$

The partial derivatives of Z with respect to a_{jp}, b_{jp}, c_{jp}, a_{js}, b_{js}, and c_{js} are given by the same expression as (20) through (25) except that P is changed to Q and i is changed to j.

2.3 Observability Analysis

We will show that the Jacobian matrix $\underline{H}$ can be rank deficient. To do that, we will show that at least two of its columns can be linearly dependent. For convenience, we will consider the case where the off-set and non-linearity terms are zero , that is $a = c = 0$. This yields

$$Z = \frac{\left(b_1 P_{mp}\right)^2 + \left(b_2 Q_{mp}\right)^2}{\left(b_3 P_{ms}\right)^2 + \left(b_4 Q_{ms}\right)^2} \quad . \quad (28)$$

The partial derivatives of Z are then given by

$$\frac{\partial Z}{\partial b_1} = \frac{2 p_{mp}^2 b_1}{D} , \quad (29)$$

$$\frac{\partial Z}{\partial b_2} = \frac{2 q_{mp}^2 b_2}{D} , \quad (30)$$

$$\frac{\partial Z}{\partial b_3} = \frac{-2 p_{ms}^2 b_3 C}{D^2} , \quad (31)$$

$$\frac{\partial Z}{\partial b_4} = \frac{-2q_{ms}^2 b_4 C}{D^2}\,, \tag{32}$$

$$\frac{\partial Z}{\partial b_1} + \frac{\partial Z}{\partial b_2} = \frac{2b_1 p_{mp}^2 D + 2b_2 q_{mp}^2 D}{D^2}\,, \tag{33}$$

$$\frac{\partial Z}{\partial b_3} + \frac{\partial Z}{\partial b_4} = \frac{-2C\left(b_3 p_{ms}^2 + b_4 q_{ms}^2\right)}{D^2}\,, \tag{34}$$

where

$$C = \left(b_1 p_{mp}\right)^2 + \left(b_2 q_{mp}\right)^2\,, \tag{35}$$

and

$$D = \left(b_3 p_{ms}\right)^2 + \left(b_4 q_{ms}\right)^2\,. \tag{36}$$

Let us now show that there is a linear dependency between the columns of $\underline{H}$. In other words, we want to show that there exist non-zero coefficients $\lambda_1, \lambda_2, \lambda_3$, and, λ_4 so that

$$\lambda_1 \frac{\partial Z}{\partial b_1} + \lambda_2 \frac{\partial Z}{\partial b_2} + \lambda_3 \frac{\partial Z}{\partial b_3} + \lambda_4 \frac{\partial Z}{\partial b_4} = 0\,. \tag{37}$$

Let

$$\boldsymbol{E} = \lambda_1 \frac{\partial Z}{\partial b_1} + \lambda_2 \frac{\partial Z}{\partial b_2} + \lambda_3 \frac{\partial Z}{\partial b_3} + \lambda_4 \frac{\partial Z}{\partial b_4}\,. \tag{38}$$

Substituting (29) through (32) into (38) we get $E = \frac{F}{D^2}$, where

$$F = 2P_{mp}^2 P_{ms}^2 b_1 b_3 (b_1 \mathsf{l}_1 - b_3 \mathsf{l}_3) + 2P_{mp}^2 Q_{ms}^2 b_1 b_4 (b_4 \mathsf{l}_1 - b_1 \mathsf{l}_3)$$
$$+2Q_{mp}^2 P_{ms}^2 b_2 b_3 (b_3 \mathsf{l}_2 - b_2 \mathsf{l}_4) + 2Q_{mp}^2 Q_{ms}^2 b_2 b_4 (b_4 \mathsf{l}_2 - b_2 \mathsf{l}_4). \quad (39)$$

F is equal to zero if we have

$$\begin{aligned} \mathsf{l}_1 b_1 - \mathsf{l}_3 b_3 &= 0, \\ \mathsf{l}_1 b_4 - \mathsf{l}_3 b_1 &= 0, \\ \mathsf{l}_2 b_3 - \mathsf{l}_4 b_2 &= 0, \\ \mathsf{l}_2 b_4 - \mathsf{l}_4 b_2 &= 0. \end{aligned} \quad (40)$$

Solving for $\mathsf{l}_1, \mathsf{l}_2, \mathsf{l}_3, \mathsf{l}_4$, we get three solutions that are different from the trivial solution $\mathsf{l}_1 = \mathsf{l}_2 = \mathsf{l}_3 = \mathsf{l}_4 = 0$. These non-trivial solutions are given by

$$\mathsf{l}_1 \neq 0, \mathsf{l}_3 \neq 0, \mathsf{l}_2 = 0, \mathsf{l}_4 = 0, \qquad \text{for } b_1^2 = b_3 b_4 \text{ or } b_1 = \sqrt{b_3 b_4},$$
$$\mathsf{l}_1 = 0, \mathsf{l}_2 \neq 0, \mathsf{l}_3 = 0, \mathsf{l}_4 \neq 0, \qquad \text{for } b_3 = b_4.$$

Consequently, the values of $a_i = 0, b_i = 1$ can not be chosen as initial conditions for the iterative procedure because the Jacobian matrix is singular. Even if the initial condition is chosen so that the Jacobian matrix associated with them is non-singular, the calculated values for the b's may become close to each other as the iterations are progressing, which makes the Jacobian matrix ill-condition.

To further investigate the matter, let us consider the case where the non-linear terms are zero, that is $c = 0$. This yields

$$Z = \frac{\left(a_1 + b_1 P_{mp}\right)^2 + \left(a_2 + b_2 Q_{mp}\right)^2}{\left(a_3 + b_3 P_{ms}\right)^2 + \left(a_4 + b_4 Q_{ms}\right)^2} \quad . \tag{41}$$

Let us show that there is a linear dependency between the columns of $\underline{H}$. In other words, we need to show that there exist non-zero coefficients $\{l_1, l_2, ..., l_8\}$ so that

$$l_1 \frac{\partial Z}{\partial a_1} + l_2 \frac{\partial Z}{\partial b_1} + l_3 \frac{\partial Z}{\partial a_2} + l_4 \frac{\partial Z}{\partial b_2} + l_5 \frac{\partial Z}{\partial a_3} + l_6 \frac{\partial Z}{\partial b_3} + l_7 \frac{\partial Z}{\partial a_4} + l_8 \frac{\partial Z}{\partial b_4} = 0 \quad . \tag{42}$$

Due to the non-orthogonality of the columns of the Jacobian matrix, it is found that the problem of showing analytically that these columns are linearly dependent is intractable. Therefore, we will try to show it numerically by investigating several cases.

2.4 Numerical Results

The rank of the Jacobian matrix has been determined in two cases with the following true values for the a's and b's:

$$a_1 = .01, \ a_2 = .02, \ a_3 = .04, \ a_4 = .03$$
$$b_1 = .95, \ b_2 = .96, \ b_3 = .98, \ b_4 = .95.$$

The initial conditions for both cases are chosen to be

$$a_1 = .1, \ a_2 = .02, \ a_3 = .5, \ a_4 = .4$$
$$b_1 = .95, \ b_2 = .9, \ b_3 = 1.1, \ b_4 = .97.$$

The values for P_m and Q_m are given in Tables 1 and 2.

Table 1. P_m and Q_m Values for Case 1.

P_{mp}	Q_{mp}	P_{ms}	Q_{ms}
0.2190	0.0346	0.6868	0.7012
0.0470	0.0535	0.5890	0.9103
0.6789	0.5297	0.9304	0.7622
0.6793	0.6711	0.8462	0.2625
0.9347	0.0077	0.5269	0.0475
0.3835	0.3834	0.0920	0.7361
0.5194	0.0668	0.6539	0.3282
0.8310	0.4175	0.4160	0.6326

Table 2. P_m and Q_m Values for Case 2.

P_{mp}	Q_{mp}	P_{ms}	Q_{ms}
-0.5650	0.8070	-0.2480	0.7130
0.9970	-0.5750	0.7150	-0.4100
0.9430	0.2420	0.4610	0.4330
-0.6420	0.4560	-0.3510	0.2160
0.2920	-0.5100	0.7690	-0.9700
0.5499	0.8208	0.2867	0.7965
0.9671	0.5852	0.7714	0.4694
0.2894	0.8366	-0.3948	0.6285

They range between 0 and 1 in case 1 and between -1 and 1 in case 2. They have been randomly obtained using the Personal Computer internal clock as the random number generator seed provider. The Jacobian matrices for both cases have been calculated by using equations (20) through (27) while putting $c_i = 0$. For case 1, it is given by

$$\underline{\underline{H}} = \begin{bmatrix} 0.2246 & 0.0492 & 0.1685 & 0.0058 & -0.0495 & -0.0340 & -0.0426 & -0.0299 \\ 0.0976 & 0.0046 & 0.1675 & 0.0090 & -0.0216 & -0.0127 & -0.0241 & -0.0219 \\ 0.4117 & 0.2795 & 0.3740 & 0.1981 & -0.2356 & -0.2192 & -0.1762 & -0.1343 \\ 0.6021 & 0.4090 & 0.6495 & 0.4359 & -0.5611 & -0.4748 & -0.2567 & -0.0674 \\ 1.4481 & 1.3535 & 0.3033 & 0.0023 & -1.1816 & -0.6226 & -0.4882 & -0.0232 \\ 0.5795 & 0.2222 & 0.6803 & 0.2608 & -0.2401 & -0.0221 & -0.4448 & -0.3274 \\ 0.5926 & 0.3078 & 0.2598 & 0.0174 & -0.2553 & -0.1669 & -0.1504 & -0.0494 \\ 0.9149 & 0.7603 & 0.5922 & 0.2472 & -0.5687 & -0.2366 & -0.6019 & -0.3808 \end{bmatrix} \quad (43),$$

whereas for case 2, it is given by

$$\underline{\underline{H}} = \begin{bmatrix} -0.7026 & 0.3970 & 1.4901 & 1.2025 & -0.3083 & 0.0765 & -1.4814 & -1.0563 \\ 1.2654 & 1.2616 & -0.3837 & 0.2206 & -1.1246 & -0.8041 & -0.0020 & 0.0008 \\ 1.1808 & 1.1135 & 0.4954 & 0.1199 & -0.8257 & -0.3807 & -0.6723 & -0.2911 \\ -2.6524 & 1.7028 & 3.1751 & 1.4479 & -0.9748 & 0.3421 & -5.2164 & -1.1267 \\ 0.3587 & 0.1048 & -0.2462 & 0.1256 & -0.1274 & -0.0980 & 0.0512 & -0.0497 \\ 0.6103 & 0.3356 & 0.9204 & 0.7555 & -0.4972 & -0.1425 & -0.7150 & -0.5695 \\ 0.7990 & 0.7727 & 0.5699 & 0.3335 & -0.6494 & -0.5010 & -0.4119 & -0.1933 \\ 0.7329 & 0.2121 & 1.8627 & 1.5583 & -0.1237 & 0.0493 & -2.0226 & -1.2712 \end{bmatrix} \quad (44).$$

These foregoing matrices are singular, and have zero determinants. They have a rank of 7 instead of 8. To see that, let us show that the last vector column $\underline{b}$ of the Jacobian matrix can be expressed as a linear combination of the seven first columns. Let $\underline{S}$ be the (7×7) dimensional sub-matrix consisting of the first 7 rows and columns of $\underline{\underline{H}}$ given by either (43) or (44). Solving for $\underline{l}$ in

$$\underline{S}\underline{l} = \underline{b}\ , \tag{45}$$

yields

$$\underline{l} = [-0.1031\ \ -0.9794\ \ -0.2062\ \ -0.9278\ \ -0.5155\ \ -1.1340\ \ -0.4124]^{T}\ ,$$

a solution that is different from zero. The same vector $\underline{l}$ is found when considering the Jacobian matrix of case 2.

Repeating the whole procedure while using the P_m and Q_m values given in Table 1 and the initial conditions given by

$$a_1 = .04,\ \ a_2 = .03,\ \ a_3 = .02,\ \ a_4 = .01$$
$$b_1 = .95,\ \ b_2 = .96,\ \ b_3 = .97,\ \ b_4 = .98,$$

we get a solution for (45) that is expressed as

$$\underline{l} = [-0.0408\ \ -0.9694\ \ -0.0306\ \ -0.9796\ \ -0.0204\ \ -0.9898\ \ -0.0102]^{T}.$$

In all cases that have been considered, the Jacobian matrix is found to be singular regardless of the values of the initial conditions and those of the $P_m s$ and $Q_m s$. In addition, it is found that the vector $\underline{l}$ that satisfies the system of equations (45) depends only on the initial conditions and not on the values of the $P_m s$ and $Q_m s$. It has also been observed that the Jacobian matrix has a rank deficiency of one. If one of the parameters a or b is assumed to be known and the associated partial derivatives deleted, then the ($mx7$) Jacobian matrix becomes non-singular and well defined. However, in that case, another problem arises; it is the existence of multiple solutions. To overcome this deficiencies alternative approaches have been developed they are described in the following chapters.

CHAPTER 3

MINIMAL FIELD CALIBRATION METHOD

3.1 Motivation and Outline of the Approach

As it has been stated in previous chapters, one way to obtain the proper measurements needed to run the state estimation is to make use of the RMC method. An example of a calibration at the substation level using the RMC method has been analyzed in Chapter 2. It is found that this method requires field calibration of at least one power measurement on each RTU in order for the model to be observable. However, if an RTU is disconnected for calibration, it makes sense to calibrate all its measurement points. Hence, no remote calibration needs to be performed when using the RMC method. Does any other method exist that allows us to calibrate only a fraction of the RTUs? Fortunately, the answer is positive, as it will be explained in the following chapter.

To solve this problem, a two-step method will be developed. The first step consists of minimizing the number of RTUs that need to be field-calibrated in order to observe the whole system. Observing the system means that the whole state vector can be calculated. The second step consists of remotely calibrating, at the control center, the rest of the RTUs through a robust estimation method. The minimum number of RTUs for field calibration will be found through a dual-search method. It consists of finding an initial set of RTUs that observe the system and then reducing the size of this RTU set by using simulated annealing. A few assumptions in the system have been made. We assume that (i) the topology of the system is known, (ii) all the parameters are known, and (iii) the system is balanced. These assumptions guarantee that the one-line diagram of the network is a good model of the system.

3.2 Minimum Number of RTUs for System Observability

3.2.1 Initial Selection Set of RTUs

The flow chart given in Figure 3 summarizes the procedure for finding the initial set of RTUs that make the system observable. The initial guess is given by a graph theoretic procedure that builds a spanning tree of full rank across the system [10]. This tree connects all the nodes through branches that are each assigned a calibrated metered value. The procedure extends a tree, step-by-step, from the observable part of the system into the remaining unobservable parts. The procedure starts from a bus with the largest number of branches on which RTUs exist. Then it moves from one node to an adjacent one according to a depth-first search methodology [11]. At each step, the observable part of the system is expanded by selecting RTUs that are within the unobservable region, adjacent to the observed sub-network. When the tree cannot be expanded any further along a given path, another path is explored by moving back one node at a time. At the completion of the procedure the number m of RTUs in the selected set is counted. By construction, this RTU set observes the system. It is used as an initial guess by the simulated annealing-based procedure in an attempt to decrease its size. The method will be described next.

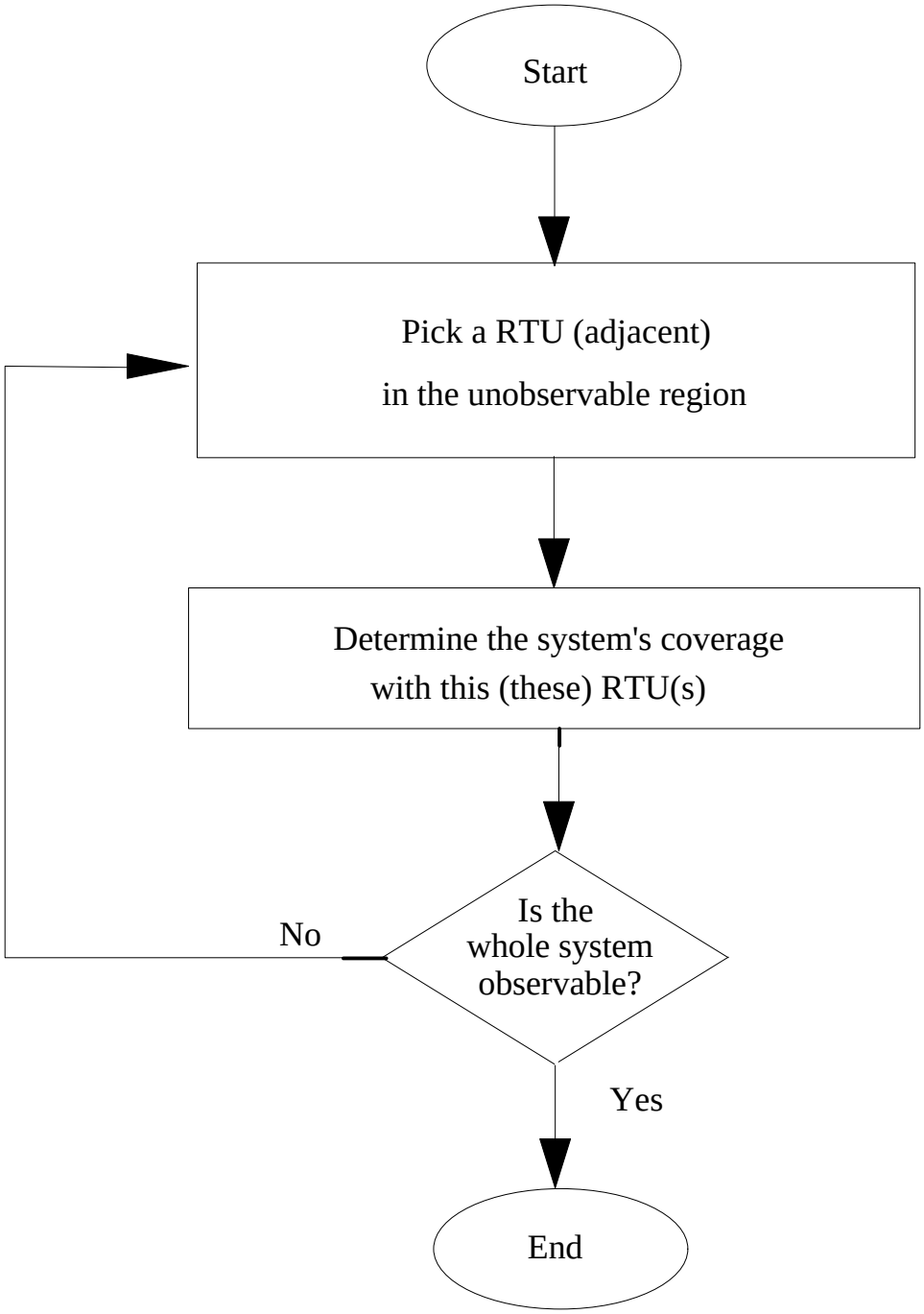

Figure 3. Flow Chart for Initial Selection Set of RTUs

3.2.2 Reducing the Size of the Initial Selection of the RTU Set

The RTU set is minimized through an algorithm that makes use of the bisecting search technique and the simulated annealing method [20-25]. It consists of five main steps, which are

Step 1: Initialize $i = 0$, set $m' = 1$, and $m'' = m$.

Step 2: Set $i = i + 1$. Let m_i be the mid-point of the interval [m', m''].

Step 3: Consider the m_i as the first numbered RTUs obtained from the graph theoretic procedure. By construction, the system is not observable for this subset.

Step 4: Execute the simulated annealing-based procedure. This procedure carries out random selections of m_i RTUs in an attempt to form an observable system.

Step 5: If the procedure is successful, then consider the lower half by setting $m'' = m_i$, otherwise consider the upper half by setting $m'=m_i$. Go to step 2.

The number of re-selections required by this procedure is $log_2 m$. Consequently, the computational burden does not increase rapidly with the size of the system [24].

3.2.3 Simulated Annealing Algorithm

The optimization process carried out in Step 4 attempts to maximize the observable region of the system associated with a given number of RTUs. This process makes use of the simulated annealing method [20-25], which is a convenient way to find the global minimum of a function that has many local minima and may not be smooth. The method is a biased random walk that samples the objective function in the space of the independent variables. It is executed in the following manner. At a fixed temperature, the

objective function is sampled according to the Metropolis algorithm. Namely, let the objective function be $E(x_1,x_2,\ldots,x_n)$, which is a function of n independent variables. Choose a convenient maximum step-size, Δr (constant for a given run), which can be taken along each of the n directions, and a starting temperature T. A single annealing stage then proceeds as shown in the Pseudo-code below. The (E,x) found at the higher temperature is used to continue the search (gradually downhill) at a reduced temperature. A brief representation of this method is presented below.

```
Initialize
        x = Starting point in domain
        E1 = E(x)
        T = (starting temperature)
    Loop over N samples
        Y = Random vector of length n. (each entry on interval -1,+1)
        x' = x + Δ rY      (Δ r = maximum allowed step size)
        E2=E(x')
        if  (E2 < E1) then    (accept the new configuration x')
                x = x'
        else     (accept new configuration with probability p)
                z = random number        (0 ≤ z ≤ 1)
                p = exp (-(E2-E1)/T)
                if (p≥z) then
                        x = x'
                endif
        endif
        E1 = E2
    End loop
```

Numerical experience indicates that a kind of adaptive divide and conquer strategy emerges as a result of the annealing process. Downhill steps are taken with a unit probability in such a way that the optimization heads toward local minima, yet the algorithm needs not get trapped in local minima (as do steepest descent searches). The algorithm "jumps out" of these local minima with a probability given by the Boltzman Probability distribution given by

$$Probability(accept\ worse\ case) = \exp\{-[E(x2)-E(x1)]/T\}. \quad (46)$$

It expresses the idea that a system in thermal equilibrium at temperature T has its energy probabilisticly distributed among all the different energy states E. Even at low temperatures, there is a chance for the system to be in a high energy state. Therefore, there is a corresponding chance for the system to get out of a local energy minimum in favor of finding a better, more global one. If the new ΔE value is positive, then (46) is used to generate the probability of accepting the new configuration. As relationship (46) suggests, choosing a worse configuration is a function of temperature. The higher the temperature, the higher the "energy" of the system. The higher the energy of the system the more easily it will move in an unfavorable direction. Simulated annealing automatically adapts itself to the magnitude of the features of a good configuration. The most profitable features of the surface are found first and the least profitable features later. If the temperature is very high, these features may be forgotten from time to time, but as the temperature decreases, eventually they are frozen. Thus, the gross feature of the terrain is sampled at a high temperature, finer details at a lower temperature. Eventually, at a low enough temperature, the system is trapped or "freezes" in a local minimum from which it can escape with small probability.

3.3 Observability Analysis

3.3.1 Algorithm for Analyzing the Observability of the System

The method requires the availability of a system observability algorithm, whose main function is to test whether the whole state vector $\underline{x}$ can be calculated[10-20]. The state vector $\underline{x}$ of an n-bus power system is a $(2n-1)$-dimensional vector that consists of n bus voltage magnitudes and $n-1$ bus voltage angles. The measurement model for static state estimation is given by

$$\underline{z} = \underline{h}(\underline{x}) + \underline{e} \ , \tag{47}$$

where $\underline{z}$ is the $(m \cdot 1)$ dimensional measurement vector, $\underline{h}(\underline{x})$ is the $(m \cdot 1)$ dimensional non-linear vector function, and $\underline{e}$ is the $(m \cdot 1)$ dimensional noise measurement vector. One way to test the observability of a network is to calculate the rank of the $(m \cdot n)$ dimensional Jacobian matrix $\underline{H}$. The computing time of such a method, called numerical observability method, grows with the size of the system and the number of measurements. Another drawback stems from the fact that this method does not allow us to identify the cause of the unobservability of a network, which can be due to either a rank deficiency or ill-conditioning of the Jacobian matrix. Ill-conditioning means at least two column vectors are nearly collinear.

To overcome this problem, we may use instead a topologically-based observability algorithm [10,13-14,19]. This algorithm is based on the graph theory [35]. It makes use of concepts such as graph, branch, path, tree, forest and spanning tree. A *graph* consists of a finite set of *nodes* and unordered pairs of nodes called *branches.* A *path* between two nodes is a sequence of branches connecting them. A *tree* is a path or connected paths that do not make a loop. A *forest* is the union of trees. A *spanning tree* is a tree that contains every node of any graph. Figure 4 is a graphical representation of these concepts.

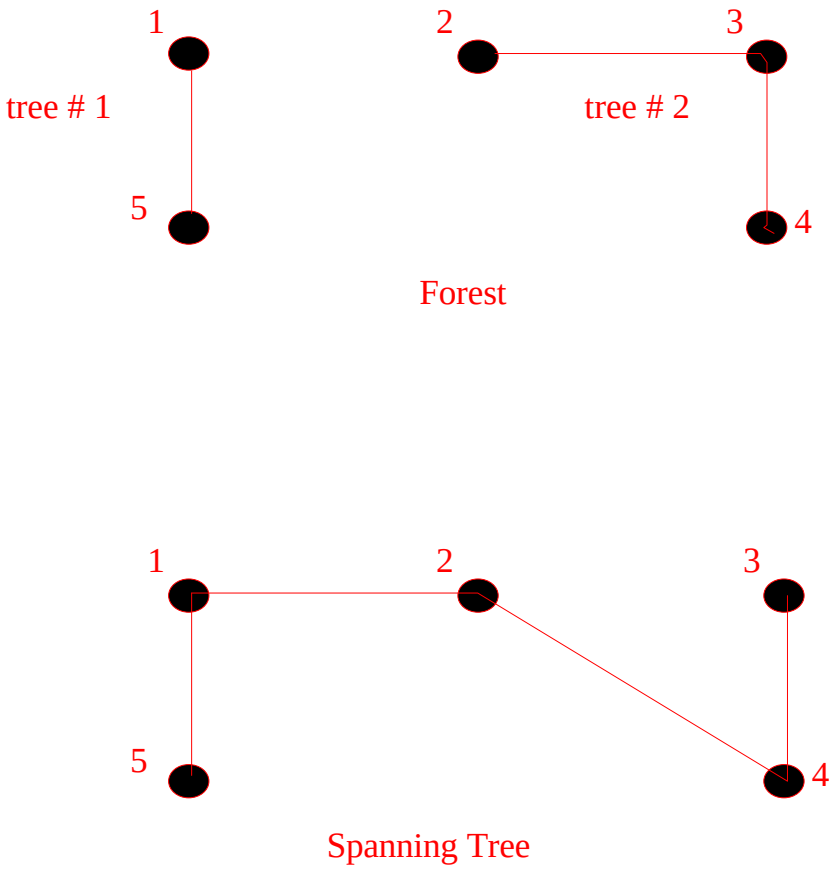

Figure 4. Representation of Graph Theory

Applications of these concepts to power systems are described in [10,13-14,19]. A review of these methods is given next. Krumpholtz, Clements & Davis [10,13-14,19] state that a network is observable if there is at least one tree in the network that has an assigned measurement to each of its branches. A measurement assignment is the association between a branch and the corresponding measurement on that branch. A sub-network is topologically observable if and only if there exists a measurement assigned tree that contains all buses of that sub-network. Therefore, the goal is to find a maximal tree of full rank [10,13-14,19], a spanning tree, in the system by assigning the system measurements to a set of branches using the following rules:

(1) A flow measurement in a branch is assigned to that branch;

(2) An injection measurement at a bus is assigned to one of the incident branches of that bus.

When a spanning tree of full rank exists, we say that the system is topologically observable. The objective here is to determine to which branch each measurement should be assigned resulting in a spanning tree. Also, as it has been stated before, each injection measurement can be assigned to one of several branches. All of these give a lot of possibilities for measurement assignments. One way to obtain a spanning tree is to examine all the possible combinations, which is not recommended due to the large size of power systems. It has been stated in [10] that in the worst case and in a completely connected n-bus network, there are n^{n-2} trees. For example the IEEE 14-bus depicted in Figure 5, has 3,909 trees while the IEEE 30-bus shown in Figure 6, contains 7,823,711 trees. There has been a lot of research done to find a feasible answer to this problem [10,13-19]. One of the most efficient ways to achieve topological observability analysis is to use the augmenting sequence technique [17,18] and the matroid theory [36]. More details on the algorithm and its applications are presented in [18] and [37].

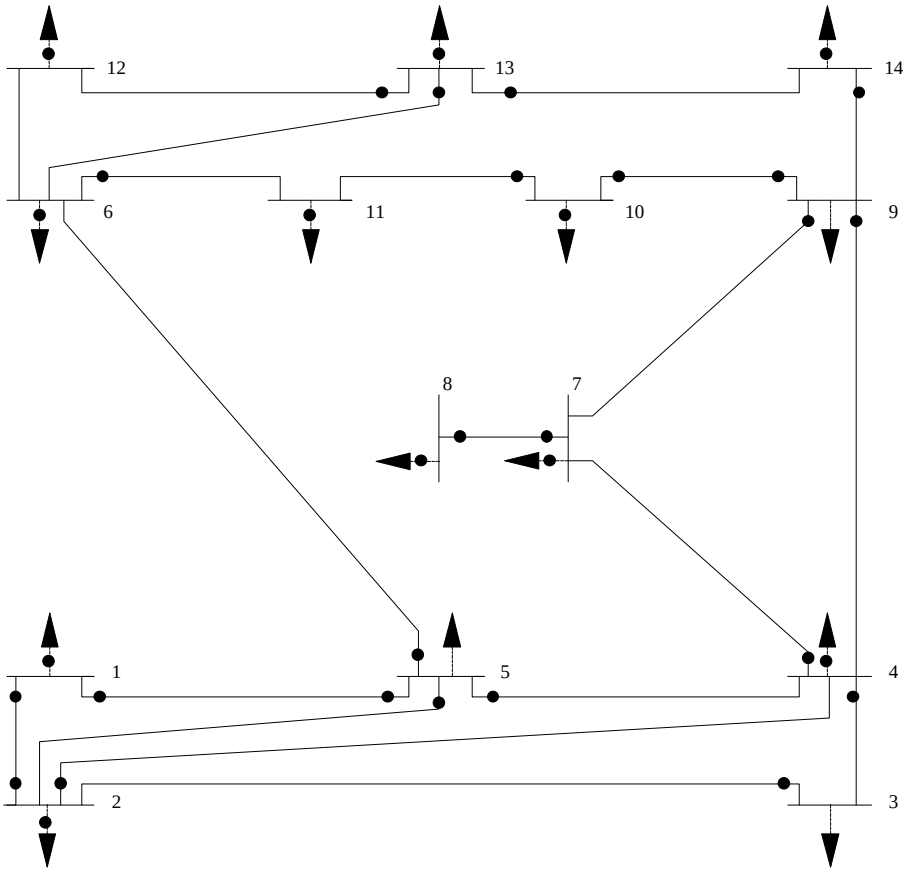

Figure 5. IEEE 14-Bus System

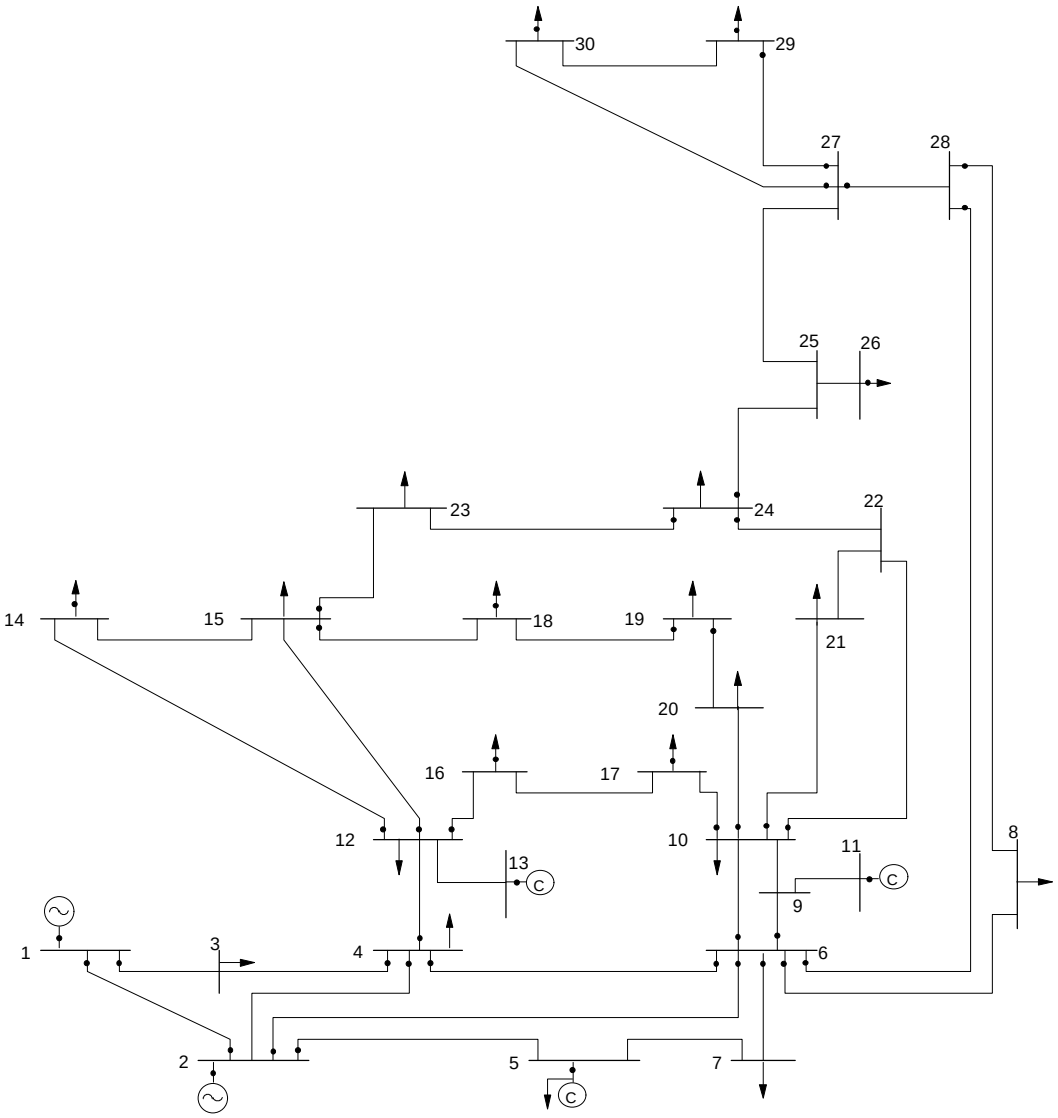

Figure 6. IEEE 30-Bus System

3.3.2 Example of the IEEE 5-Bus System

In this section, an example of topological observability is illustrated in the IEEE 5-bus system shown in Figure 7. In this system, voltage measurement on bus one, power flow measurements on buses one and four, and power injection measurement on bus three are assumed to be calibrated. The power flow on bus one will cover lines connecting buses one to two and one to three. The power injection on bus three can be used as a power flow covering the line connecting buses three to four. The power flow on bus four also covers the line connecting buses four to five. Figure 7 is also a graphical representation of the formation of a spanning tree. As it can be seen, a measurement has been assigned to each branch and, therefore, all the buses have been covered and no loop has been made. Hence, we conclude that the system is observable.

Now we will show that thanks to these measurement points, the whole state of the system can be calculated. The reader is referred to Figure 8. By knowing the voltage at bus one and power flows at nodes 1 and 2, and using basic power equations, the voltages at buses two and three are calculated. Also the power flows at nodes 4 and 8 are obtained. Knowing the voltages at buses two and three, the flows at nodes 5 and 9 are found. By equating the sum of the powers at bus three to zero, we calculate the power flow at node 10. Using the voltage and power flow at node 10, we find the voltage at bus four and the power flow at node 12 in the same manner as we did before. From the power equality at bus four, we obtain the injection at this bus. Similarly, by knowing the voltage at bus four and flow at node 13, we calculate the voltage at bus five and the power flow at node 15. Using the voltages at buses two and five, we calculate the power flows at nodes 6 and 16. Power flow summing equality at buses one, two, and five yields the corresponding power injections. Therefore as it can be seen, by having only 4 measurement points, we are able to calculate all the other electrical quantities of the system.

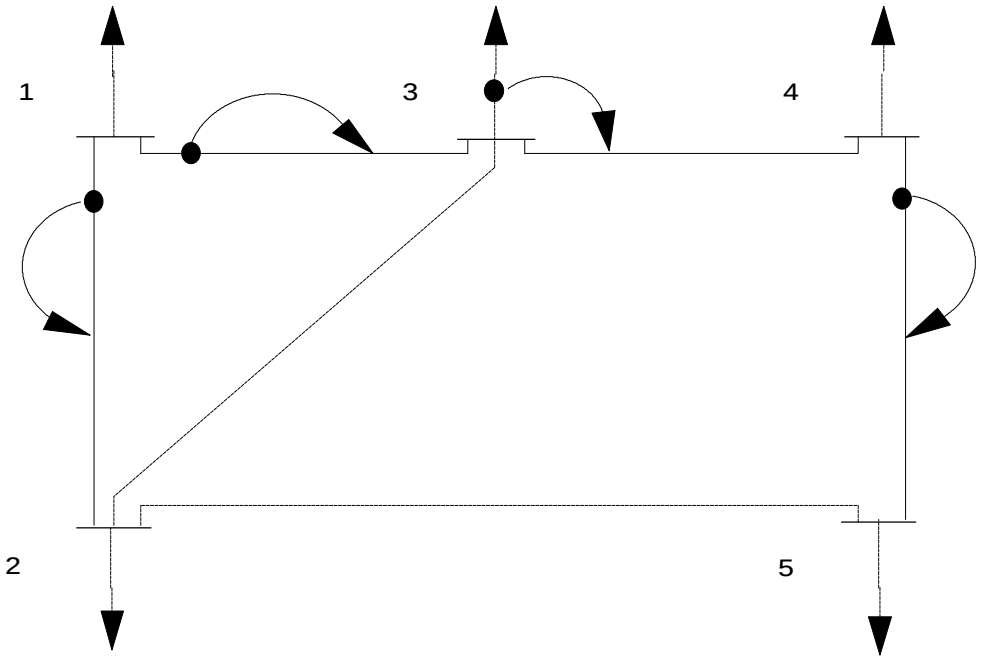

Figure 7. Implementation of Forming a Spanning Tree for IEEE 5-Bus System

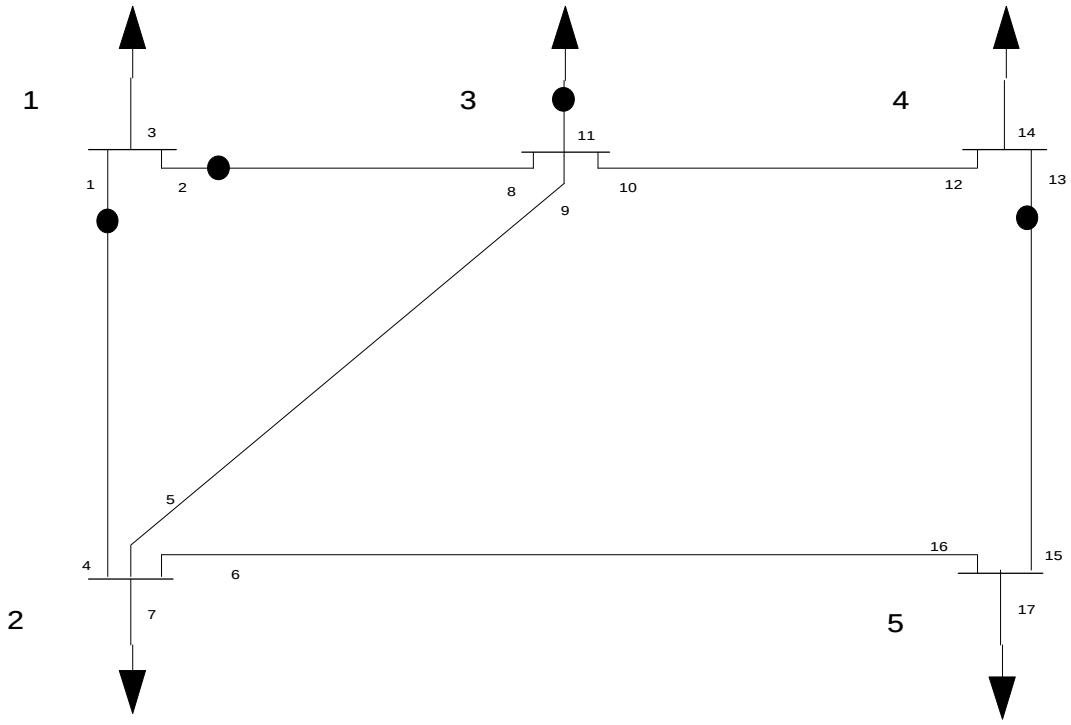

Figure 8. IEEE 5-Bus System With Marked Measurement Points

Hence as seen from the example, using the foregoing algorithm enables us to calculate all the state variables of the system. This information will be exploited in the following section.

3.4 The Applied Algorithm

3.4.1 Statement of the Problem

Simulated Annealing will be used to find the global minimum number of RTU's that observe the system. These minimum number of RTU's are needed to be field calibrated in the system.

3.4.2 The Procedure

The first step is to define the objective function. The objective function in our case is chosen as the number of unobserved buses in the system, and the objective of the program is to minimize this number. The next step is to find the proper range, for our parameters. This is called *parameter tuning*. This is a task of great importance, because by choosing wrong parameter values, the algorithm may not get to the global minimum that we are looking for. In some cases, it may get stuck in a local minimum point, and thereby, may not even converge. The parameters of interest are the initial temperature and the step size. There is no special formulas that can be used to achieve this task, and as observed during the application of the algorithm to different cases, the tuning is found to be very case dependent. In the following paragraph, a general procedure for parameter tuning is outlined.

The first step in tuning is to start from a large temperature with a very slow moving function. This can be achieved by selecting the step size in such a way that the change is small. Let the program run until it gets to the desired reasonable solution of the problem. In doing this, the number of iterations or the running time of the program needs to be recorded. In the next step, by taking this solution as a reference value, start the program again. This time, first reduce the temperature and fix the step size, and second, reset the temperature at the original value, and increase the step size so that the procedure moves faster down the function. Record the results for these steps and, if any improvement is observed, use that as a base case and execute the whole procedure again. The intent here is to vary these parameters so that the efficiency of the algorithm is improved. Sometimes, after a while of working on the problem and not achieving any significant improvement in the results, a change may be necessary in the random generator seed number in order to have a different sequence of numbers.

3.4.3 Application to IEEE 14 and 30 Bus Systems

The above algorithm is applied to the 14- and 30- bus systems. They are depicted in Figures 5 and 6 respectively. The objective function in both cases is the number of unobserved buses in the system. The algorithm starts from some temperature and step size. It first selects a smaller number of RTU's than the total one, and then check if with this number of RTU's the system may be observed. If the answer is positive, then the number of RTU's is reduced further and system observability tested again. If not, then the number of the unobserved buses is determined and recorded for later usage. It then reuses the same RTU's, in terms of number and locations, and changes only one location from the list of available unused RTU's. Now the system should be observed again. If it is not observed then the number of the unobserved buses is determined and compared with the previous one. If it is smaller than the previous one, this means that the objective function is decreasing toward the minimum and, therefore, the RTU set is a good choice. If the number of unobserved buses is more than the previous one, the Boltzman distribution probability is evaluated and compared with some randomly generated functions. If it is

larger, accept the new configuration, otherwise, the previous RTU set is considered and a change of RTU is made randomly. This procedure is repeated until a local minimum is found to be, with a high probability, a global one.

The following Tables 3, 4, and 5 are the results of the applications of the above algorithm to the 14- and 30-bus systems. As observed, the tuning parameters are different from one another in each case.

Table 3. Tuning Parameters for 14- and 30-Bus Systems

System:	Initial guess:	Initial temperature	Step size:	Seed number
14 Bus	14 Bus	40.0	0.8	156
30 Bus	21 Bus	100.0	0.9	156

Table 4. The Results for 14-Bus System

Number of RTUs used to observe the system:	RTU's locations:	Number of iterations:
14	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1
11	1,3,4,5,6,7,8,10,11,12,14	1
9	1,2,3,5,6,7,8,10,14	1
7	4,5,8,9,11,12,13	28
5	4,5,7,10,13	93

The system is not observable for less than 5 RTUs.

Table 5. The Results for 30-Bus System

Number of RTUs used to observe the system:	RTU's locations:	Number of iterations:
21	1,2,3,4,...,21	1
17	2,4,5,6,10,11,12,13,14,15,16, 17,18,19,24,26,29	1
12	1,4,5,6,10,11,12,13,19,24,26, 27	12

The system is not observable for less than 16 RTUs.

3.5 Remote Calibration of the Remaining RTUs

After identifying the minimum number of RTUs required to observe the whole system in one island, we need to field calibrate them. The field calibration results in obtaining reference points required for remote calibration of the remaining RTUs. This task involves the adjustment of the measuring devices by reporting the offset, linear and nonlinear correction factors for each device. Once the field calibration has been carried out and the power flows and the power injections of the system calculated, we use them as references to estimate the coefficients a_i, b_i, and c_i of the *ith* RTU through

$$a_i + b_i \times M + c_i \times M^2 = C, \tag{48}$$

where M denotes the measured value and C denotes the calculated value.

We should record a sufficiently large number of metered values so that they cover all the possible loading conditions of the power system under study. This namely includes the high, medium, and lightly loaded system, as well as high, medium and low voltage profiles. Once the data are gathered, the next task is to estimate the calibration coefficients. In order to do this, the use of a robust estimator is required. This will give us protection against large errors such as change in temperature error, error in data transmission, etc. By using a high breakdown estimator, this problem could be solved [2-5,29-34]. The high breakdown estimator that has been used in our algorithm is the Least Median Square (LMS) estimator. This estimator has the ability to handle both vertical outliers and bad leverage points (the points that are far from the bulk of data in the design space)[4]. Interestingly, this estimator is capable of withstanding a case where half the measurements are outliers.

We should recognize that the reference RTUs may experience a bias drift with time and temperature. Therefore, the whole procedure should be carried out periodically.

3.5.1 Overview of Parameter Estimation

The objective is to estimate the coefficients a_i, b_i, and c_i , involved in Equation (48). This estimation is carried out at the control center from measured and calculated values. To achieve this task, a regression technique needs to be used. Errors present in the data are largely due to changes in temperature affecting the bias of the instruments and errors in data transmission to the control center, etc. In error-filled data, the response variable and the explanatory part can be outliers. The outlying explanatory part can lead to leverage points, which are points far from the pattern of the bulk of the data in the design space. Outliers in the measurements may spoil an ordinary regression analysis such as Least Squares (LS). Despite this weakness, the LS is the most commonly used regression technique, due to its ease of computation. To overcome the outlier problem, we use the LMS estimator, which is highly robust [4,5,38].

3.5.2 Least Median of Squares (LMS) Estimator

The LMS estimator minimizes the median of the squared (weighted) residuals in the one-dimensional and simple regression case. In the multiple regression case, it minimizes the n-th ordered squared residual, yielding

$$J(\underline{x}) = r^2_{w(\mathsf{n})}, \tag{49}$$

where

$$\mathsf{n} = \left[\frac{m}{2}\right] + \left[\frac{n+}{2}\right]. \tag{50}$$

In the above equations, m denotes the number of measurements, n denotes the number of state variables, and [y] denotes the integer part of the real number y. Here, $_{w(v)}$ is the - th ordered weighted residual, which is squared first and then ordered by increasing value;

$$r^2_{w(1)} \le r^2_{w(2)} \le \ldots \le r^2_{w(m)}. \tag{51}$$

In estimation of location, the LMS is the midpoint of the shortest half of the data having the smallest distance between its extreme observations (the smallest of the differences). The halves are subsets containing ν ordered observations, with $n = \left[\frac{m}{2}\right] + 1$, which yields

$$z_{(n)} - z_{(1)}, z_{(n+1)} - z_{(2)}, \ldots, z_{(m)} - z_{(m-n+1)}, \tag{52}$$

where $\{z_{(1)}, z_{(2)}, \ldots, z_{(m)}\}$ is the ordered sample.

In simple regression, the LMS line lies at the middle of the narrowest strip covering half the observations. In multiple regression, one has hyper-strip and hyper-planes instead. All these properties imply that the LMS estimator fits the majority of data. Therefore, LMS can withstand outliers of up to half of the observations. The following example will demonstrate this property.

3.5.3 Description of the Simulation Procedure

Considering a regression model given by

$$\underline{y} = a_i + b_i \underline{x} + c_i \underline{x}^2 \tag{53}$$

where the values for $\underline{x}$ are randomly selected. They are drawn from a normal distribution with zero mean and unit variance. Second, true values for the coefficients a_i, b_i, and c_i are selected and the responses y_i are calculated. Third, Gaussian errors with zero mean and a variance of 0.01 are added to both the explanatory variables x_i and the response variables y_i. Finally, a number of data points, say b, have been made outliers. Specifically, it is the

maximum number of outliers that any regression equivariant estimator can handle has been introduced. It is given by

$$\beta_{max} = \left[\frac{m-1}{2}\right]. \tag{54}$$

Rousseeuw and Leroy [50] showed that LMS can withstand up to β_{max} outliers. The LMS solution is found by means of the resampling technique described in [50]. The standardized residuals are defined as

$$rs_i = \frac{r_i}{\hat{\sigma} w_i}. \tag{55}$$

Where

$$w_i = \min\left(1, \left(\frac{b}{RD_i}\right)^2\right). \tag{56}$$

In (56), RD_i is a robust distance of x_i with respect to the median of the x_j, which is given by

$$RD_i = \frac{x_i - \underset{j}{med}\left(x_j\right)}{1.4826\, MAD}, \tag{57}$$

and MAD is a robust estimator of scale defined as

$$MAD = \underset{i}{median}\left|x_i - \underset{j}{median}(x_j)\right|, \tag{58}$$

Note that for a leverage point, RD_i is large, implying that w_i is small and for a non-leverage point, where RD_i is small, implying that w_i is close to one. The value of b, which is a cutoff value, should be properly chosen. Because when the x_i are drawn from a Gaussian distribution, the RD_i follows a χ^2 distribution with one degree of freedom, we fix

$$b = \chi^2_{1,0.975}. \tag{59}$$

As for $\hat{\sigma}$, it is given by

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{m} u_i r_i^2}{\sum_{i=1}^{m} u_i - p}}, \tag{60}$$

where

$$u_i = \begin{cases} 1 & if \left|\frac{r_i}{s^0}\right| \le 2.5 \\ 0 & otherwise \end{cases}, \tag{61}$$

and

$$s^0 = 1.4826\left(1+\frac{5}{m-n}\right)\sqrt{med_i r_i^2}\ . \tag{62}$$

Any measurement with $|rs_i| > 2.5$ is flagged as an outlier and eliminated from the data set.

3.5.4 Simulation Result

Let us now give some simulation results that illustrate the above algorithm. The parameters a_i, b_i, and c_i have been fixed to 0.01, 0.9, and 0.1, respectively. A data set of 150 randomly selected numbers drawn from the standard normal distribution $N(0,1)$ have been chosen for the explanatory variables x_i. Plugging these values in (53), we obtain the responses y_i. Then random errors following a normal distribution $N(0,0.01)$ are added to both the x_i and y_i. Next, 74 outliers, which is the maximum allowed number of outliers obtained from (54), have been added randomly to the values taken by the explanatory and/or response variables in our data set. Finally, the LMS estimator followed by a residual analysis using equations (55 - 62) has been applied to these data. The value of b given by (59) is equal to 5.024. The magnitudes of the standardized residual has been compared to 2.5. After rejecting the outliers, we apply the LS to the remaining points and obtain the estimated values of the coefficients, which are displayed in Table 6. As can be seen from Table 6, even with 74 outliers, the model still gives good estimates. Also, it is shown that the LS without outlier rejection is not a good estimator.

Table 6. LS and Robust Estimation Results

	A	*B*	*C*
True Values	0.01	0.9	0.1
LS Estimates	-0.53754	0.41114	1.20156
Robust Estimates Without Outliers	0.00992	0.90003	0.10018

As a second example, we have fixed the true values for a_i , b_i and c_i to 0.4, 0.92, and -0.15, respectively. A set of 30 values for x have been taken from 0 to 3 pu with a step size of 0.1, and y have been calculated using (53). Then, random errors following N(0, 0.0001) have been added to both x and y values. Finally, 11 outliers have been introduced in the data set. The calibration curve estimated by means of the LMS estimator is displayed in Figure 9. Note that it fits nicely the majority of the data set.

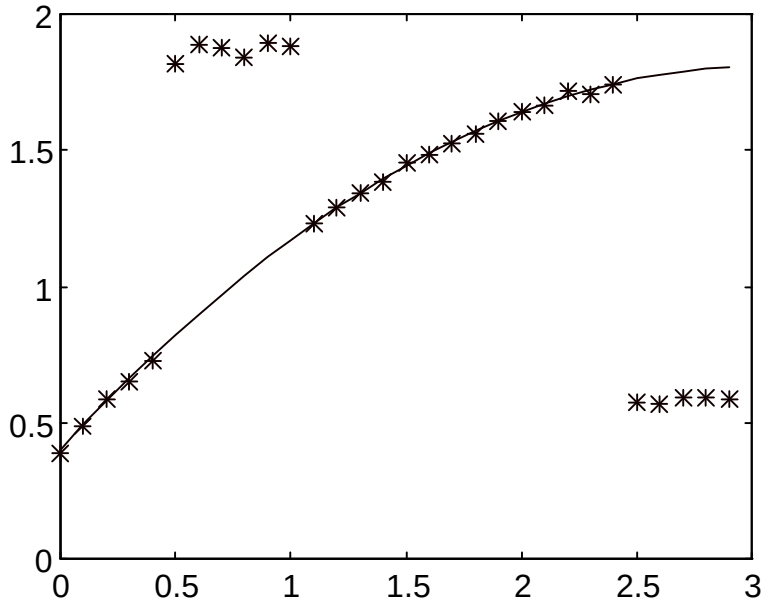

Figure 9. Calibration Curve Estimated Through LMS

CHAPTER 4

SYSTEM DECOMPOSITION

4.1 Introduction

Our experience has shown that Simulated Annealing [20-25] is a slow technique. Its parameter tuning becomes harder, and the time of execution grows exponentially as a system size gets larger. Because power system networks are relatively large, Simulated Annealing is not a practical choice for them. To overcome this deficiency, one possible solution is to tear the power system network into smaller sub-networks and to apply Simulated Annealing [20-25] on each of them. The final solution would be the union of the Simulated Annealing results of all the sub-networks that form the whole network. This chapter describes a technique based on Dynamic Programming (DP) [42-46] to tear the system into smaller sub-systems [47,48] in an optimal way.

4.2 Network Tearing

The topology of the power system network is described by a set of nodes connected in pair by branches. The objective of network tearing is to assign each node to a sub-network. This assignment should be chosen so as to minimize some cost function that is usually designed to form a balance between the number of nodes in a sub-network and the number of branches connecting the sub-networks to each other. To achieve this task, Bellman's Optimality Principle [42-46] is used. The application of this method to power system network tearing is introduced in [47-48]. This method reduces the storage requirement from N^2, where N is the total number of network nodes, to $N_1^2 + N_2^2 + \ldots + N_m^2 + N_c^2$, where $\{N_i,\ i = 1,\ldots,\ m\}$ are the number of nodes in each sub-

network. Here N_c is the total number of cut branches among sub-networks. The criterion for network tearing is to minimize both the number of cut branches connecting different sub-networks to each other and the storage requirement. It can be seen that there exists an optimum storage value for any network. A problem arises when we take into account the maximum number of branches in each sub-network as a constraint. This limits the $\{N_i, i = 1,..., m\}$ to a pre-specified range, which changes the order of the optimum solution. In our problem, the maximum number of nodes in each sub-network is chosen arbitrarily by the user. However, it should not be too small, resulting in a large number of sub-networks; in addition, it can't be too large, resulting in a large computational time for the Simulated Annealing [20-25].

4.3 Principle of Dynamic Programming

4.3.1 Introduction

The objective of Dynamic Programming [42-46] is to create an analytic structure within which we may study the multi-stage decision processes that have certain desirable basic properties. Multi-stage processing relies on a basic idea of mathematical induction that states that if a problems can be break down into smaller subproblems and if a solution of subproblems may easily be found, then the solution to the original difficult problem may be obtained by linking the solution to the subproblems to each other. Specifically, Dynamic Programming seeks the optimal solution by appealing to a basic principle, called the "Principle of Optimality". It will be discussed next.

4.3.2 Bellman s Principle of Optimality

In order to guarantee an optimum solution to our problem, we need to obtain a recurrence relation as a representation of the model. This relation is an immediate consequence of the “Principle of Optimality,” by Bellman [42-46] which states:

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision."

Conceptually, this principle may be thought of as follows: Given an optimal trajectory from point A to C, which passes through point B as shown in Figure 10, the trajectory connecting points B to C must be an optimal trajectory from B to C.

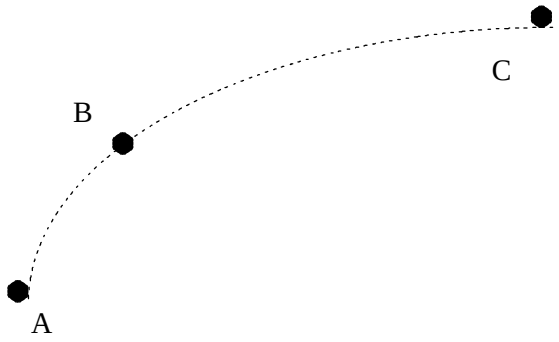

Figure 10. Optimal Trajectory From A to C

At this point, some terminology needs to be defined here:

State: It is a structure of a system. It is described by a label that indicates the characteristics of that state.

Stage: It is a single step. It corresponds to the transition of a system from one state to an adjacent one.

Action: It corresponds to taking a particular stage from one state to another.

Plan: It is a set of actions, one for each state. An optimal plan is the best set of actions for a given objective.

Return: It is the value of an action.

Value of State: It is the sum of the returns generated when the system starts in that state and a specific plan is followed. The value of a state under an optimal plan is the optimal value.

The application of the Principle of Optimality to the recurrence relation is to minimize cost from a state and stage. This can be achieved by minimizing the sum of the current single stage cost plus the minimum cost of going to the end of the process from the resulting next state. Thus, by carrying the minimum cost function backward one stage at a time, we can find the minimum cost at any state and stage. This simple concept provides the foundation for our ability to derive recurrence relations.

In addition to the recurrence relations for the minimum cost function, it is important to determine the optimum sequence of decisions to achieve the minimum cost for a given state and stage. This would be achieved by keeping track of the values that are being considered and by determining which value actually leads to the minimum cost value. This way, we can obtain the optimum decision sequence from any state and stage.

4.3.3 Example

To explain the Dynamic Programming and the Principle of Optimality, the shortest path problem is considered [43]. The objective of this problem is to find the shortest distance from city A to city H as shown in the Figure 11.

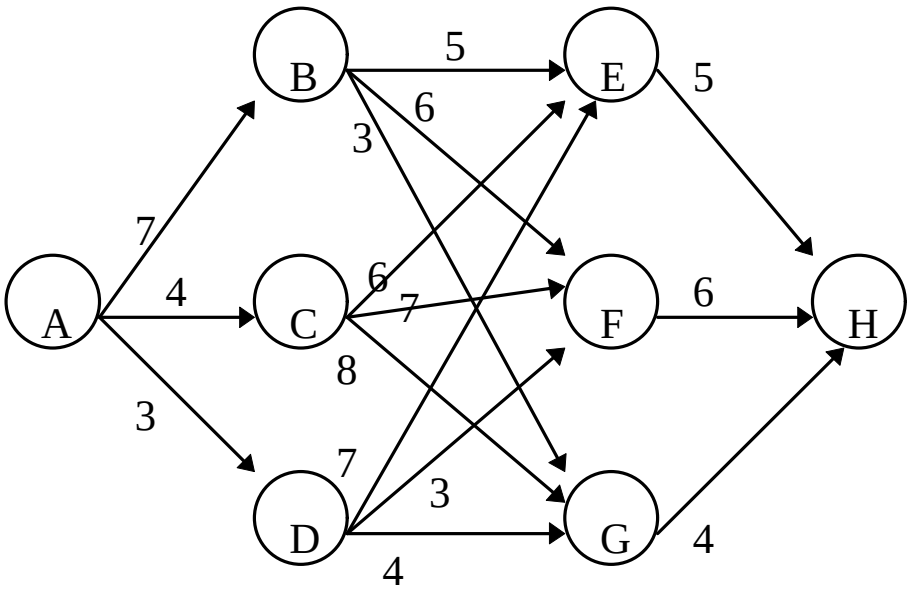

Figure 11. The Graphical Representation of The Shortest Path Problem

A description of the used terminology are as follows: a city is a state; movement from one city to another is a stage; taking a particular route from one city to another is an action; a set of actions, one for each of a number of cities is a plan; a distance from a city to an adjacent one is a return, and the sum of the short distances from city A to H is the value of state. The optimal plan is to find the shortest distance from A to H.,

The graphical representations of these terms are in Figure 12.

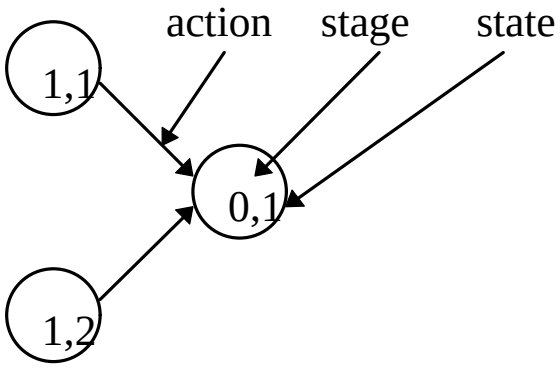

Figure 12. Graphical Representation of The Action, Stage, and State

Representing these terminology in the chosen network will results in Figures. 13 and 14.

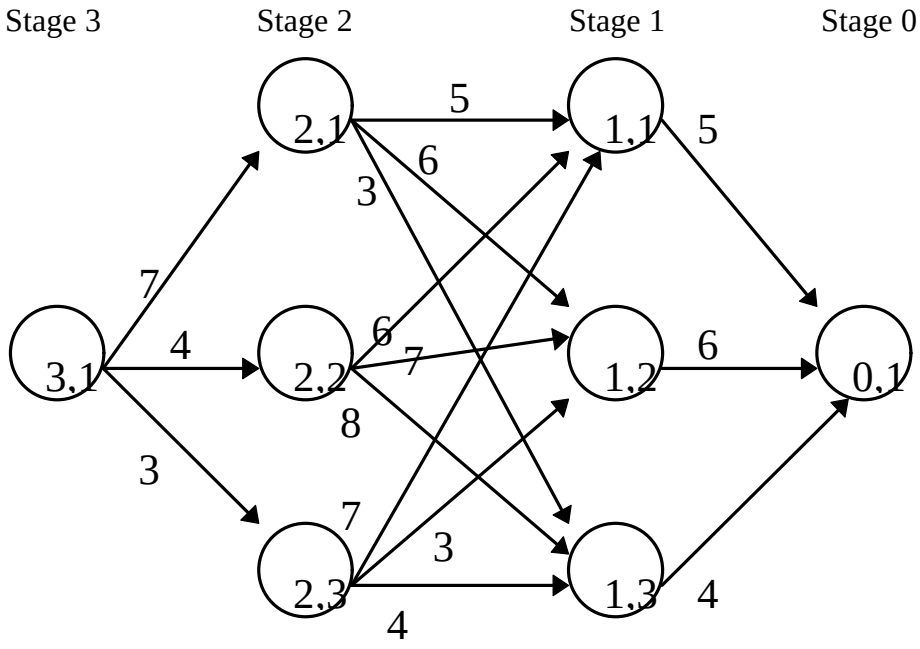

Figure 13. Stage, State and Value in the Shortest Path Problem

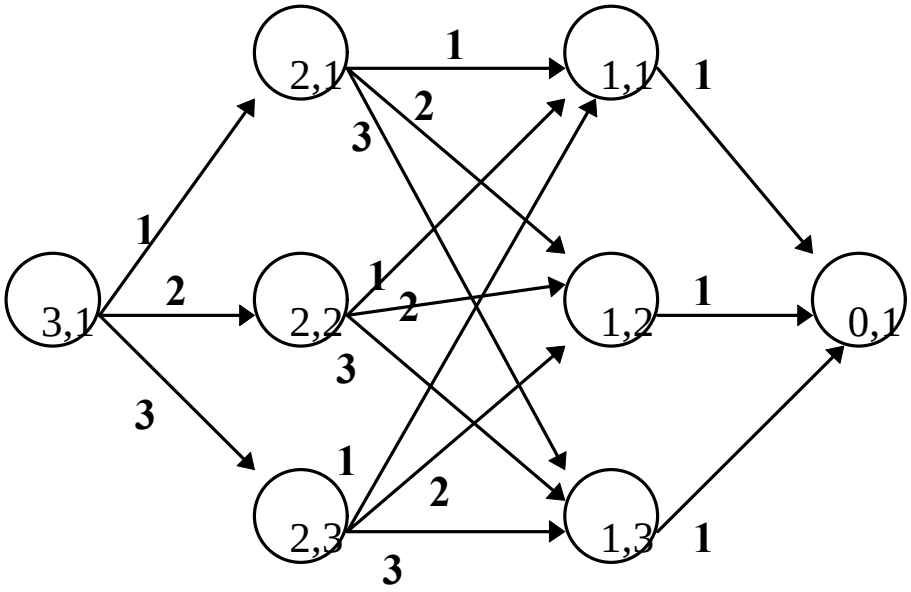

Figure 14. Stage, State and Action in the Shortest Path Problem

The objective is to start from the end of the network (H or (0,1)) and find the shortest path to get to the beginning of the network (A or (3,1)) one step at a time. First, start from stage zero (0,1), and find the shortest paths to go from all states at stage one, namely from states (1,1), (1,2), and (1,3) to state one and stage zero, (0,1). Next, find the shortest paths to go from all states at stage two, namely from states (2,1), (2,2), (2,3) to states at stage one. Then from there to state at stage zero using the already calculated shortest distances of stage one to stage zero for various states. Finally, look for the shortest paths to go from states of stage three to states of stage two, and by the help of the previously found shortest distances, determine the shortest distance from A to H. As can be seen, at each of these attempts, the shortest distance to go from a particular stage and state to the one stage before has been found. The calculated values at different states and stages are presented in the Table 7.

Table 7. Results For the Shortest Path Problem

Stage	State	Action	Value
1	1	1	5
1	2	1	6
1	3	1	4
2	1	1	10
2	1	2	12
2	1	3	**7**
2	2	1	**11**
2	2	2	13
2	2	3	12
2	3	1	12
2	3	2	9
2	3	3	**8**
3	1	1	14
3	1	2	15
3	1	3	**11**

The graphical representation of the results presented in table 7 are shown in Figure 15. Bolded lines are the selected shortest paths between any two consecutive states and stages. The bolded continuous line with the value of 11, is the optimal solution to the problem, and is the shortest path between cities A and H. Its path is described as

(3,1) ---> (2,3) ---> (1,3) ---> (0,1)

The recurrence relation in this example, as can be extracted from Table 7, is that the value of a state is the minimum overall relevant actions of the sum of the return and the value of the next state.

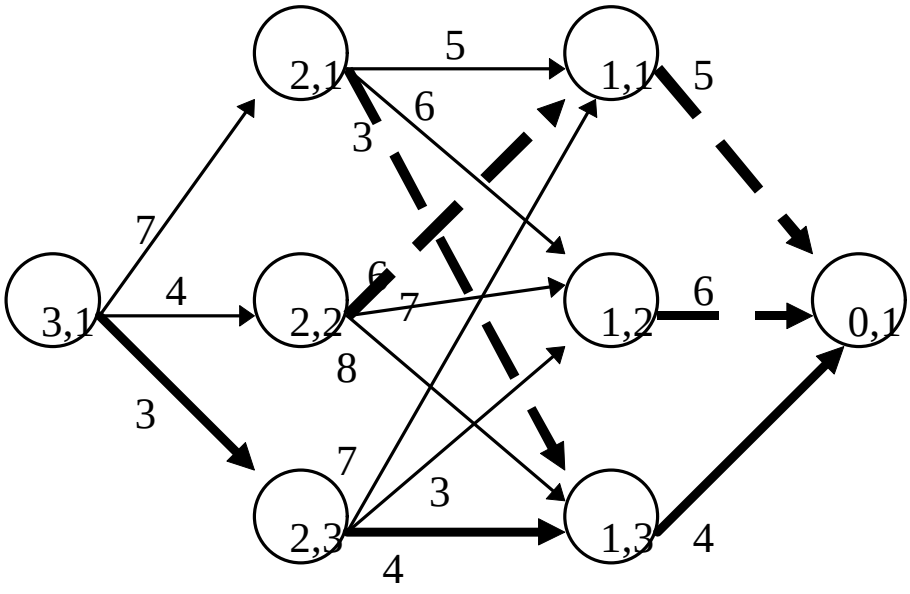

Figure 15. Shortest Paths Among States in Different Stages

4.4 Outline of the Proposed Approach

The proposed method is an improved version of the one proposed in [47], which is based on Dynamic Programming [42-46] in conjunction with some logical decisions. The objective is to tear the system into smaller sub-systems. The method described in [47] is called Networks Decomposition for Network Solution (NDNS) method, that will be described next. The algorithm starts ordering the nodes by connecting them to each other, one at a time, while keeping the number of cut branches between the ordered and unordered nodes at a minimum. Based on the Dynamic Programming [42-48] principle, summation of optimal solutions is optimal. Because at each stage the number of connected lines between selected (ordered) nodes and un-selected (unordered) nodes is minimized, and because this rule is carried out until all the nodes are accounted for, an absolute optimal solution of minimum number of lines between the sub-networks will be obtained. However, if we enforce the size of the nodes in each sub-network to be within a pre-specified limit, then the final answer will not be optimal anymore. The NDNS approach suffers from several deficiencies which make it unsuitable to the problem that we intend to solve. First, it explores all the possibilities of the connections between nodes while connecting them to each other; this in turn, makes the method slow. This is an unnecessary step in our application because our intent is to carry out observability analysis on measured networks. Therefore, only those branches that have measurements associated with them, need to be considered. Also in the NDNS method, the nodes in sub-networks are not logically ordered for the application of topological observability. This means that the nodes in sub-networks may not satisfy the observability criteria. By contrast, the proposed approach can handle this case easily because only the branches provided with measurements are considered. As a result, all grouped nodes will meet the observability requirement.

4.4.1 The Procedure

The procedure of the approach is as follows:

Step 1: Assign the measurements to the lines. Keep only the assigned lines and delete the rest of them;

Step 2: Reorder the network nodes to minimize the number of cut branches between groups;

Step 3: Group these ordered nodes into different sub-networks;

Step 4: Merge these groups to minimize the cost function and, therefore, to reduce the storage requirement;

In Steps 3 and 4, the maximum number of nodes in each group is a constraint, and it has to be taken under consideration.

4.4.1.1 Measurements Assignment

The measurement assignment is carried out according to the following procedure. A flow is assigned to the line on which it is located. An injection can be assigned to only one incident line. A flow constitutes a *link* between nodes. An injection measurement or a zero injection creates *virtual links*. These are links between the node on which the injection is located and all adjacent nodes that are connected to this node by a line. Only one of the virtual links will be converted to an actual link and the rest will be deleted. In a case of unmeasured injection, no link exists.

Let us now illustrate the measurement assignment procedure in the 7-bus system depicted in Figure 16. As can be seen from Figure 16, there are five measured flows and four injections -- one measured, one zero, one unmeasured, and one without telemetered value. If we apply our method to this system, it will be changed to the one shown in Figure 17.

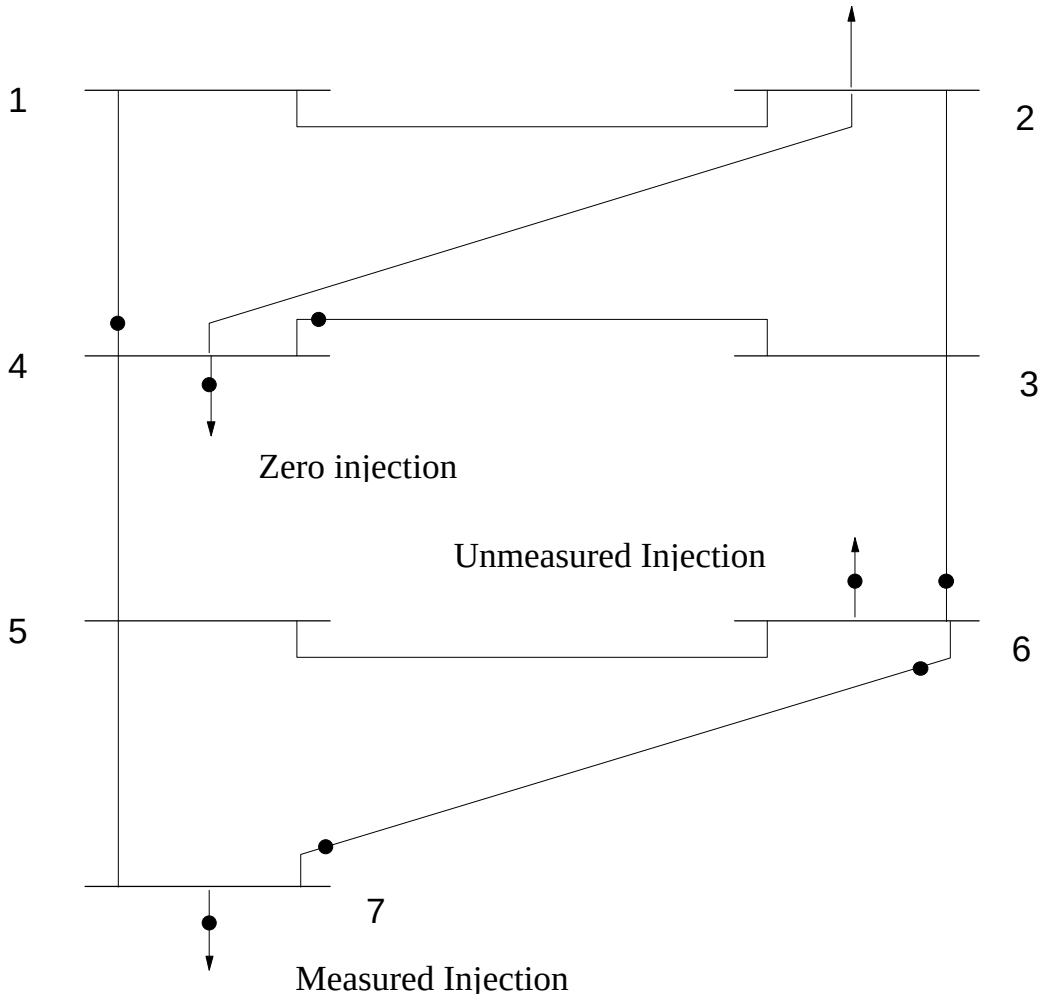

Figure 16. Seven Bus Test System With All The Branches

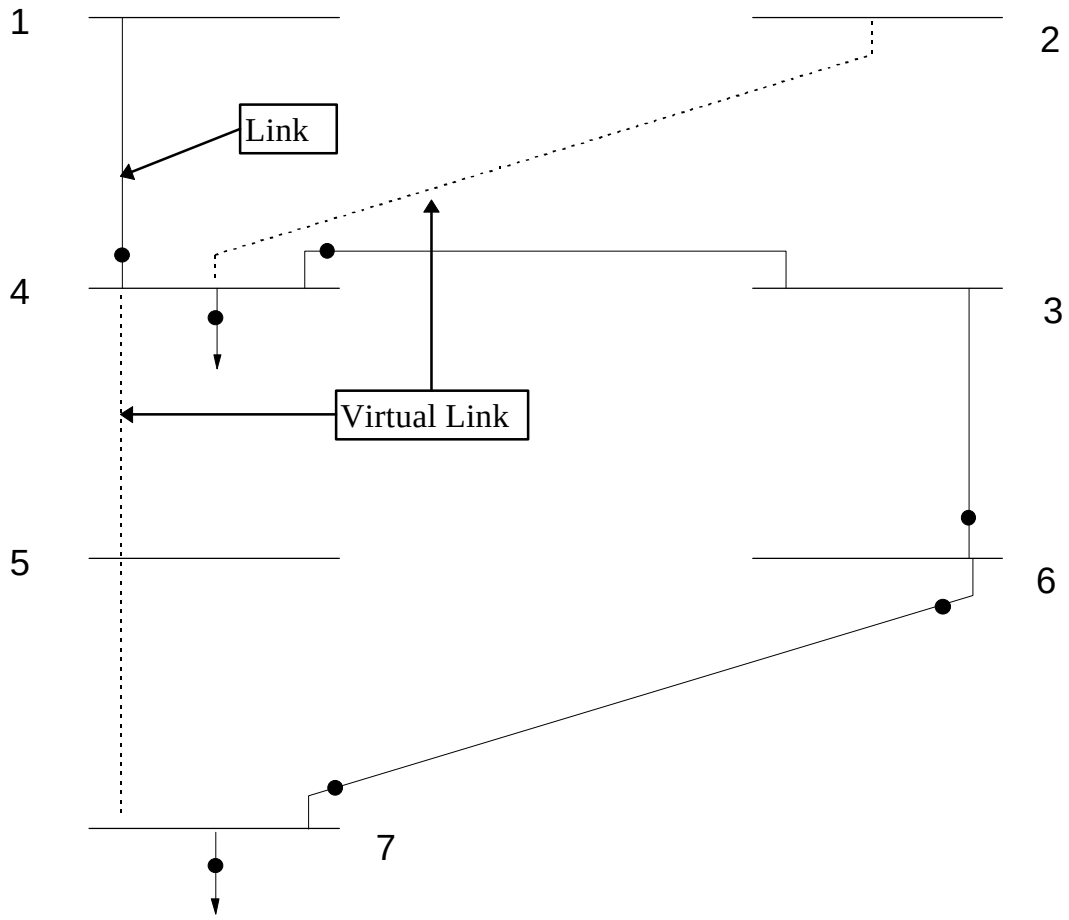

Figure 17. Seven Bus Test System With Only The Selected Branches

4.4.1.2 Ordering the Nodes

The system nodes are going to be reordered using Bellman's Optimality [42-48] Principle. The procedure consists of the following steps:

Step 1: Create all the links.

Step 2: Create an empty vector of size N named $\underline{P}$, where N is the total number of nodes.

Step 3: Select the node that has the fewest numbers of links associated with it. A measured or zero injection is only considered as one link. If more than one node has this property, just select one randomly from that set of nodes. Put it in P_1.

For each element of P_i from P_2 to P_N, do the following:

Step 4: Pick out all nodes that are connected to the nodes already included in P_1 through P_{i-1} of Vector $\underline{P}$.

Step 5: Find the total number of cut links (links that are connecting nodes in $\underline{P}$ to the ones that are not in $\underline{P}$) when P_i is included.

Step 6: Find the one P_i that gives the fewest numbers of cut links and then assign that to the P_i. If more than one node has this property, use the one that has the fewest number of connections to the unordered nodes. Again, if more than one has this property, select one randomly.

Because we keep the minimization of the number of cut links throughout the whole procedure, the formed sub-divisions have the minimum number of cut links relating them to each other as well.

4.4.1.3 Subdivisions Assignment

Put P_i in the subdivision S_j, if (i)P_{i-1} is already in S_j, (ii)P_i is connected to P_{i-1}, and (iii)the number of nodes in S_j does not exceed the pre-specified maximum number of nodes in a sub-division. If any of these criteria is not satisfied, then put P_i in the S_{j+1}.

4.4.1.4 Subdivisions Merging

The purpose of subdivision merging is to reduce the number of subdivisions by merging them into each other, as long as the limit of the maximum number of nodes in a subdivision is not violated and the merging reduces the cost function. The cost function that we are considering is given by

$$C = N_1^2 + N_2^2 + \ldots + N_m^2 + N_c^2. \tag{63}$$

Here $N_1^2, N_2^2, \ldots, N_m^2$ are the numbers of nodes in subdivisions $S_1, S_2, \ldots, S_m$ respectively, N_c is the total number of cut links that connects the subdivisions to each other, and m is the maximum number of subdivisions.

Subdivision merging proceeds as follows. Start with the smallest subdivision S_i., $i=1, \ldots, m$. Locate all other subdivisions connected to this one, by at least one branch, and merge them to S_i one by one. Calculate the cost for each case. Select the one with the largest reduction in the cost that satisfies the maximum size of the subdivision nodes limit. The cost reduction is the difference between the cost before and after the merging. This value is equal to

$$\begin{aligned} C_{reduced} = C_{before} - C_{after} &= \left[N_i^2 + N_j^2 + N_c^2\right] - \left[\left(N_i + N_j\right)^2 + \left(N_c - N_r\right)^2\right] \\ &= 2N_c N_r - 2N_i N_j - N_r^2. \end{aligned} \tag{64}$$

In (64) $C_{reduced}$ denotes the total reduction in the cost, C_{before} denotes the cost before merging and C_{after} is the cost after merging. Also, N_c is the total number of cut links and N_r is the reduced number in the number of cut links when subdivisions S_i, and S_j are merged to each other. As for N_i, and N_j, they are the number of nodes in the subdivisions S_i, and S_j respectively.

If no subdivision to merge with the S_i can be found, then just discard it from the subdivision merging list and start processing another one. In a case where the cost reduction is the same, pick one of them randomly. A successful candidate is one with the largest cost reduction value that meets the limit of the maximum number of nodes in a subdivision.

4.4.2 An Illustrative Example

The above algorithm has been applied to the IEEE 14-bus system displayed in Figure 5. It proceeds as follows. Assign measurements to the branches of the network and create the required links. Delete the unnecessary branches. This in turn decreases the number of branches needed to be considered when the algorithm explores the possible branch cutting. The results are graphically depicted in Figure 18.

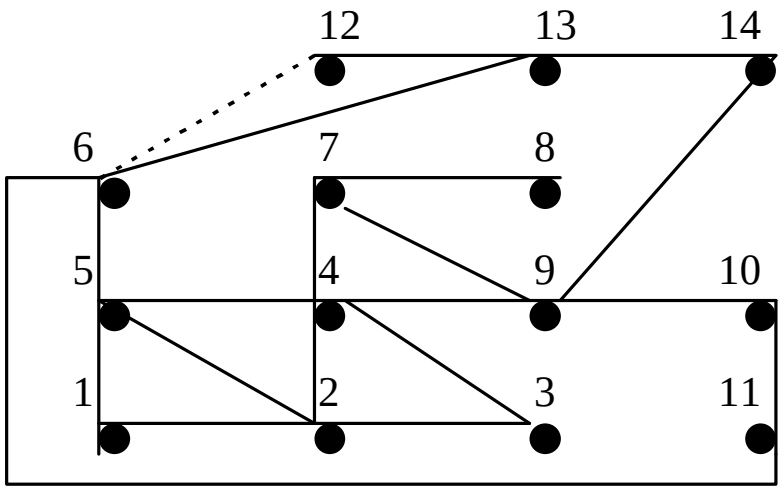

Figure 18. Graphical Representation of IEEE 14-Bus System Links

4.4.2.1 Reordering the Nodes

Create an empty vector P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}

Start from a node that has the smallest number of links attached to it, node 8. Select this node and place it on P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8													

Now, consider all the nodes that are attached to node 8. Because there is only one node, take it and put it on P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7												

In the next stage, there are two choices, Nodes 9 and 4. First, consider Node 9. Count the number of cut links between the ordered Nodes (8,7,9) and the rest of the unordered nodes. This number is 4. Apply the same procedure to Node 4. Here the number of cut links between the ordered set (8,7,4) and the rest of the unordered nodes is 5. Because the number of cut links for Node 9 is smaller than the one for Node 4, we add 9 to the vector P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9											

Now the possible nodes that can be added are 4, 10, and 14. The number of cut links for these nodes are 5, 4, and 4, respectively. Because the last two are the same, we need to look at the number of unordered nodes that are attached to 10 or 14 as if they were included in the set of ordered nodes. In the case of 10, it is Node 11, and in the case of 14, it is Node 13. Because these two numbers are the same, just pick one randomly, say 10, which yields the following vector P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10										

In the next stage, possible choices are 4, 11, and 14. Again, the cut links are 5, 4, and 4, respectively. The attached unordered nodes for the case of, 11 is 6, and for 14, is 13. Because they have the same number, again randomly pick one, say Node 14, which yields to the following vector P.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14									

Now, adding 4, 11, and 13 gives us 5, 4, and 5. Eleven is the obvious choice because it has the smallest number of cut links.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11								

The possibilities at this next stage are 4, 6, and 13. These would give us 5, 6, and 5. Because there are two choices with the same number, we need to look at the unordered nodes attached to any of them. In the case of 4, the unordered Nodes are 2, 3, and 5, while for 13, it is only 12. Therefore, 13 is our choice and P will be as follows

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13							

Next, consider adding 4, 12, or 6. They result in 6, 5, and 5, respectively, as the number of cut links. For 12 and 6, we need to look at the list of unordered nodes attached to them. For Node 6, it is only Node 5, and for 12, there is none, which is a smaller number. Therefore, we select 12.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12						

Next, the possible choices are 4 and 6, which in turn, give us 6 and 3 as the number of cut links. The choice is 6.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12	6					

Consider 5 or 4, which in turn gives us 5 and 4 as the number of cut links. Node 4 will be our preferred choice.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12	6	4				

Now, look at the number of cut links in the case that we add 5, 3, or 2. The numbers are 4, 4, and 6, respectively. Because 5 and 3 results in the same number of cut links, consider the list of unordered nodes attached to them. In the case of 5, Node 1 is the only one, and 3 does not have any unordered nodes attached to it. Therefore, we select 3.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12	6	4	3			

The next possible choices are 2 and 5, which results in 4 as the number of cut links for both of them. The unordered node attached to both of them is 1. Therefore, we select one randomly, say Node 5.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12	6	4	3	5		

In the last attempt, the possible choices are 1 and 2, which results in 4 and 2 as the number of cut links. Two is the selected one, and 1 comes at the end because it is the last.

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
8	7	9	10	14	11	13	12	6	4	3	5	2	1

At this point, we have all the nodes in the vector P. The reordering is completed. Now, we need to form the preliminary subdivisions.

4.4.2.2 Preliminary Subdivisions

The preliminary subdivisions follow the rules that were mentioned before. First, all the nodes that are connected in series to each other will be put in a subdivision, as long as they don't violate the pre-specified maximum allowed number of nodes in a subdivision. These rules will result in the following subdivisions.

S_1				S_2	S_3	S_4			S_5		S_6		
8	7	9	10	14	11	13	12	6	4	3	5	2	1

Figure 19 gives a graphical representation of the above subdivisions. The next step is to merge them together to reduce the number of subdivisions.

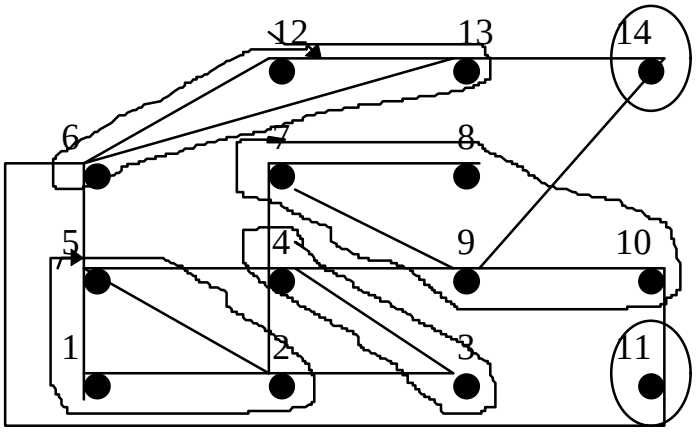

Figure 19. Preliminary Subdivisions of IEEE-14 Bus System

4.4.2.3 Subdivisions merging

Start from the smallest subdivision, such as S_2. Find a set of other subdivisions so that at least one of their nodes is connected to one of the nodes in S_2. Add their nodes to each other, if the total number of nodes is not exceeding the maximum number of nodes limit; those subdivisions meet the requirements for merging. Start calculating the reduction of cost function.

Application of this algorithm to our example is as follows: The subdivisions that can be attached to S_2 are S_1 and S_4. The cost reduction for S_2 and S_1 merging is 11, and for S_2 and S_4, merging is 13. Therefore, S_2 merges to S_4. Rename the new subdivisions. There will be as follows:

S_1				S_2	S_3				S_4		S_5		
8	7	9	10	11	14	13	12	6	4	3	5	2	1

Now, start from the smallest one, again, S_2. This subdivision can only be merged into S_1 and S_3. The cost reduction for S_2 and S_1 merging is 9, and for S_2 and S_3, is also 9. Therefore, we randomly pick one that is S_3. The new subdivisions after merging and renaming are shown below.

S_1				S_2					S_3		S_4		
8	7	9	10	14	13	12	6	11	4	3	5	2	1

The only possible merge that is not violating the maximum allowed nodes in the subdivisions is S_3 and S_4. The cost reduction for these two after merging is 27. Therefore, these two will merge, and the final result of the sub-networks after merging and renaming is depicted below.

S_1				S_2					S_3				
8	7	9	10	14	13	12	6	11	4	3	5	2	1

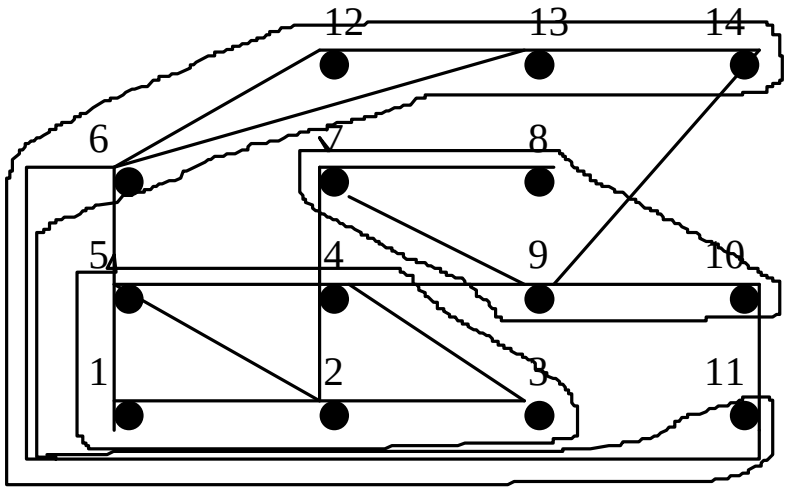

Figure 20. Final Subdivisions of IEEE 14-Bus System

Figure 20 gives a graphical representation of the subdivisions after merging. At this point, we run the simulated annealing on each of these subdivisions. The procedure is as follows: First, find an initial condition for the simulated annealing for the whole system. This could include all of the RTUs. Then, try to minimize the number of RTUs in one subdivision at a time using simulated annealing for that subdivision. Next run the observability algorithm for this reduced number of RTUs from that subdivision in conjunction with the RTUs from the other subdivisions to check whether the whole system is observable. Continue this procedure until this subdivision cannot be further reduced. Then, start to reduce another subdivision. Perform this task until all subdivisions are reduced. The final answer will be a set of minimum RTUs that makes the whole system observable.

4.5 Comparative Study

As was mentioned before, there are distinctive differences between our methodology and the NDNS method described in [47]. To make the matter clearer both methods will be applied to the 9-bus system depicted in Figure 21.

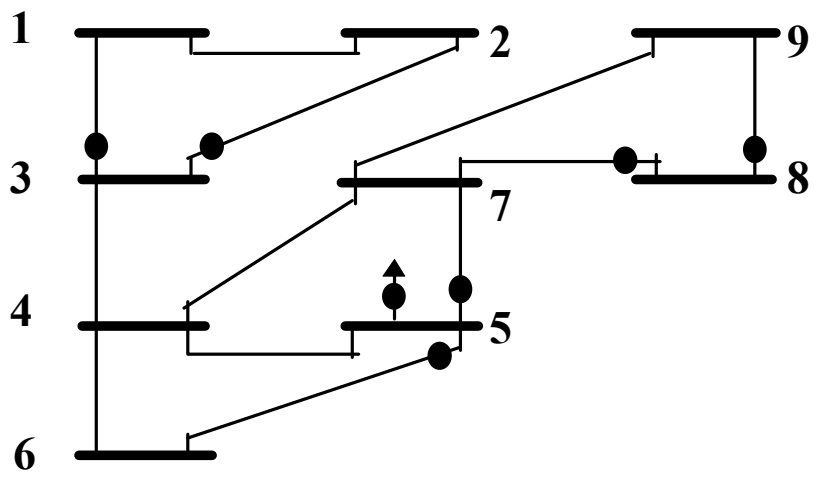

Figure 21. 9-Bus Test System

The results provided by the NDNS method summarized in Table 8. As can be seen, the NDNS method only forms one island that contains all the nodes. To reduce this system further we can set the maximum number of nodes in subnetworks to be anything from 1 to 9. Let us say we chose 5 as the limit for the number of nodes in each subnetwork. This would give us two set of subnetworks. Subnetwork I contains Nodes 1, 2, 3, 4 and 6, and subnetwork II contains Nodes 5, 7, 8 and 9. As can be seen from the measurement

locations, the first subnetwork is not observable, but the second one is. This is a major deficiency of NDNS method. To overcome this problem and make the method works the size of the subnetworks should be carefully selected which is not an easy task. Also, the NDNS method is slow since all the possibilities for branch cutting must be explored.

Table 8. Nodes Assignment For 9-Bus Test System
(Topology Oriented Method)

Selected Nodes	Adjacent Node	Num. of Cut Branches
1	**2** 3	2 3
1.**2**	**3**	
1.2.**3**	**4**	
1.2.3.**4**	5 **6** 7	4 3 5
1.2.3.4.**6**	**5** 7	2 5
1.2.3.4.6.**5**	**7**	
1.2.3.4.6.5.**7**	**8** 9	2 2
1.2.3.4.6.5.7.**8**	**9**	
1.2.3.4.6.5.7.8.**9**		

Now let us apply our method to the same 9-bus system. Figure 22 provides a graphical representation of the 9-bus system with all the unnecessary lines removed. The results are presented in Table. 9.

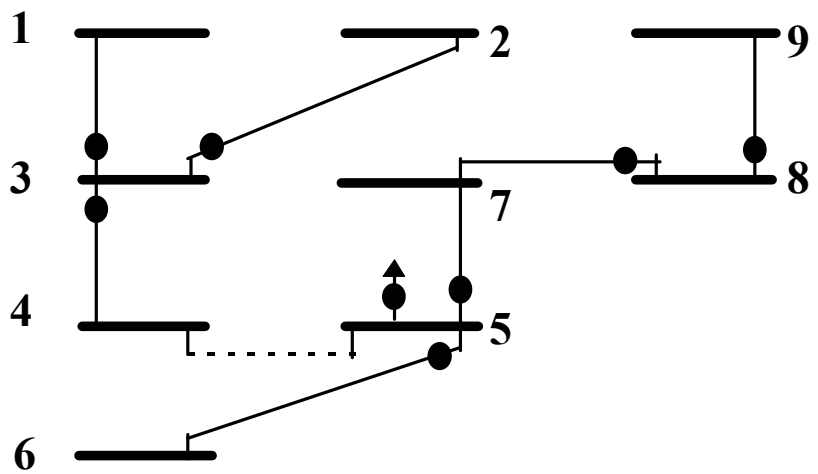

Figure 22. Links Assignments For The 9-Bus Test System

Table 9. Nodes Assignment For 9-Bus Test System

(Measurement Oriented Method)

Selected Nodes	Adjacent Node	Num. of Unordered Nodes
1	**3**	
1.**3**	**2** 4	1 2
1.3.**2**	**4**	
1.3.2.**4**	**5**	
1.3.2.4.**5**	**6** 7	1 2
1.3.2.4.5.**6**	**7**	
1.3.2.4.5.6.**7**	**8**	
1.3.2.4.5.6.7.**8**	**9**	
1.3.2.4.5.6.7.8.**9**		

The subdivisions are given by

1,3,2	4,5,6	7,8,9

We observe that there are three distinctive subdivisions which can not be merged, because they violate the maximum number of nodes in a subdivision limit. The Simulated Annealing is run on each of subdivision separately. The algorithm recognizes RTUs on bus 3 in subdivision one, bus 5 in subdivision two and bus 8 in subdivision three as the ones needed to observe the whole system. As can be seen from the above example, the new algorithm has a distinctive advantage over the NDNS method in terms of forming observable subdivisions.

4.6 Simulation Results

The Simulated Annealing has been applied both directly to the whole system and to the decomposed system of the IEEE 14- and 30-bus systems. A brief description is as follows.

Application of Simulated Annealing to the whole 14 Bus system directly, Figure 5, results in a selection of 5 RTUs (4,5,7,10,13), as a required set that observe the whole system. Network tearing results in 3 distinctive subnetworks.

S_1					S_2				S_3				
14	13	12	6	11	8	7	9	10	4	3	5	2	1

Run Simulated Annealing on S_1 with all the Buses in S_2 and S_3 included. This results in the selection of RTUs 7 and 10 as the one required in conjunction with all of the RTUs in S_2 and S_3 to observe the whole system. Next freeze these two RTUs and all the one in S_3 and run Simulated Annealing on S_2. This results in the selection of RTU 13. Now freeze all the

selected RTUs namely 7, 10 and 13 and run Simulated Annealing on S_3. This results in the selection of RTUs 4 and 5. Therefore the whole system can be observe by a set of RTUs at location 4, 5, 7, 10 and 13. This selection is the same as the one resulted from applying the Simulated Annealing to the whole system. This results and the one for IEEE 30 bus system are reported in Table 10.

Table 10. Simulated Annealing Applied to the Whole and Decomposed Systems

Whole System 14 Bus System RTUs Selected	4,5,7,10,13

	S_1	S_2	S_3
Subdivisions RTUs	7,8,9,10	6,11,12,13,14	1,2,3,4,5
Selected RTUs	7,10	13	4,5

Whole System 30 Bus System RTUs Selected	1,4,5,6,10,11,12,13,19,24,26,27

	S_1	S_2	S_3	S_4	S_5
Subdivisions RTUs	1,2,3,4,12, 13,14,15	5,6,7,8,9, 11,28	10,16,17,18, 19,20,21	22,23,24, 25,26	27,29,30
Selected RTUs	2,4,12,13	6,11	10,19	24,26	27

CHAPTER 5

REMOTE VOLTAGE MEASUREMENT CALIBRATION

5.1 Introduction

Voltage availability, due to its nature in power systems, is different from real and reactive power (*P*s and *Q*s respectively). For *P*s and *Q*s, we only have one measurement for each side of the line at any snapshot. However, for voltages around transformers, as can be seen in Figure 23, there are usually more than one meter value for the same voltage level. This gives us an alternative method for calibrating voltages, called Comparative Voltage Calibration (CVC). This is a problem of finding the largest voltage cluster, as well as satisfying the maximum tolerance requirement. Tolerance is the maximum allowable deviation in voltage-biases of any voltage-pairs located on the same bus. We are stating that, if the bias difference in pairs of voltages coming from the same bus is small and within the accepted tolerance, then there is a high probability that these voltages are correct, and the median of them can be used as a calibrated value. This probability increases if the largest cluster of reported voltages are close to each other and within the tolerance limit.

The developed method assumes that the measurement bias is constant for the processed metered values. The CVC algorithm has two steps. First, it identifies all pairs of voltage measurements at the same location whose metered values are statistically close to each other and within the acceptable tolerance over all possible operating conditions of a power system. Then, the median of all these pairs will be used as a calibrated value for the voltage at that location.

This chapter deals with the voltage calibration. It states the problem and presents statistical methods using multiple comparison to solve it. Three different sets are considered and their performances are compared to each other through a real data example. The first set is a set of classical multiple comparison methods that are incapable of dealing with any outliers. The second set is a method based on the Winsorized trimmed mean, which is a robust method. The third set is the Friedman rank sum test, which is a nonparametric technique that is not assuming any distribution. When this method was applied to the real data, despite our expectations, it failed. This opened a new problem, and part of this chapter addresses this problem as well. A new version of the Friedman test has been proposed and successfully applied to a real life example.

5.2 Statement of the Problem

Assume that a voltage at a given bus is measured by *nt* independent measurements, yielding *nt* metered values (*treatments*) at every snapshot for a period of *nb* snapshots (*blocks*). Let V_{ij} denote the metered voltage for the *j*th meter at time *i*, where *i*=1,2,...,*nb* and *j*=1,2,...,*nt*. We model the voltage measurements of V_{ij} as

$$V_{ij} = V + \beta_i + \tau_j + e_{ij}, \tag{65}$$

where V is the overall median, β_i is the effect due to the *i*th time, τ_j is the effect of *j*th meter, and e_{ij} is a random error (i.i.d.).

$$\overbrace{\quad}^{n_t \text{ meters (treatments)}}$$

$$n_b \text{ snap-shots (blocks)} \left\{ \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1j} & \cdots & v_{1n_t} \\ v_{21} & v_{22} & \cdots & v_{2j} & \cdots & v_{2n_t} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{i1} & v_{i2} & \cdots & v_{ij} & \cdots & v_{in_t} \\ \vdots & \cdots & \ddots & \cdots & \ddots & \cdots \\ v_{n_b1} & v_{n_b2} & \cdots & v_{n_bj} & \cdots & v_{n_bn_t} \end{bmatrix} \right.$$

We assume that the largest group of meters is well calibrated with a small bias, while the other groups have a larger bias. We apply different statistical tests to each measurement pair to identify whether or not its metered values are statistically close to each other. Different kinds of methods have been used, which are explained in the following sections.

5.2.1 Classical Parametric Multiple Comparison Methods

Applying the following methods to (65), we are trying to identify if there are any differences among the meters. Meaning if $t_1 = t_2 = ... = t_{nt}$. This in turn can be considered as an equivalent to test if $m_1 = m_2 = ... = m_{nt}$ where m_i is the mean of the ith meter over n_b measurement values.

5.2.1.1 Fisher s Least Significant Difference (LSD)

This is one of the least conservative methods and is as follows:

1) perform ANalysis Of VAriance (ANOVA) to test

   H_0: $m_1=m_2= m_3= .... =m_{nt}$, against

   H_1: at least one of the means differs from the rest

2) if H_0 is satisfied, then quit and conclude that voltages are calibrated

3) if H_0 is rejected, calculate LSD using

$$\text{LSD}=t_{\left(\frac{a}{2},n_i+n_j-2\right)}\sqrt{sw^2\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}, \tag{66}$$

where n_i and n_j are sample sizes for population i and j respectively, t is the critical value, and sw^2 is mean squares within the samples and is presented as

$$sw^2 = \frac{\sum_{i,j}\left(x_{ij} - \overline{x_i}\right)^2}{\left(\left(n^* n_t^* n_b\right) - n_t\right)}. \quad (67)$$

4) if $\left|\bar{y}_i - \bar{y}_j\right| \geq$ LSD, then μ_i and μ_j are different. Repeat for all $i \neq j$, $i, j \in (1,...,nt)$; and

5) The probability of a type-I error is fixed at the value of α for each pair-wise comparison.

5.2.1.2 Tukey Kramer Procedure

Tukey's method is more conservative than Fisher's LSD and is obtained as follows:

1) rank the *nt* sample means

2)Calculate the test statistics using

$$w_{ij} = q_\alpha(t, \nu)\sqrt{\frac{sw^2}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, \quad (68)$$

where n_i and n_j are the number of observations in samples i and j, t is the number of samples, sw^2 is the mean square within samples with ν degrees of freedom, $q_\alpha(t,\nu)$ is the upper-tail critical value of the studentized range for comparing *nt* different populations, and α (the experimental error rate) is the probability of observing an experiment with one or more pair-wise comparisons.

5.2.1.3 Student-Newman-Keuls Procedure

This method is less conservative due to the absence of experiment-wise and per-comparison error rates. For means, the error rate is defined as the same number of ordered

steps apart. Because the critical value W decreases as the number of steps between the means being compared decreases, the SNK is less conservative. The procedure is as follows:

1) rank nt sample means from the lowest to highest

2) μ_i and μ_j are declared different if $\left|\bar{y}_i - \bar{y}_j\right| \geq w_r$, where $\bar{y}_i$ and $\bar{y}_j$ are r steps apart and

$$w_r = q_\alpha(r,\nu)\sqrt{\frac{sw^2}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}, \tag{69}$$

where n_i and n_j are the number of observations for sample i and j respectively, sw^2 is the mean square within samples with ν degrees of freedom, and $q_\alpha(r,\nu)$ is the critical value of the studentized range.

5.2.1.4 Classical Methods Result

All of the above-mentioned algorithms identify pairs of measurements whose mean values are statistically close to each other. They all assume that the measurements follow a Gaussian distribution, which may not be the case. Also, because the comparison performs on the mean values, they are not robust; an outlier can break down the multiple comparison test results. Simulation results in the following sections represent these facts. Therefore, none of these tests are suitable candidates for our problem.

5.2.2 Winsorized Trimmed Mean Method

This method applies the test to every pair of measurements to check whether or not the median of their absolute differences recorded at successive times is small. Consider a pair of the n_t voltage measurements, say $V_{\underline{p}}$ and $V_{\underline{q}}$. Let

$$\left\{\left(V_{p1}, V_{q1}\right), \left(V_{p2}, V_{q2}\right), \ \ldots \ \left(V_{pn_b}, V_{qn_b}\right)\right\} \tag{70}$$

be a series of n_b pairs of metered values associated with these voltage measurements recorded at successive times $\{t_1, t_2, \ldots, t_{nb}\}$. Let d_i be the absolute difference of the metered pair recorded at time t_i, namely $d_i = \left|V_{p_i} - V_{q_i}\right|$. The sample $\left\{d_1, d_2, \cdots, d_{n_b}\right\}$ may be regarded as n_b realizations of a random variable D. If the median of D, denoted by D_{med}, is within the tolerance, then the measurements $V_{\underline{p}}$ and $V_{\underline{q}}$ are assumed to be calibrated; otherwise, at least one of them is not. This test has been implemented through the following algorithm [32]:

1) Calculate the differences $d_i = |V_{pi} - V_{qi}|$, for $i = 1, \ldots, n_b$;

2) Calculate the 2β-trimmed mean of the d_i, $\overline{D}_{2\beta}$, defined as

$$\overline{D}_{2\beta} = \frac{1}{n_b - 2n} \sum_{i=n+1}^{n_b - n} d_{(i)}, \tag{71}$$

where the d_i are the ordered differences from small to large, yielding

$$d_1 \le d_2 \le \cdots \le d_{n_b}. \tag{72}$$

$n = [n_b \beta]$ is the number of data that is trimmed from each side of the ordered data sample, and $[x]$ denotes the integer part of the real number x. Here, the trimming fraction β is set equal to a given value, say $\beta = 0.25$.

3) Calculate the standard error $\hat{S}$ of $\overline{D}_{2\beta}$. It is given by

$$\hat{S} = \frac{1}{1 - 2\beta} \frac{s_{w_{2\beta}}}{\sqrt{n_b}}, \tag{73}$$

where

$$s^2_{w2\beta} = \frac{1}{n_b - 1} \sum_{i=1}^{n_b} (X_i - \bar{X})^2, \tag{74}$$

$$X_i = \begin{cases} d_{(\gamma+1)}, & i \le \gamma \\ d_{(i)}, & \gamma \le i \le n_b - \gamma, \\ d_{(n_b-\gamma)}, & n_b - \gamma < i \end{cases} \tag{75}$$

and

$$\bar{X} = \frac{1}{n_b} \sum_{i=1}^{n_b} X_i . \tag{76}$$

Note that $s^2_{w2\beta}$ given by Equation (74) is the sample variance of the 2β-Winsorized sample.

4) Calculate the standardized distances $\bar{D}_S = (\bar{D}_{2\beta} - \lambda) / \hat{S}$, where λ is the tolerance.

5) Compare $\bar{D}_S$ to the critical point $t_{n-1,1-\alpha}$ of a t-distribution with $n-1$ degrees of freedom, where $n = n_b - [2\beta n_b]$. The significance level α is given by

$$\alpha = \frac{0.05}{n(n-1)}. \tag{77}$$

Specifically, accept the null hypothesis

$$H_0 : D_{med} < \lambda, \tag{78}$$

if

$$\bar{D}_S < \lambda \ ; \tag{79}$$

reject H_0 otherwise. Note that when the hypothesis H_0 is accepted, the pair of voltage measurements $V_{\underline{p}}$ and $V_{\underline{q}}$ are assumed to be calibrated.

5.2.3 The Nonparametric Test

5.2.3.1 The Friedman Rank Test

We are using the Friedman rank test as our selection of a nonparametric test, which is a test that does not assume any distribution for data. Application of this test to our problem is as follows.

Applying the Friedman test to (65), we are trying to identify if there are any differences among the meters, meaning if $t_1 = t_2 = ... = t_{nt}$. This in turn is equivalent to test if $M_1 = M_2 = ... = M_{nt}$ where M_i is the median of the ith meter over n_b measurement values. To achieve this task, first, rank the voltages of all the meters within every time-snap separately from 1 to n_t. If there are ties, then use the midranking, meaning average all the ranks that would be assigned to the tied observations and then change all the rank values of these observations to their calculated average ones. R_{ij} is the rank corresponding to voltage V_{ij}. Add up all ranks for each meter over all the snapshots, that is

$$R_j = \sum_{i=1}^{n_b} R_{ij} \quad \text{for } j=1,2,...,n_t. \tag{80}$$

The Friedman test statistic T is defined as

$$T = \left[\frac{12}{n_b n_t (n_t+1)} \sum_{j=1}^{n_b} R_j^2 \right] - 3n_b(n_t+1). \tag{81}$$

T is compared to the *chi-squared* distribution with (n_t-1) degrees of freedom to determine if the null hypothesis should be rejected or not. When the Friedman test rejects the null hypothesis, the M_i's (medians) are not all equal, and we need to determine which voltages have a statistically equal meter effect. To achieve this, we use a multiple comparison procedure based on the Friedman rank sum test that involves a direct comparison between every pair of meters (treatments) [51,60]. The procedure starts by choosing an overall experiment-wise error rate (α). The selected value in our case is 0.01. Next, set

$$\alpha^* = \frac{\alpha}{n_t(n_t - 1)}. \tag{82}$$

α^* (individual error rate) is related to the probability of type-I error and is defined as the probability that a specific pair of meters that are not different will be declared significantly different by the procedure. The threshold value is given by

$$L = z_{\alpha^*}\sqrt{\frac{n_b n_t(n_t + 1)}{6}}, \tag{83}$$

where z_{α^*} is such that $P\left(Z > z_{\alpha^*}\right) = \alpha^*$ and $Z \approx N(0,1)$. As an example, for n_b=20, n_t=5, and α=0.1, then α^*=0.005, and z_{α^*}=2.575, which in turn results in L=25.75.

The Friedman rank sums between the meters j and r, which are R_j and R_r, are declared to be significantly different if

$$| R_j - R_r | \geq L. \tag{84}$$

When this test was applied to the real data, where there was no crossing between voltage lines and every line always stayed above or below the other one, it failed. When

there is no crossing between the data lines, then the top line at each snapshot has the highest rank, and the one after has the second highest rank, and so on. This way the difference between the sum ranks of different treatments are at the maximum, and the test rejects the null hypothesis. Also, this test is unable to identify how far the voltages are from each other. The test works on the ranks, and the ranks are obtained based on the relative differences among the data; therefore, as long as one voltage is above the other one, it doesn't matter how much, it has a higher rank. Due to these deficiencies, a modified algorithm is proposed.

5.2.3.2 The Modified Friedman Test

One way to overcome the Friedman test deficiencies for solving our problem is to add random noise to all the measurements. This is to ensure that the lines are crossing, which, in turn, corrects the ranks and makes the Friedman test work in our case. The amount of noise that needs to be added has a directly proportional to tolerance, which is selected by the operator. Tolerance is the maximum allowable deviation in voltage-biases applied to every voltage-pair located on the same bus. A conservative operator may select a tolerance that may lead to rejection of all data; on the other hand, another operator may select another one that can result in accepting some or all of the pairs.

Added noise adds a new term to our voltage model defined in Equation (65). In the new model, we have

$$V_{ij} = V + \beta_i + \tau_j + e_{ij} + z_{ij}, \tag{85}$$

where e_{ij}, as mentioned before is a random error that is assumed to be a continuous function with unknown probability distribution function (*pdf*), and z_{ij} is a random noise with $N(0,\sigma^2)$. The objective is to choose σ^2 based on the pre-selected λ (reciprocal of tolerance) so that

$$P\left(\frac{-l}{2} \le z_{ij} \le \frac{l}{2}\right) = 0.90. \tag{86}$$

This in turn gives

$$P\left(\frac{-l}{2s} \le z \le \frac{l}{2s}\right) = 0.90, \tag{87}$$

where z is the test statistic for a normal test. This also results in the following

$$\frac{l}{2s} = z_{0.95}. \tag{88}$$

Substituting for $z_{0.95}$=1.645 yields

$$s = \frac{l}{3.29}. \tag{89}$$

Therefore, the noise distribution is

$$N\left(0, \frac{l^2}{10.8241}\right). \tag{90}$$

Application of this method to our data set resulted in identifying the voltages at T_1, T_2 and S_1 as the closest ones to each other. The next section illustrates the results.

5.3 Illustrative Example

Let us illustrate the present methods through a real-life example kindly provided to us by Adibi. Consider the substation displayed in Figure 23. It consists of two autotransformers that connect three 500 kV-transmission lines to four 230 kV-transmission lines. The voltages of the transmission lines and those of the 230 kV-sides of the autotransformers are metered, yielding 9 voltage measurements. They are represented by circles in Figure 23. Three of these measurements are on the 500 kV-lines and are labeled P1 through P3, four are on the 230 kV lines and are labeled S1 through S4, and two are on the 230 kV-sides of the autotransformers and are labeled T1 and T2.

Figure 24 displays the actual hourly metered values in p.u. recorded at the 230 kV side over a 12-day period, yielding 286 data points for each of the six measurements T1, T2, and S1 through S4. We observe that 3 of these time series, namely those recorded at points S1, T1, and T2, remain close to each other throughout the time period, whereas the 3 others depart from them. As a result, the former measurements are assumed to be calibrated, while the latter are not.

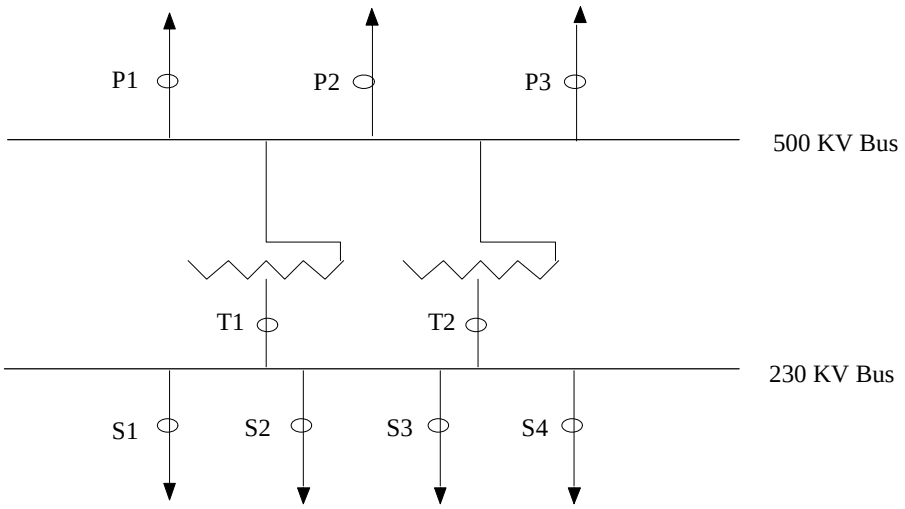

Figure 23. One-line Diagram of a Transmission Substation.

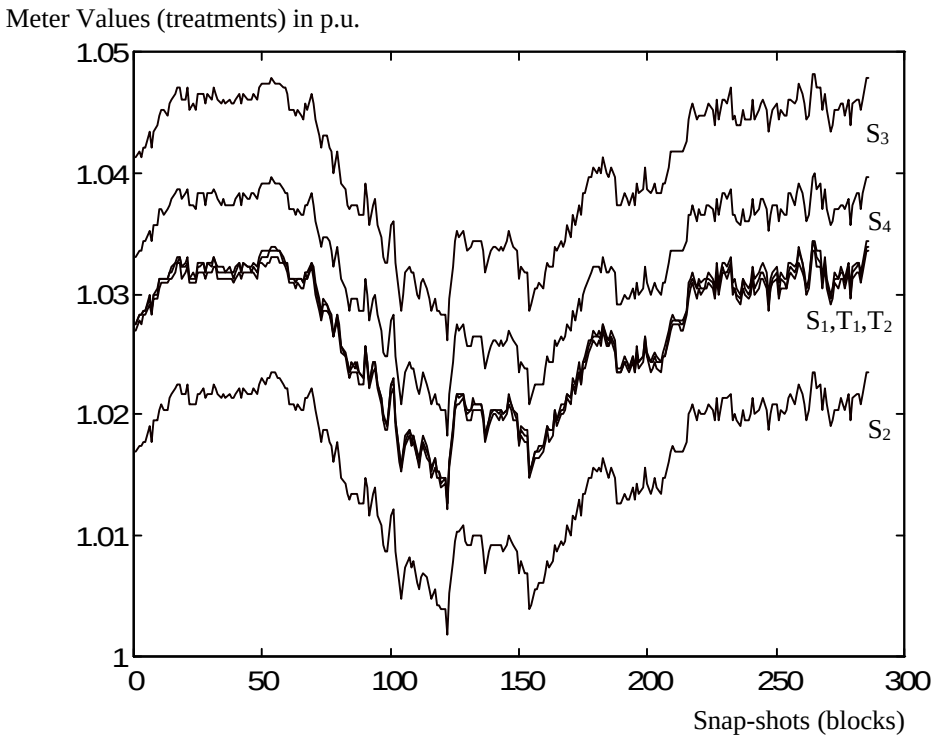

Figure 24. Hourly Voltage Metered Values in p.u. at Points T1, T2, S1- S4

Let us now apply the statistical tests described in this chapter to these 6 time-series. Tables 11 through 14 summarize the results of these tests. The solid bar under the values in Table 11 represents the statistically close voltages.

Table 11. Student-Newman-Keuls Test Results

Group	S2	T2	T1	S1	S4	S3
Mean	1.016	1.026	1.027	1.027	1.032	1.041

Table 12. Simultaneous Confidence Intervals for All Pair-wise Differences of Means

(95.0% Confidence Interval)

(Tukey Method)

Test Results	Group i	Group j	Mean(i)-Mean(j)	Lower Limit	Upper Limit
R	S1	S2	0.010906	0.005952	0.015860
R	S1	S3	-0.013560	-0.018514	-0.008606
R	S1	S4	-0.005504	-0.010458	-0.000550
A	**S1**	**T1**	0.000257	-0.004696	0.005211
A	**S1**	**T2**	0.000709	-0.004245	0.005663
R	S2	S3	-0.024466	-0.029420	-0.019512
R	S2	S4	-0.016410	-0.021363	-0.011456
R	S2	T1	-0.010649	-0.015602	-0.005695
R	S2	T2	-0.010197	-0.015151	-0.005243
R	S3	S4	0.008056	0.003102	0.013010
R	S3	T1	0.013817	0.008863	0.018771
R	S3	T2	0.014269	0.009315	0.019223
R	S4	T1	0.005761	0.000807	0.010715
R	S4	T2	0.006213	0.001259	0.011167
A	**T1**	**T2**	0.000452	-0.004502	0.005406

Table 13. Simultaneous Confidence Intervals for All Pair-wise Differences of Means

(95.0% Confidence Interval)

(Fisher--LSD Test)

Test Results	Group i	Group j	Mean(i)-Mean(j)	Lower Limit	Upper Limit
R	S1	S2	0.010906	0.007916	0.013895
R	S1	S3	-0.013560	-0.016549	-0.010570
R	S1	S4	-0.005504	-0.008493	-0.002514
A	**S1**	**T1**	0.000257	-0.002732	0.003247
A	**S1**	**T2**	0.000709	-0.002280	0.003698
R	S2	S3	-0.024466	-0.027455	-0.021476
R	S2	S4	-0.016410	-0.019399	-0.013420
R	S2	T1	-0.010649	-0.013638	-0.007659
R	S2	T2	-0.010197	-0.013186	-0.007207
R	S3	S4	0.008056	0.005067	0.011045
R	S3	T1	0.013817	0.010828	0.016807
R	S3	T2	0.014269	0.011280	0.017258
R	S4	T1	0.005761	0.002772	0.008751
R	S4	T2	0.006213	0.003224	0.009202
A	**T1**	**T2**	0.000452	-0.002538	0.003441

Table 14. Statistical Test Applied to the Pair-wise Differences of the 230 kV Voltages Measurements Using the Winsorized Trimmed Mean.

Test Results	Voltage Pairs	Trimmed Mean	Standard Error	Standardized Differences	$t_{k-1,1-\alpha}$
R	S1-S2	0.01091	0.000018	4631.244	5.075377
R	S1-S3	0.01354	0.000014	6103.112	5.075377
R	S1-S4	0.00556	0.000009	4247.415	5.075377
A	**S1-T1**	0.00026	0.000012	-540.414	5.075377
A	**S1-T2**	0.00072	0.000011	-216.967	5.075377
R	S2-S3	0.02448	0.000015	12841.85	5.075377
R	S2-S4	0.01639	0.000010	12773.65	5.075377
R	S2-T1	0.01063	0.000013	6064.496	5.075377
R	S2-T2	0.01016	0.000015	5271.96	5.075377
R	S3-S4	0.00806	0.000012	4799.05	5.075377
R	S3-T1	0.01383	0.000009	12426.56	5.075377
R	S3-T2	0.01424	0.000009	11719.90	5.075377
R	S4-T1	0.00572	0.000017	2282.54	5.075377
R	S4-T2	0.00617	0.000017	2595.429	5.075377
A	**T1-T2**	0.00043	0.00000001	-0.000565	5.075377

Note that in all of the above tables, the letters R and A stand for rejection and acceptance of the hypothesis H_0 , respectively.

The first three tests performed well in the absence of bad data, as can be seen from Tables 11, 12 and 13. However, introducing only one outlier caused all the tests to fail. The results of the Winsorized trimmed mean test with $\lambda = 10^{-3}$ p.u. is represented in Table 14. As expected, the test identifies as small (smaller than 10^{-3} p.u.) the differences between the metered values recorded at measurement points S1 and T1, S1 and T2, and T1 and T2. This test is robust and able to handle up to [2bn_b] bad data. The Friedman test, when applied directly to the data, failed. However, the modified Friedman test worked well with the tolerance of 0.012 and identified S1, T1, and T2 as the close voltage measurements within the tolerance. After introducing the noise to the data, they look like as displayed in Figure 25.

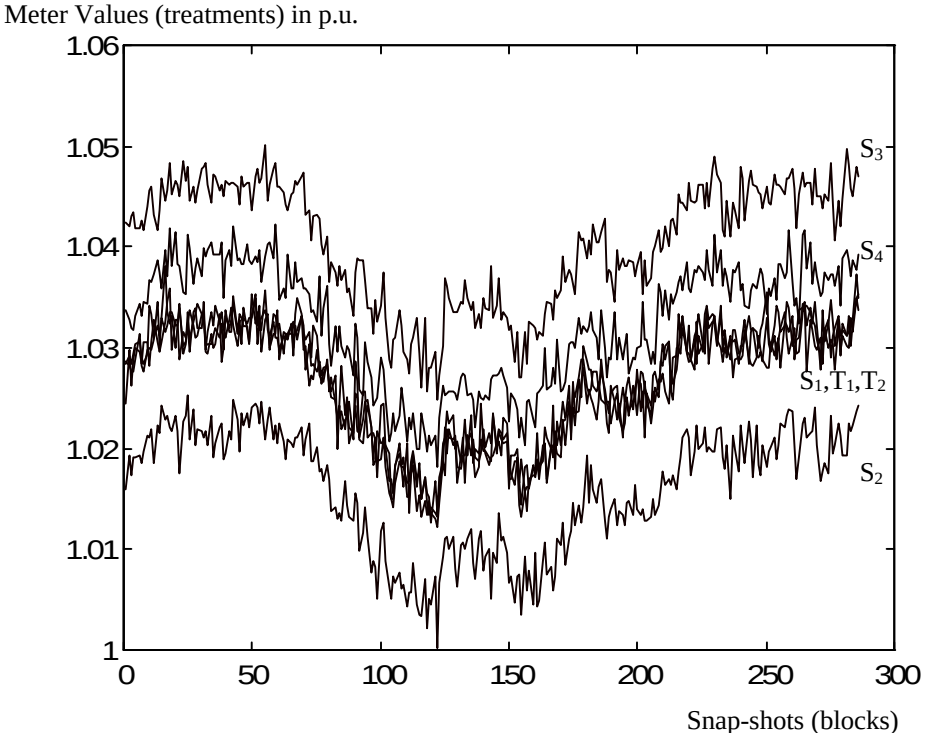

Figure 25. Voltage Measurements With Added Noise

The following table represents the sum of ranks for both methods: the regular and the modified Friedman rank test.

Table 15. Sum of Ranks

	S1	S2	S3	S4	T1	T2
Regular	1077.5	286.0	1716.0	1430.0	908.0	588.5
Modified	930.5	286.0	1716.0	1427.0	864.5	782.0

In this example, n_b=286, n_t=6, and α=0.01, then $\alpha^*=0.000\overline{3}$, and z_{α^*}=3.4462, which in turn results in L=154.2. After comparing the difference between the sum of the ranks of different treatments, it can be seen that with this value of L none of the treatment pairs in the regular method are selected; however, in the modified method T1, T2, and S1 are the selected ones. Table 16 represents this.

Table 16. Modified Friedman Rank Test Results

Groups	S2	T2	T1	S1	S4	S3
Sum of Ranks	286.0	782.0	864.5	930.5	1427.0	1716.0

5.4 Resistance of the Friedman Rank Test

5.4.1 Resistance Concept

Collected data are often contaminated by recordings, transmission, and round-off errors. A logical question is to find how many contaminated observations a test statistic can handle. This introduces the concept of test resistance. This section studies the resistance of the Friedman rank test. Resistance of a test describes how a test reacts in the presence of outliers and also how many outliers or contaminated observations a test can handle before it breaks down and becomes unreliable. One approach to answer this question is to use the testing model proposed by Ylvisaker [52,53], which is the addition contamination model. In this model, contaminating points were added to the sample, and this is carried out until the test breaks down. This model is described as follows

Let $x_1, x_2, \ldots, x_N$ be a sample with c.d.f. of F_N, test statistics of $T(F_N)$, and critical value of C_N. The procedure is

$$\text{Reject } H_0 \text{ if } T(F_N) > C_N. \tag{91}$$

Ylvisaker's definition for the resistance to rejection is defined in [52-53] as "the smallest $\frac{m_0}{N}$ where $\{ m_0 \in Z^+ \}$ and for any choice of the remaining $(N-m_0)$ points there exists $x_1, x_2, \ldots, x_{m0}$ such that $T(F_N) > C_N$."

An alternative method to Ylvisaker's method is the replacement contamination method. This method is more consistent with the actual cases where collected data are subject to errors. In this method, the observations are changed one at a time with the one at the extreme position until the test breaks down.

When the test statistic is close to critical value, then it takes less contaminated data to get it past the critical value. Therefore, a concept of maximum test resistance [53], needs to be used. This new concept starts from the value of the test statistic at the origin. Hence, Maximum test Resistance to Rejection(MRR) is a measure of the resistance of the test to rejection. It is defined as the number of contamination replacements required to reject a test divided by the total sample size when the test statistic starts from zero. The mathematical definition of it is as follows:

> Maximum resistance to rejection of (T, C_N) is defined in [53] as "the smallest proportion $\frac{m}{N}$ such that for any sample S there exists a contaminated sample S_m^* such that
>
> $$T > \left(S_m^*\right) > C_N . \tag{92}$$

In the above definition, S is the combined sample of size N, and S_m^* is any sample obtained by replacing m of the points in S with arbitrary values."

5.4.2 Maximum Resistance Calculation

To evaluate the resistance of the Friedman test, we start from a balanced-block model. This is a model with the same sum rank number for each treatment to ensure that we are starting from the test statistic of zero. We then try a perturbation scheme, which consists of a series of contaminations. Each contamination is a move of one observation from its position to one of the extreme positions. To have the most effect, we start from one treatment and change its smallest entry first, then the second smallest one, and so on. This way we are trying to make this treatment as large as possible. The test statistic gets closer to critical value each time, and when it passes this value, the test fails. The contamination scheme is optimal in the sense that it requires the fewest number of contaminated observations to reach rejection. A formula that is derived to mathematically

model the procedure for testing the Friedman rank test is as follows. Rejection takes place if the test statistic is greater than critical value. The critical value comes from a *chi_squared* distribution with (n_t-*1*) degree of freedom and confidence interval of a. The test statistics is defined as

$$T = \frac{12}{n_t n_b n(n_t+1)} \sum R_i^2 - 3n_b n(n_t+1), \tag{93}$$

where n is the number of groups, n_t is the number of treatments, n_b is the number of blocks of data, and R_i is the sum of the ranks of each treatment.

The sum of squares of treatments-rank is defined as

$$\sum_{i=1}^{n_t} R_i^2 =$$

$$\left(\left[\begin{array}{l} A_1 * \left(R + \sum_{i=1}^{\left[\frac{F-1}{n}\right]} (n_t - i) * U(n_t - 1 - i) \right) + A_2 * \left(R + \sum_{i=1}^{\left[\frac{F-1}{n}\right]+1} (n_t - i) * U(n_t - 1 - i) \right) + \\ (n-F) * R * U(n-F) \end{array} \right]^2 + \right.$$

$$\sum_{j=2}^{n_t+1} \left(\left(\begin{array}{l} \left(R - \left[\frac{F-1}{n}\right]\right) * U\left(n_t - \left[\frac{F-1}{n}\right] + 1 - j\right) + \\ \left(R - (n_t + 1 - j)\right) * U(n_t - j) * U\left(j - \left(n_t - \left[\frac{F-1}{n}\right] + 2\right)\right) * U(n_t + 1 - j) \end{array} \right) * A_1 + \right.$$

$$\left(\begin{array}{l} \left(R - \left[\frac{F-1}{n}\right] - 1\right) * U\left(n_t - \left[\frac{F-1}{n}\right] - j\right) + \\ \left(R - (n_t + 1 - j)\right) * U(n_t - j) * U\left(j - \left(n_t - \left[\frac{F-1}{n}\right] + 1\right)\right) * U(n_t + 1 - j) \end{array} \right) * A_2 +$$

$$\left. \left. (n-F) * R * U(n-F) * U(n_t - j) \right)^2 \right) * U\left(\left((n_t - 1) * n\right) - F\right), \tag{94}$$

where

$$A_1 = \left[(n-1) - rem\left(\frac{F+(n-1)}{n} \right) \right] * \ U(F-n), \tag{95}$$

and

$$A_2 = \left[1 + rem\left(\frac{F+(n-1)}{n} \right) \right]. \tag{96}$$

In the above equation, F is the number of bad data, $[y]$ is the integer part of y, $rem(y/x)$ is the remainder of y/x, the unit step function is

$$U(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}. \tag{97}$$

R is the initial sum of the ranks in one group and is equal to

$$R = \frac{n_b * (n_t + 1)}{2}. \tag{98}$$

Using Equations (93-98), we can calculate the number of bad data (F) for any number of groups (n) and treatments (n_t). Figures 26, and 27 are graphical representations of the ratio between the number of bad data and the total number of sample size for different groups and treatments.

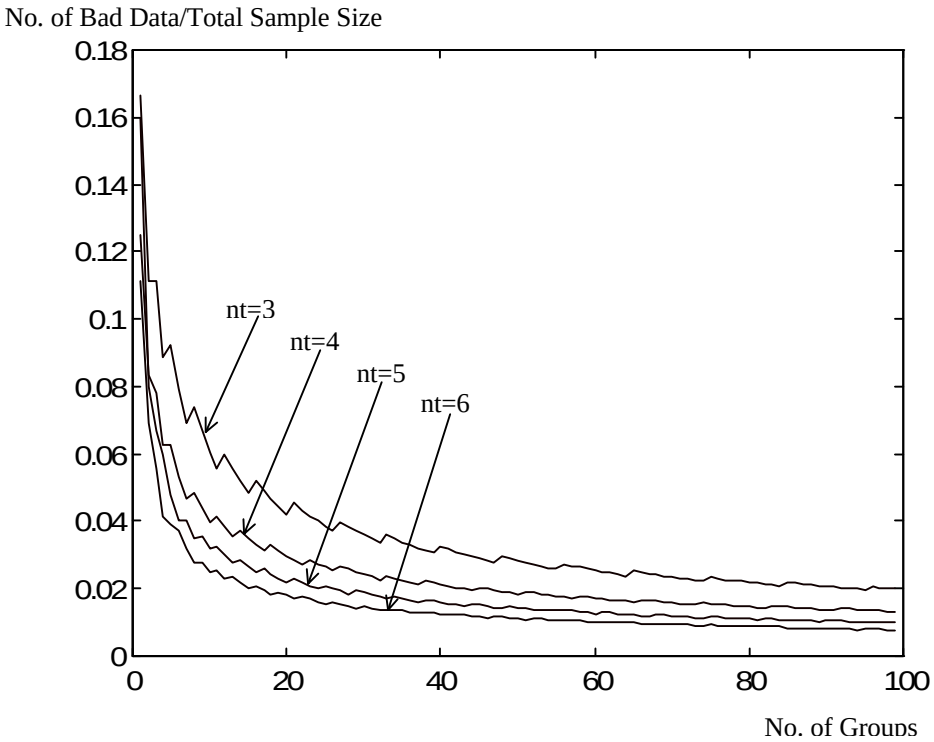

Figure 26. MRR of Friedman test for nt=3,...,6

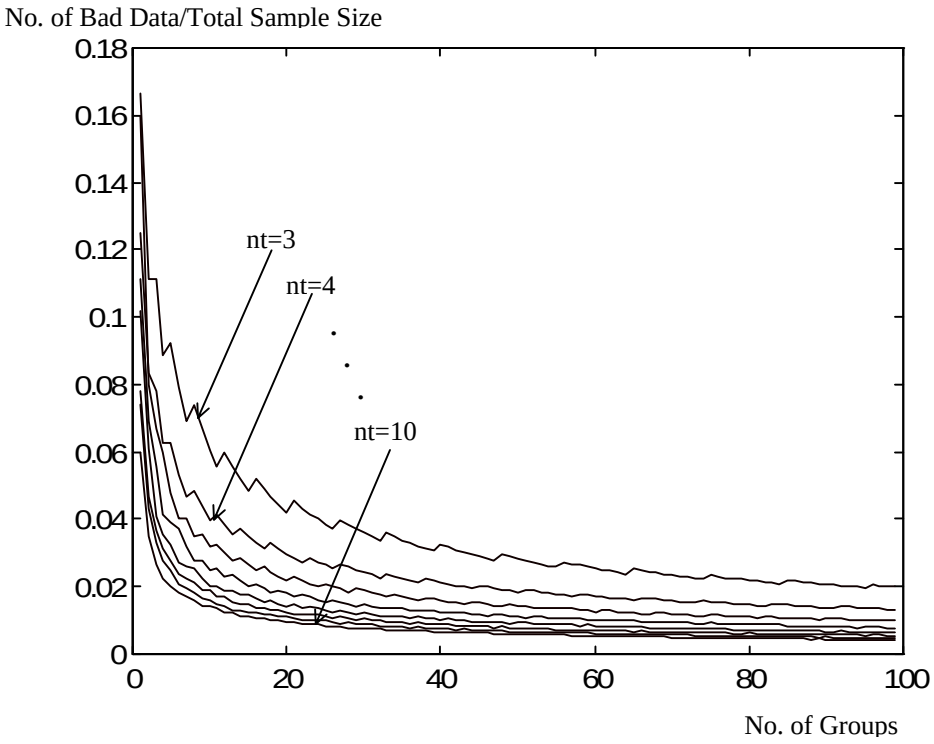

Figure 27. MRR of Friedman test for nt=3,...,10

As can be seen from Figures 26 and 27, for the same number of treatments, as the number of groups increases, the ratio between the number of bad data and the total number of the sample size becomes smaller and approaches zero. Also, as the number of treatments increases, the ratio approaches zero faster, as represented in the figures.

Let us look at a simple case where there are only three treatments and blocks in each group, that is $n_t = n_b = 3$, and also the number of groups is $n = 5$. This reduces Equations (94-98) to

$$A_1 = \left[4 - rem\left(\frac{F+4}{5}\right)\right] * \ U(F-5), \tag{99}$$

$$A_2 = \left[1 + rem\left(\frac{F+4}{5}\right)\right], \tag{100}$$

$$\sum_{i=1}^{3} R_i^2 =$$

$$\left(\left[\begin{array}{l} A_1 * \left(6 + \sum_{i=1}^{\left[\frac{F-1}{5}\right]} (3-i) * U(2-i) \right) + A_2 * \left(6 + \sum_{i=1}^{\left[\frac{F-1}{5}\right]+1} (3-i) * U(2-i) \right) + \\ (30-6F) * U(5-F) \end{array} \right]^2 + \right.$$

$$\sum_{j=2}^{4} \left(\begin{array}{l} \left(\begin{array}{l} \left(6 - \left[\frac{F-1}{5}\right]\right) * U\left(4 - \left[\frac{F-1}{5}\right] - j\right) + \\ (2+j) * U(3-j) * U\left(j - 5 + \left[\frac{F-1}{5}\right]\right) * U(4-j) \end{array} \right) * A_1 + \\ \left(\begin{array}{l} \left(5 - \left[\frac{F-1}{5}\right]\right) * U\left(3 - \left[\frac{F-1}{5}\right] - j\right) + \\ (2+j) * U(3-j) * U\left(j - 4 + \left[\frac{F-1}{5}\right]\right) * U(4-j) \end{array} \right) * A_2 + \\ (30-6F) * U(5-F) * U(3-j) \end{array} \right)^2 \left. \right) * U(10-F). \tag{101}$$

Solving for F using an iterative procedure results in Table 17.

Table 17. Friedman Test Case Results

No. of Bad Data	*Chi_squared*	*Test_statistic*	$\frac{\textit{No of Bad Data}}{\textit{Total Sample Size}}$
1	5.991492	0.4	$0.0\bar{2}$
2	5.991492	1.6	$0.0\bar{4}$
3	5.991492	3.6	$0.0\bar{6}$
4	5.991492	6.4	$0.0\bar{8}$

For $F = 4$ the test statistic $T = 6.4$ becomes greater than the $c^2_{0.95,2} = 5.99$, which results in the rejection of the Null Hypothesis (H_0). The MRR in this case is $\frac{F}{nn_t n_b} = \frac{F}{45} = 0.0\bar{8}$ as is indicated in Table 17.

5.4.3 Comparison to Other Tests

As can be concluded from Figures (26-27) and the example in the previous section, the MRR of the Friedman rank test is small. This is because to break down a test we only need to move the test statistic across the boundary of critical region, which is a fixed number and in not depended on the block size. So, as the number of measurement values grows the sample size gets larger than the number of bad data. Therefore, the ratio of MRR approaches zero. This agrees with the results reported by Coakley and Hettmansperger [53] for the two sided sign test. In terms of robustness of the classical methods, such as t-test as indicated in [53], only 4 bad observation are required to cause rejection of a 5% level 2-sided t-test, regardless of the sample size. However, for a

Friedman rank test, as can be extracted from Figure 28, for n_t = 3 and n = 10 the number of bad data are 6 and for n_t = 3 and n = 5 the number of bad data are 13.

Now we compare numerically the Friedman test with a 2-way analysis of variance. This test's procedure is as follows:

1) Calculate Total Sum of Squares (*TSS*) using

$$TSS = \sum_{i,j} \left(y_{ij} - \bar{y} \right)^2 , \tag{102}$$

where y_{ij}s are measurements at i = 1, ..., n_t and j = 1, ..., n_b, and $\bar{y}$ is the overall mean

2) Calculate variability of the treatment means ($\overline{y_i}$) from the overall mean using

$$SST = n_b \sum \left(\overline{y_i} - \bar{y} \right)^2 . \tag{103}$$

3) Calculate between treatment variability from

$$SSB = n_t \sum \left(\overline{B_j} - \bar{y} \right)^2 , \tag{104}$$

where $\overline{B_j}$ is the sample mean for block j

4) Calculate Sum of Squares Error (*SSE*) which is the variability in y_{ij}s that are not accounted for by the blocks and treatments, and is defined as

$$SSE = TSS - SST - SSB. \tag{105}$$

5) Calculate *MST* and *MSE* using

$$MST = \frac{SST}{n_t - 1}, \tag{106}$$

$$MSE = \frac{SSE}{(n_b - 1)(n_t - 1)}. \tag{107}$$

6) Obtain test statistic from

$$F\_test = \frac{MST}{MSE}. \tag{108}$$

7) Compare test statistic with *F* distribution that has

$$\alpha = 0.05,\ df_1 = n_t - 1, \text{ and } df_2 = (n_t - 1)(n_b - 1)$$

If the test statistic (*F_test*) is greater than *F* distribution then reject the null hypothesis and conclude that at least one treatment is different from the others.

Applying this test to our data resulted in the Figures 28 and 29. As can be extracted from the figures, the number of bad data that this test can handle is much smaller than the Friedman test. As an example, for the same range such as $3 \le n_t \le 10$ and $1 \le n \le 100$ the maximum outliers for two way analysis of variance is 10 and for Friedman test is 42. Therefore, the Friedman test, relatively speaking, is able to handle more contaminated data than the 2 way analysis of variance test.

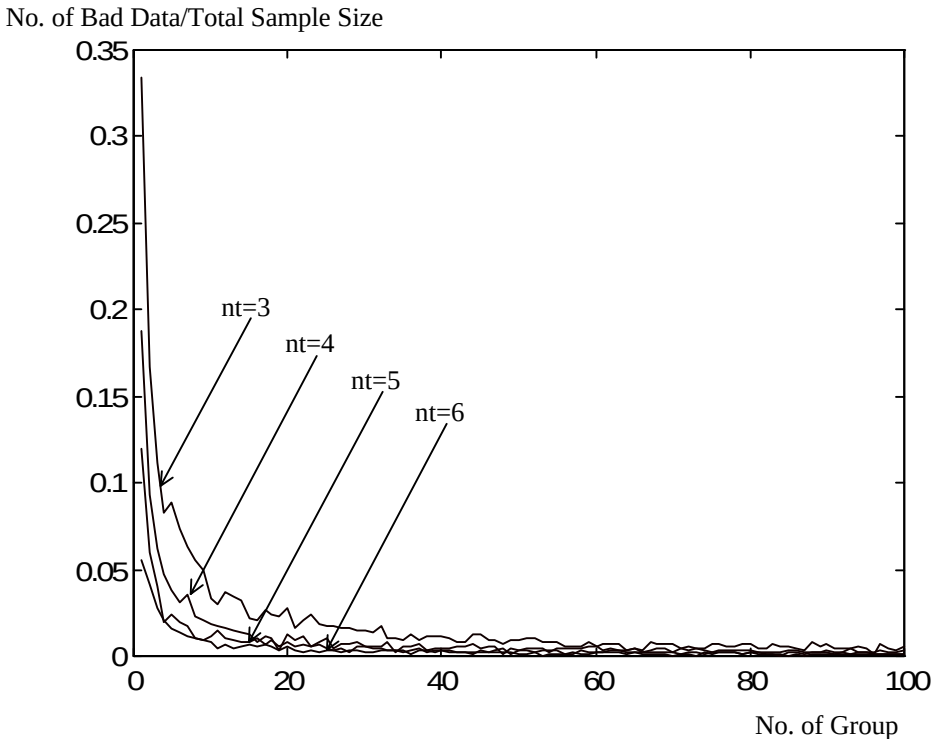

Figure 28. MRR of 2-way ANOVA for nt=3, ..., 6

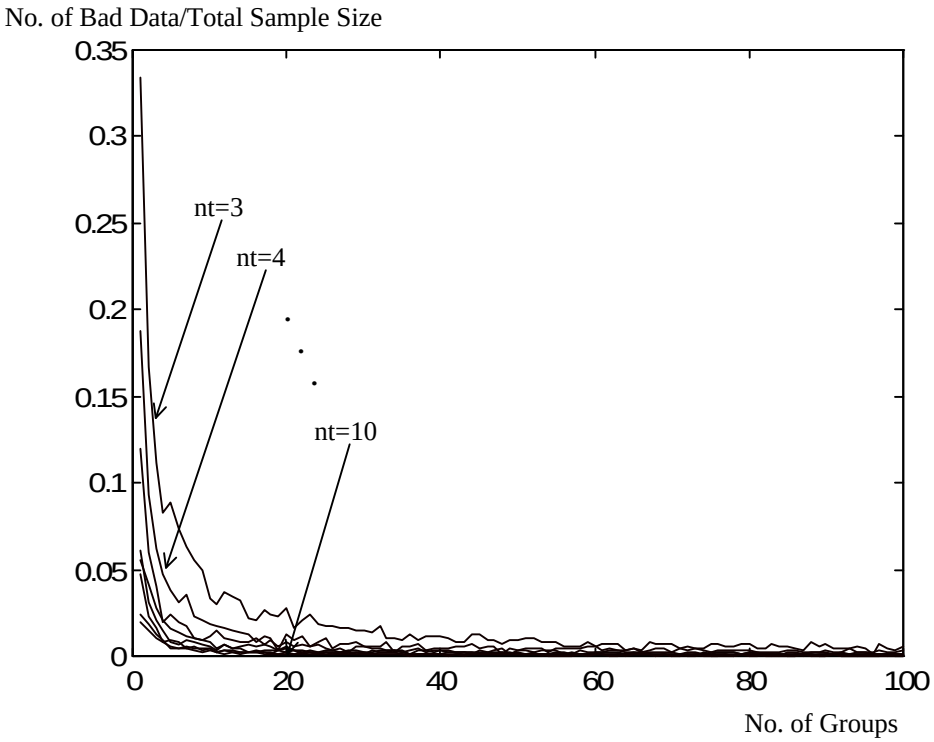

Figure 29. MRR of 2-way ANOVA for nt=3, ..., 10

CHAPTER 6

RRMC APPLIED TO TRANSFORMERS

6.1 Introduction

As was mentioned in Chapter 2, the calibration model in Equation (109) for measurements around transformers is not observable since the Jacobian matrix is rank deficient. One way to overcome this problem is to make the Jacobian matrix of full rank by estimating the offset coefficients a_i by another method. This can be done easily by recording the metered values when the bus, or the transformer, or the line is disconnected. In that case, the line values are known to be zero. The method values will be equal to a bias plus a random noise.

6.2 Algorithm

The objective, as was mentioned before, is to estimate from metered values of *P's*, *Q's* and *V's* the coefficients *a*'s, *b*'s, and *c*'s in the following equation:

$$Z = \frac{\left(a_{ip} + b_{ip}P_{ip} + c_{ip}P_{ip}^2\right)^2 + \left(a_{jp} + b_{jp}Q_{jp} + c_{jp}Q_{jp}^2\right)^2}{\left(a_{is} + b_{is}P_{is} + c_{is}P_{is}^2\right)^2 + \left(a_{js} + b_{js}Q_{js} + c_{js}Q_{js}^2\right)^2}\ , \qquad (109)$$

Where $Z = \left(\frac{V_{cp}}{\mathrm{a}\,V_{cs}}\right)^2$.The new approach uses the possibility to estimate at the control center the offset coefficients, designated by *a*'s in the above equation. This can be achieved by disconnecting the bus and the transformer and recording the meter values. We will assume in the following that the a's are available. In that case, the number of unknown

variables (*b*'s and *c*'s) is reduced to 8, and the Jacobian matrix given by 109 becomes full rank. In other words, the model (109) becomes observable and the coefficients are identifiable.

6.3 Procedure

The estimation method that will be used, as was mentioned in Chapter 3, needs to be robust and not sensitive to the presence of outliers. Again, as in Chapter 3, our choice is the least median of squares estimator. The relationship (109) yields a nonlinear regression model that can be put in a matrix form given by

$$\underline{z} = \underline{h}(\underline{x}) + \underline{e}. \tag{110}$$

The observation vector $\underline{z}$ can be regarded as a function of 8 parameters contained in $\underline{x} = \left[b_{pi}, c_{pi}, \ldots, b_{sj}, c_{sj}\right]$. Its values come from the ratio between the voltages in the primary and the secondary sides of the transformer, which are calibrated using a comparative calibration method as mentioned in Chapter 5. Suppose that *m* values of *Z* are calculated from *m* snapshots. Let Z_k be the value of *Z* obtained at the *k*th snapshot. Assume that we have *m* snapshots and let $\underline{z} = \left[Z_1, Z_2, \ldots, Z_m\right]^T = \underline{h}(\underline{x})$. A first order Taylor series expansion of $\underline{z}$ about an operating point yields

$$\Delta\underline{z} = \underline{\underline{H}}\Delta\underline{x}\ , \tag{111}$$

where $\underline{\underline{H}} = \dfrac{\P\,\underline{h}(\underline{x})}{\P\,\underline{x}}$ is the $(m \times 8)$ Jacobian matrix given by

$$\underline{\underline{H}} = \left[\frac{\P Z}{\P b_{ip}}, \frac{\P Z}{\P c_{ip}}, \frac{\P Z}{\P b_{jp}}, \frac{\P Z}{\P c_{jp}}, \frac{\P Z}{\P b_{is}}, \frac{\P Z}{\P c_{is}}, \frac{\P Z}{\P b_{js}}, \frac{\P Z}{\P c_{js}}\right]. \tag{112}$$

The partial derivative equations are

$$\frac{\P Z}{\P b_{ip}} = \frac{2P_{ip}\left(a_{ip} + b_{ip}P_{ip} + c_{ip}P_{ip}^2\right)}{B}, \tag{113}$$

$$\frac{\P Z}{\P c_{ip}} = \frac{2P_{ip}^2\left(a_{ip} + b_{ip}P_{ip} + c_{ip}P_{ip}^2\right)}{B}, \tag{114}$$

$$\frac{\P Z}{\P b_{jp}} = \frac{2Q_{jp}\left(a_{jp} + b_{jp}Q_{jp} + c_{jp}Q_{jp}^2\right)}{B}, \tag{115}$$

$$\frac{\P Z}{\P c_{jp}} = \frac{2Q_{jp}^2\left(a_{jp} + b_{jp}Q_{jp} + c_{jp}Q_{jp}^2\right)}{B}, \tag{116}$$

$$\frac{\P Z}{\P b_{is}} = \frac{-2P_{is}\left(a_{is} + b_{is}P_{is} + c_{is}P_{is}^2\right)A}{B^2}, \tag{117}$$

$$\frac{\P Z}{\P c_{is}} = \frac{-2P_{is}^2\left(a_{is} + b_{is}P_{is} + c_{is}P_{is}^2\right)A}{B^2}, \tag{118}$$

$$\frac{\P Z}{\P b_{js}} = \frac{-2Q_{js}\left(a_{js} + b_{js}Q_{js} + c_{js}Q_{js}^2\right)A}{B^2}, \tag{119}$$

$$\frac{\P Z}{\P c_{js}} = \frac{-2Q_{js}^2\left(a_{js} + b_{js}Q_{js} + c_{js}Q_{js}^2\right)A}{B^2}, \tag{120}$$

where

$$A = \left(a_{ip} + b_{ip}P_{ip} + c_{ip}P_{ip}^2\right)^2 + \left(a_{jp} + b_{jp}Q_{jp} + c_{jp}Q_{jp}^2\right)^2, \tag{121}$$

$$B = \left(a_{is} + b_{is}P_{is} + c_{is}P_{is}^2\right)^2 + \left(a_{js} + b_{js}Q_{js} + c_{js}Q_{js}^2\right)^2. \tag{122}$$

$\Delta\underline{x}$ is an *8*-dimensional vector given by

$$\Delta\underline{x} = \left[\Delta b_{ip}, \Delta c_{ip}, \Delta b_{jp}, \Delta c_{jp}, \Delta b_{is}, \Delta c_{is}, \Delta b_{js}, \Delta c_{js}\right]^T, \tag{123}$$

and the residual vector is written as

$$\underline{r} = \underline{z} - \underline{\underline{H}}\underline{x}. \tag{124}$$

The LMS estimator is defined as the minimizer of the nth ordered squared weighted residuals, yielding the objective function given by

$$J(\underline{x}) = r_{w(\mathsf{n})}^2, \tag{125}$$

where

$$\mathsf{n} = \left[\frac{m}{2}\right] + \left[\frac{n+1}{2}\right]. \tag{126}$$

Here, *m* denotes the number of measurements, *n* denotes the number of parameters to be estimated. The minimum of (125) is found through the resampling technique [3,4]. It will be described next.

6.3.1 The Resampling Technique

Although the LMS has many advantages, it is a computationally intensive method. Its computation involves the selection of different measurement subsets of size n followed by the calculation of $\underline{x}$ and $J(\underline{x})$. Due to the relatively large size of the measurement set, selection of all possible subsets of size n is not advisable or feasible. Therefore, the resampling technique has been advocated in [3,4] as a more efficient way to minimize (125). The following section will outline this technique.

In principle, the combinatorial search includes $\binom{m}{n}$ measurement subsets, where

$$\binom{m}{n} = \frac{m!}{n!(m-n)!} \ . \tag{127}$$

This is a large number, even for small values of m and n. The resampling technique is used to reduce the number of measurement subsets to be investigated. One can assume that each measurement has an equal chance of being selected and included in a subset. One can also consider a certain number k of random selections such that the probability P of obtaining at least one uncontaminated subset is high. A typical value of P is 0.95. This probability is given by

$$P = 1 - (1 - (1 - \varepsilon)^n)^k \ , \tag{128}$$

where ε is the assumed fraction of contaminated data among the m observations. This is the ratio of the number of bad data to the total number of measurements against which we want to be protected.

From (128), it follows that

$$k = \frac{\ln(1-P)}{\ln(1-(1-\varepsilon)^n)} \ . \tag{129}$$

Table 18 reports values of k for P = 0.95, n = 8, m = 1000 and for different fraction of contamination ε. Since we usually need to have more than one good sample, we use $2k$ instead of k samples. There is a dramatic decrease in the number of subsets to be investigated when using the resampling technique since the total number of subsets is equal to $\binom{m}{n}$ is 1.8609e+11 in this case.

Table 18. Resampling Technique

ε	k	$2k$
0.01	1	2
0.05	3	6
0.1	5	10
0.15	9	18
0.2	16	32

6.4 Simulation Results

The following is a detailed description of the simulation. A set of *a*'s, *b*'s, and *c*'s is pre-selected and values for P_p, Q_p, P_s, and Q_s are drawn randomly from the standard normal distribution *N(0,1)* using the computer's internal clock as a seed number for a Gaussian random generator function. These values are substituted in (107) to obtain the vector $\underline{Z}$. Then, random noises following a normal distribution *N(0,0.01)* are added to all values of the explanatory and response variables. Some of the measurement values have intentionally been changed to values far from the original one, in order to examine the effect of outliers on the parameter estimates. Now, by knowing P_p, Q_p, P_s, and Q_s, $\underline{Z}$, and *a*'s, and by using randomly-generated initial conditions for *b*'s, and *c*'s, we estimate these parameters by means of a two-step robust estimation method. First, using the least median of squares estimator, we identify the outliers using the technique described in Chapter 3 and 6, then after removing the outliers, we run the least squares estimator on the remaining measurement set using a Monte Carlo simulations. The histogram of the estimated values for each of the parameters *b*'s and *c*'s have been obtained and displayed in Figures 30 through 37. It can be seen that none of them follow a Gaussian distribution; therefore, the medians and not the means of all these estimates are compared to the true values. The results is summarized in Table 19.

Table 19. Robust Estimation Results Based on LMS

	b_{ip}	c_{ip}	b_{jp}	c_{jp}	b_{is}	c_{is}	b_{js}	c_{js}
True	0.9600	0.0100	0.9800	0.0200	1.0100	0.0150	1.2000	0.0180
Estimate	0.9588	0.0108	0.9800	0.0197	1.0061	0.0181	1.1892	0.0225

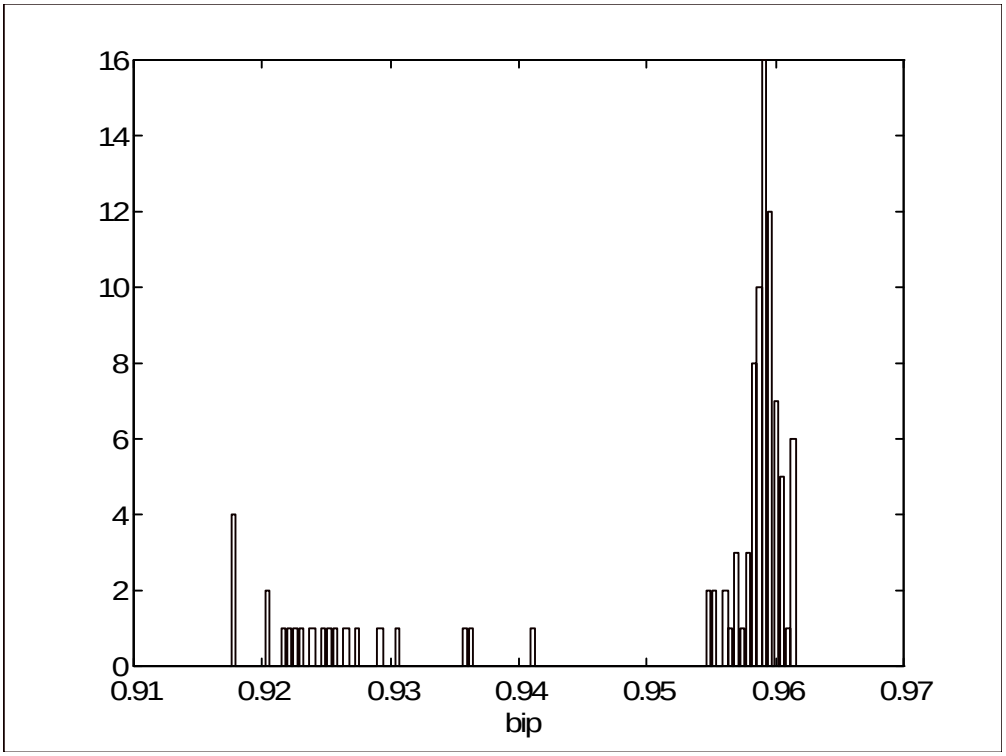

Figure 30. Histogram of Estimated b_{ip}

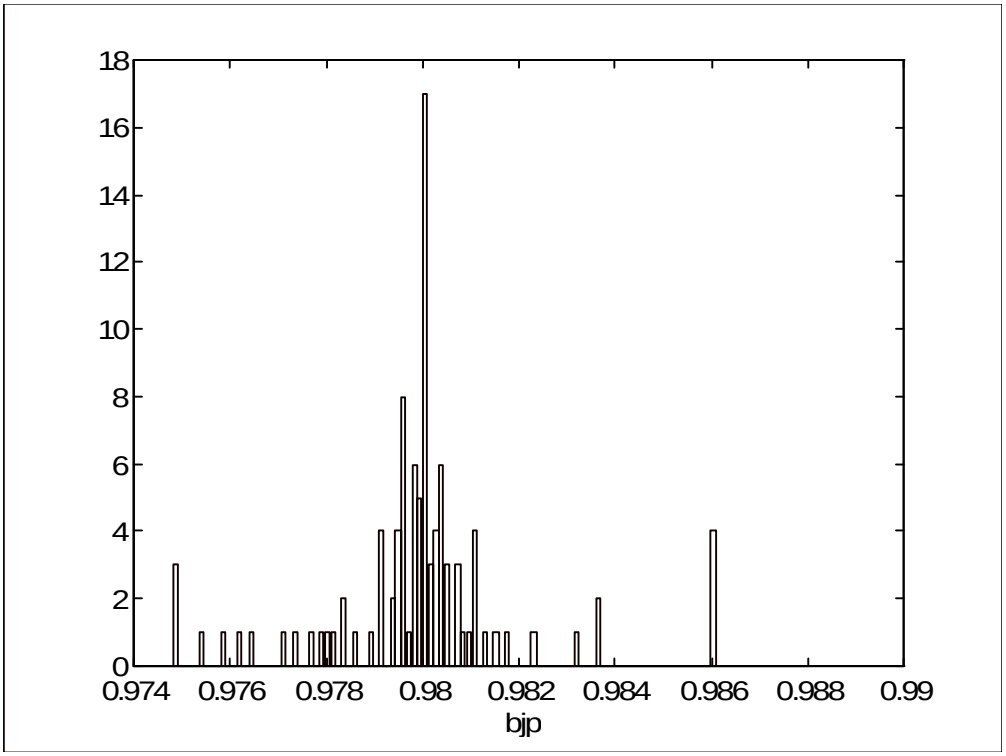

Figure 31. Histogram of Estimated b_{jp}

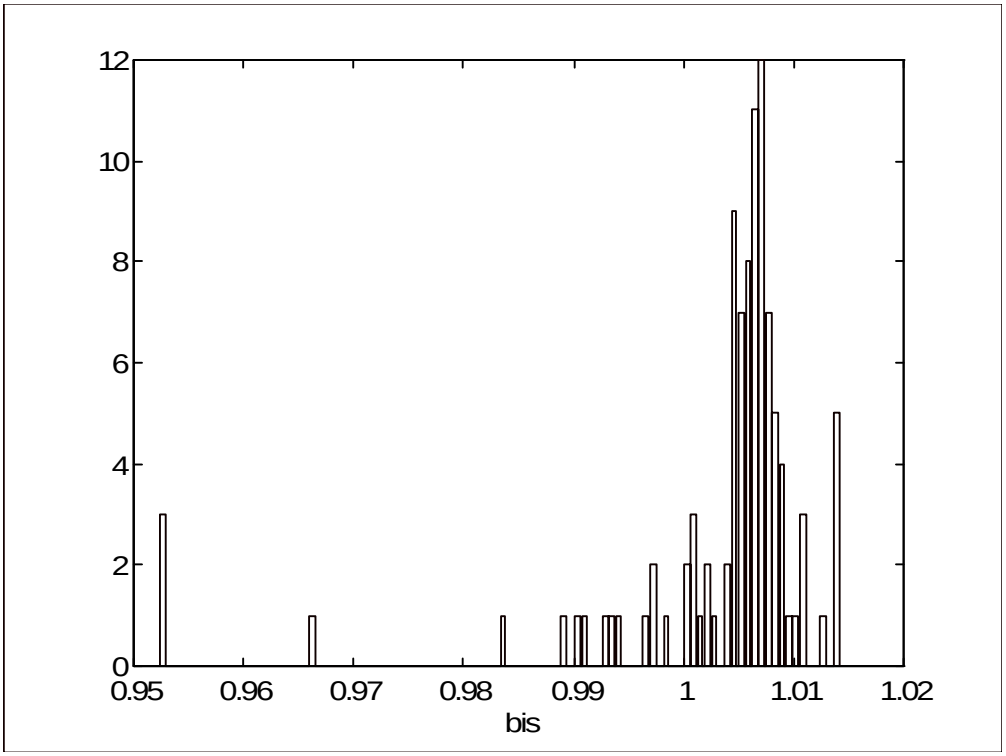

Figure 32. Histogram of Estimated b_{is}

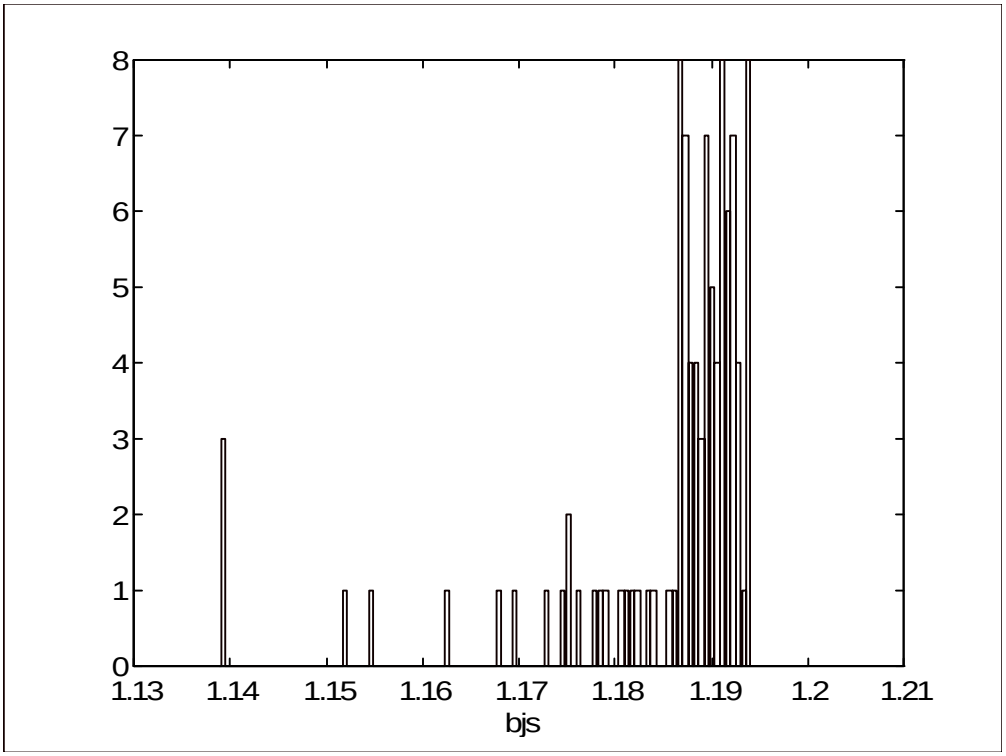

Figure 33. Histogram of Estimated b_{js}

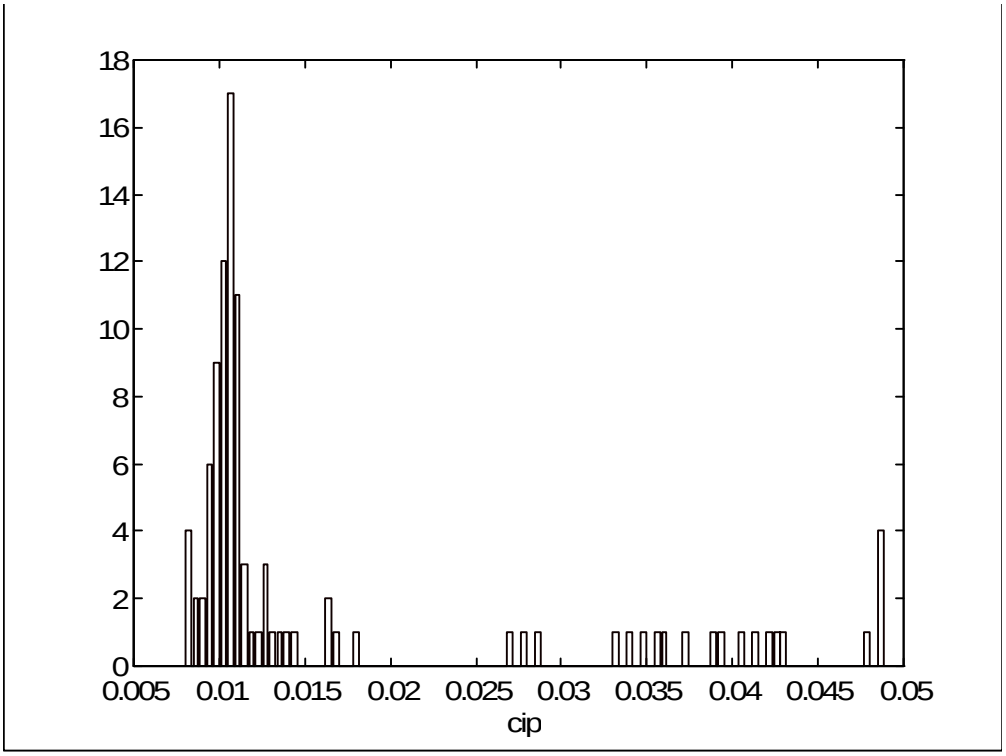

Figure 34. Histogram of Estimated c_{ip}

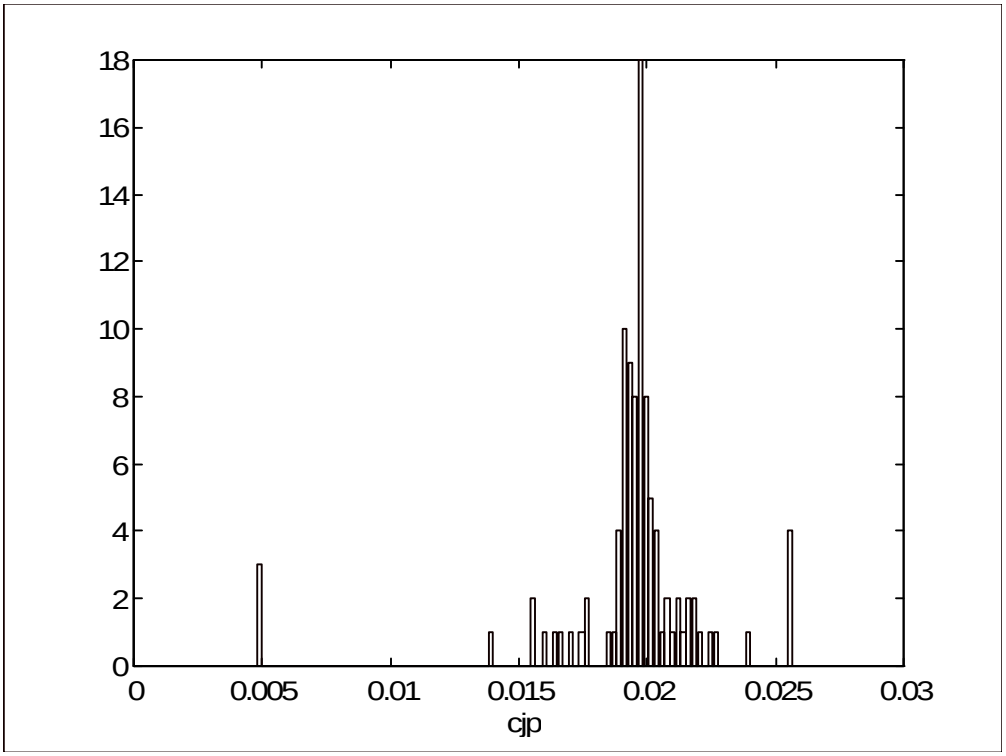

Figure 35. Histogram of Estimated c_{jp}

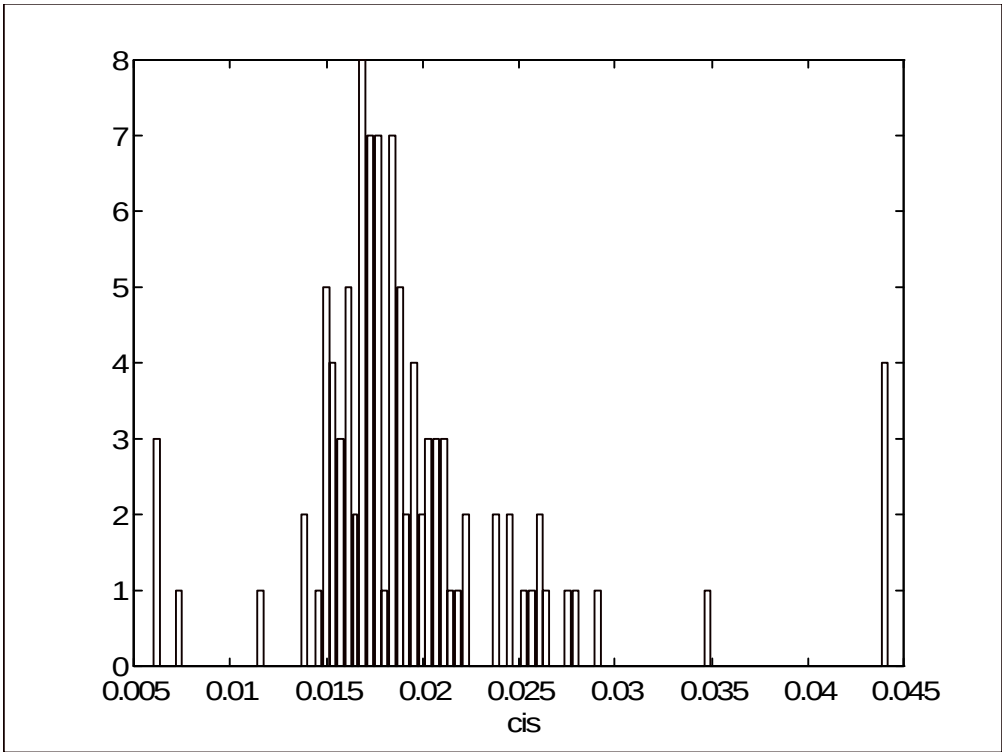

Figure 36. Histogram of Estimated c_{is}

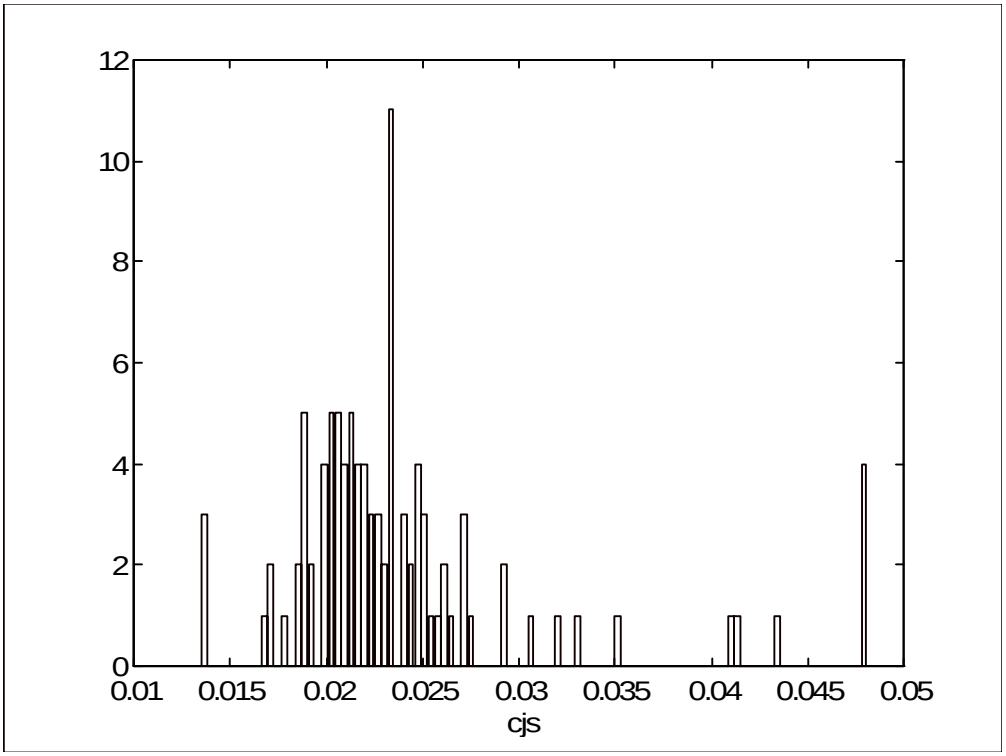

Figure 37. Histogram of Estimated c_{js}

Figures 38 and 39 are the graphical representations of the measured vs. calibrated real powers in the primary and secondary side of the transformer, respectively. The continuous line is the graph of the estimated calibration curve vs. the metered values with random outliers removed. The '*' are the data points whose coordinates are the metered values with random outliers and their associated calibrated values using the true parameters. Formally, we have

$$\hat{calibrated} = \hat{a} + \hat{b}(measured) + \hat{c}(measured)^2 \qquad (130)$$

The results as can be seen from the graphs are satisfactory.

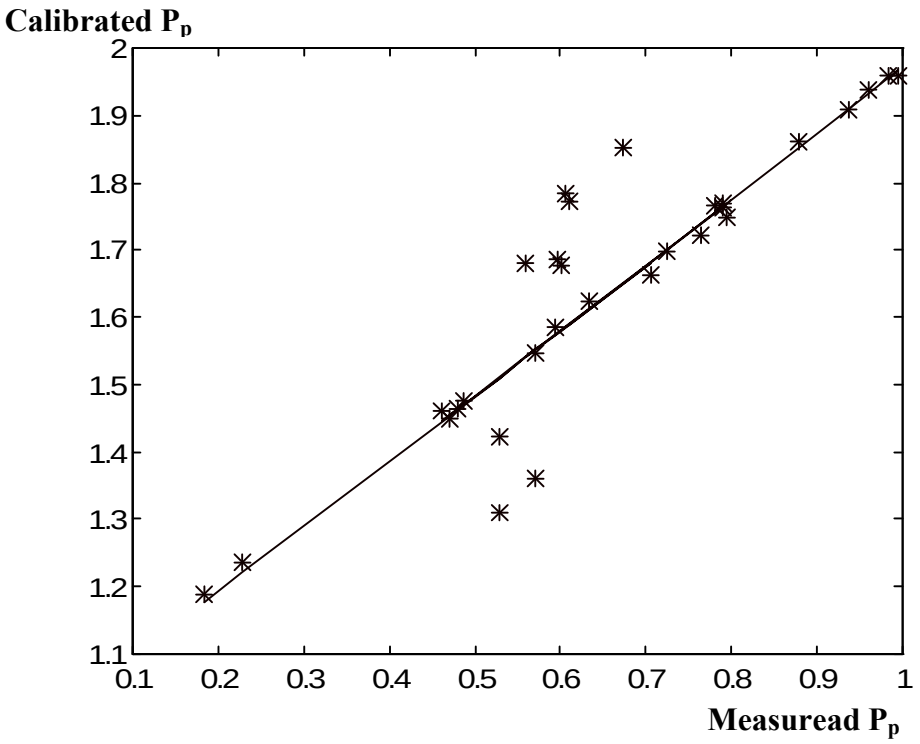

Figure 38. Measured Vs. Calibrated P_p

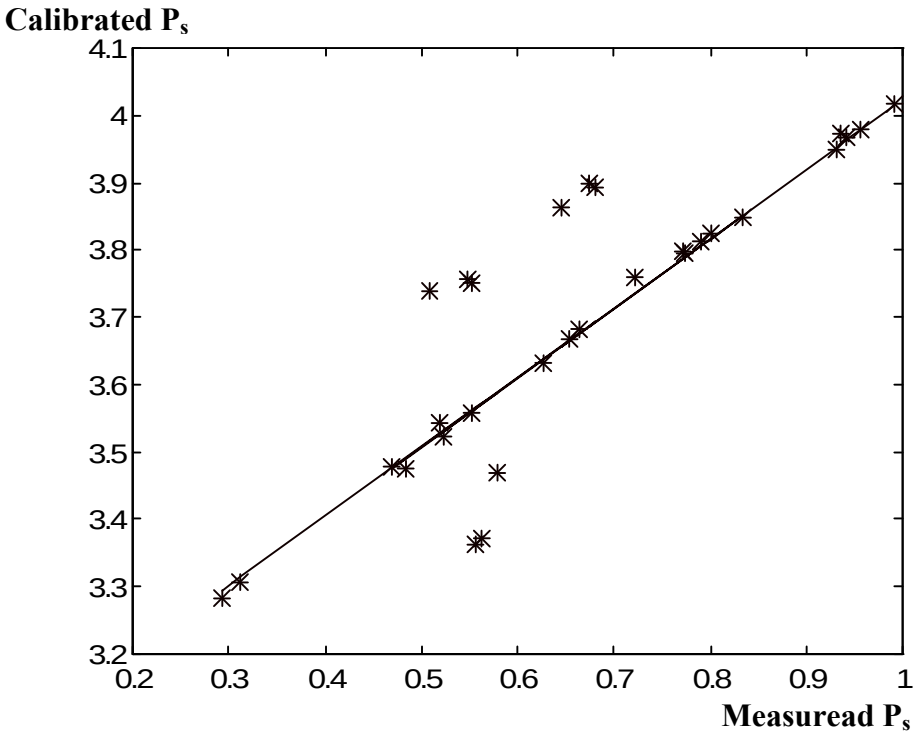

Figure 39. Measured Vs. Calibrated P_s

CHAPTER 7

CONCLUSION AND FUTURE WORK

A new method for remote measurement calibration has been developed. The objective of this method is to minimize the systematic errors of the measurements through an appropriate scaling procedure. The approach provides a significant decrease in the time execution and the cost of calibration. The main feature of the approach is that about 40% of the RTUs need to be field calibrated, and the remaining 60% are remotely calibrated. The candidate RTUs for field calibration are identified through a simulated annealing method. It is made applicable to large systems as well, by means of system tearing and dynamic programming. The number of field calibrations can further be reduced by using the availability of multiple voltage measurements at the same buses in conjunction with the remote estimation of the measurement offset biases when the equipments are not energized. To this end, different comparative calibration methods were investigated. The robustness characteristics of these methods in terms of dealing with outliers were explored. Because, the distribution of our data is unknown, a modified version of the Friedman test was developed. This test is a nonparametric test since it does not assume any a priori distribution. In addition, it is robust against outliers. An expression for calculating the maximum number of bad data that it can handle has been derived. It allows us to study its Maximum Resistance to Rejection (MRR). Comparisons of this MRR with those of other parametric tests revealed that the Friedman test can handle relatively more bad data.

As future work and to further improve the method, six improvements are proposed. First, it is suggested to use Phasor Measurement Units (PMUs) instead of the Remote Terminal Units (RTUs). PMUs deal with unbalanced systems by measuring the

three phases and calculating the positive sequence components. Also, unlike the RTUs, which are already placed in the system, the PMUs can be located appropriately in such a way that the required number of units for system observability is minimized. This in turn will further alleviate the field calibration endeavor. Another advantage of the PMUs is that they measure all the currents at the bus where they are located. These current measurements are of much better quality than the voltage magnitude measurements because they are based on magnetically coupled current transformers. The latter provide values to within 1% of the true values, whereas the capacitive coupled voltage transformers introduce 5 to 7% error in the voltage magnitude at the secondary winding. The major problem with PMUs is that they are expensive.

As a second improvement to the proposed approach, we may investigate an alternative method to simulated annealing, such as Tabu method. This is because the simulated annealing is a rather slow technique and its tuning takes time, requires expertise, and is case dependent. To further decrease the time of execution parallel processing methods may be used after tearing the large system into smaller subsystems. At the present time simulated annealing runs on each subnetwork individually one at a time. This is not efficient for large systems. However, parallel processing if adapted can speed up the process.

As a future work it would be interesting to investigate in addition to the MRR, the Maximum Resistance to Acceptance (MRA) of the Friedman rank test. MRA in conjunction with the Maximum Resistance to Rejection (MRR) define the upper and the lower limit for resistance to outliers. At this point, all our studies are based on simulated data. As a future work, it would be good to validate the RRMC method with real data recorded from large system. Also, an economical assessment of the method needs to be considered in order to show the cost saving provided by measurement calibration.

REFERENCES

[1] L. S. Van Slyck, J. Allemong, "Operating Experience With the AEP State Estimation", *IEEE Transaction on Power Systems*, Vol. 3, No. 2, May 1988, pp. 521-528.

[2] L. Mili, V. Phaniraj, P. J. Rousseeuw, "Robust Estimation Theory for Bad Data Diagnostics in Electrical Power System", in *Control and Dynamic Systems: Advances in Theory and Applications, Vol. 37: Advances in Industrial Systems*, edited by C. T. Leondes, Academic Press, 1990.

[3] L. Mili, V. Phaniraj, P. J. Rousseeuw, "Least Median of Squares Estimation in Power Systems", *IEEE Trans. Power Syst.* Vol. 6, No. 2, May 1991, pp. 511-523.

[4] P. J. Rousseeuw, A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley, 1987.

[5] F. R. Hampel, E. W. Ronchetti, P. J. Rousseeuw, W. A. Stahel, *Robust Statistics: the Approach Based on Influence Functions*, John Wiley, 1986.

[6] M. M. Adibi, D. K. Throne, "Remote Measurement Calibration", *IEEE Transactions on Power Systems*, Vol. PWRS-1, No.2, May 1986, pp. 194-203.

[7] M. M. Adibi, J. P. Stovall, "On Estimation of Uncertainties in Analog Measurements", *IEEE Transactions on Power Systems*, Vol. 5, No. 4, Nov. 1990, pp. 1222-1230.

[8] M. M. Adibi, R. J. Kafka, "Minimization of Uncertainties in Analog Measurements for use in State Estimation", *IEEE Transactions on Power Systems*, Vol. 5, No. 3, August 1990, pp. 902-910.

[9] M. M. Adibi, R. J. Kafka, K. A. Clements, J. P. Stovall, "Integration of Remote Measurement Calibration With State Estimation- A Feasibility Study", Paper No.

90 SM 274-1 PWRS, IEEE/PES Summer Meeting, Minneapolis, Minnesota, July 15-19, 1990.

[10] G. R. Krumpholtz, K. A. Clements, P. W. Davis, "Power System Observability: A Practical Algorithm Using Network Topology", *IEEE Transaction on Power Apparatus and Systems*, Vol. PAS-99, No. 4, July/August 1980.

[11] R. Tarjan, "Depth-First Search and Linear Graph Algorithms", *SIAM Journal of Comput*., Vol. 1, No. 2, pp. 146-160, June 1972.

[12] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, "Equation of State Calculations by Fast Computer Machines", *J. Chem. Phys.*, Vol. 21, 1953, pp. 1087-1092.

[13] K. A. Clements, B. F. Wollenberg, "An Algorithm for Observability Determination in Power System State Estimation", IEEE PES Summer Meeting, Paper A 75 447-3, San Francisco, July 1975.

[14] K. A. Clements, G. R. Krumpholtz, P. W. Davis, "Power System State Estimation With Measurement Deficiency: An Algorithm That Determines the Maximal Observable Subnetwork", *IEEE Transaction on Power Apparatus and Systems*, Vol. PAS-101, No. 9, pp. 3044-3052, September 1982.

[15] A. Monticelli, F. F. Wu, "Network Observability: Theory", *IEEE Transaction on Power Apparatus and Systems*, Vol. PAS-104, No. 5, pp. 1042-1048, May 1985.

[16] A. Monticelli, F. F. Wu, "Network Observability: Identification of Observable Islands and Measurement Placement", *IEEE Transaction on Power Apparatus and Systems*, Vol. PAS-104, No. 5, pp. 1035-1041, May 1985.

[17] F. F. Wu, A. Monticelli, "Observability Analysis and Bad Data Processing for State Estimation Using Hachtel's Augmented Matrix Method", *IEEE Transaction on Power Apparatus and Systems*, Vol. PWRS-3, pp. 604-611, May 1988.

[18] R. Nucera, M. L. Gilles, "Observability Analysis: A New Topological Algorithm", *IEEE Transaction on Power Systems*, Vol. 6, No. 2, May 1991, pp. 466-475.

[19] K. A. Clements, G. R. Krumpholz, P. W. Davis, "Power System State Estimation Residual Analysis: An Algorithm Using Network Topology", *IEEE Transaction on Power Apparatus and Systems*, Vol. PAS-100, No. 4, April 1981, PP. 1779-1787.

[20] I. O. Bohachevsky, M. E. Johnson, M. L. Stein, "Generalized Simulated Annealing for Function Optimization", *Technometrics*, Vol. 28, 1986, pp. 209-217.

[21] S. Kirkpatrick, C. D. Gellatt jr., M. P. Veechi, "Optimization By Simulated Annealing", *Science*, Vol. 220, 1983, pp. 671-680.

[22] S. Kirkpatrick, "Optimization By Simulated Annealing: Quantitative Studies", *J. Stat. Phys.*, Vol. 34, 1984, pp. 975-986.

[23] F. Romeo, A. Sangiovanni-Vincentelli, "A Theoretical Framework for Simulated Annealing", *Algorithmica*, Vol. 6, 1991, pp. 302-345.

[24] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press.

[25] F. Zhuang, F. D. Galiana, "Unit Commitment by Simulated Annealing", *IEEE Transactions on Power Systems*, Vol. 5, No. 1, pp. 311-318, May 1990.

[26] B. Scott, "Review of Load-Flow Calculation Methods", *Proceeding of the IEEE*, Vol. 62, No. 7, July 1974, pp. 916-929.

[27] T. L. Baldwin, L. Mili, M. B. Boisen Jr., R. Adapa, "Power System Observability With Minimal Phasor Measurement Placement", *IEEE/PES* 1992 Summer Meeting, Seattle, WA, July 12-16, 1992.

[28] L. Mili, T. Baldwin, R. Adapa, "Phasor Measurement Placement for Voltage Stability Analysis of Power Systems", *Proceeding of the 29th IEEE conference on Decision and Control*, Honolulu, HI, Dec. 5-7, 1990.

[29] M. Merill, F. C. Schweppe, "Bad Data Suppression in Power System Static State Estimator", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-90, 1971, pp. 2718-2725.

[30] M. R. Irving, R. C. Owen, M. Sterling, "Power System State Estimation Using Linear Programming", *Proceeding of the IEEE*, Vol. 125, No. 9, Sept. 1978, pp. 879-885.

[31] J. V. Beck, K. J. Arnold, *Parameter Estimation in Engineering and Sciences*, John Wiley, 1977.

[32] R. G. Staudte, S. J. Sheater, *Robust Estimation and Testing*, John Wiley, 1990.

[33] E. L. Lehmann, *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, 1975.

[34] K. A. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, Second edition, John Wiley, 1965.

[35] N. Christofides, *Graph Theory : An Algorithmic Approach*, Academic Press, 1975.

[36] D. J. A. Welsh, *Matroid Theory*, Academic Press, 1976.

[37] M. G. Cheniae, *Observability Method for the Least Median of Squares Estimator as Applied to Power System*, Master Thesis, VPI & SU, June 1991.

[38] P. J. Huber, *Robust Statistics*, John Wiley, 1981

[39] P. J. Rousseeuw, "Least Median of Squares Regression", *Journal of the American Statistical Association*, Vol. 79, No. 388, 1984, pp. 871-880.

[40] P. Bonanomi, G. Gramberg, "Power System Data Validation and State Calculation By Network Search Techniques", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 1, January 1983.

[41] N. I. Deeb, S. M. Shahidehpour, "Calibration of Measurement Sets In An Electric Power System Using An Updated Broyden Method", *Proceedings of the American Power Conference*, 1991.

[42] Leon Cooper, Mary W. Cooper, *Introduction to Dynamic Programming*, Pergamon Press, 1981.

[43] Stuart E. Dreyfus, Averill M. Law, *The Art and Theory of Dynamic Programming*, Academic Press, 1977.

[44] D. J. White, *Dynamic Programming*, Oliver and Boyd Ltd. and Holden-day Inc. , 1969.

[45] Richard Bellman, *Introduction to Matrix Analysis*, McGraw Hill, 2nd ed., 1970.

[46] Richard Bellman, Kenneth L. Cooke, Jo Ann Lockett, *Algorithms Graphs and Computers*, Academic Press, 1970.

[47] Mahendra K. Jain, N. Dharma Rao, "A Power System Networks Decomposition Algorithm for Network Solutions ", *IEEE PES Summer Meeting,* San Francisco, Calif. July 9-14, 1972.

[48] J. M. Undrill, H. H. Happ, "Automatic Sectionalization of Power System Network for Network Solutions", *IEEE Transactions of Power Apparatus and Systems,* Vol. PAS-90, No. 1, January/ February 1971

[49] John W. Pratt, Jean D. Gibbons, *Concepts of Nonparametric Theory*, Springer-Verlag New York Inc., 1981.

[50] Jason C. Hsu, *Multiple Comparison Theory and Methods*, Chapman and Hall, 1996.

[51] Wayne W. Daniel, *Applied Nonparametric Statistics*, Boston: PWS-Kent. 2nd ed., 1990.

[52] Clint W. Coakley, Thomas P. Hettmansperger, "Breakdown Bounds and Expected Test Resistance", *Nonparametric Statistics,* Vol. 1, pp 267-276, Gordon and Breach Science Publishers, 1992.

[53] Clint W. Coakley, Thomas P. Hettmansperger, "The Maximum Resistance of Tests", *Austral. J. Statist,* 36(2), pp 225-233, 1994.

[54] Hany Zayed, Dana Quade, "A Note on The Distribution of The Resistance of Rank Tests", *Commun. Statist, -Theory Meth.,* 26(8), pp 1867-1875, 1997.

[55] Hany Zayed, Dana Quade, "On The Resistance of Rank Correlation", *J. Statist Comput. Simul. ,* Vol. 58, pp 59-81, 1997.

[56] D. L. Hillhouse, O. Petersons, W. C. Sze, "A Prototype System For On-Site Calibration of Coupling Capacitor Voltage Transformers (CCVTs)", *IEEE Transactions on Power Apparatus ans Systems,* Vol. PAS-98, No. 3, May/ June 1979.

[57] M. R. Irving, M. J. H. Sterling, "Optimal Network Tearing Using Simulated Annealing", *IEE Proceedings,* Vol. 137, Pt. C, No. 1, January 1990

[58] M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance", *Journal of the American Statistical Association,* 32, 675-701, 1937.

[59] M. Hollander, D. A. Wolfe, *Nonparametric Statistical Methods*, New York: Jonh Wiley and Sons, Inc. , 1973.

[60] B. J. McDonald, W. A. Thompson, "Rank Sum Multiple Comparisons in one- and two-way Classifications", *Biometrika,* 54, 487-497, 1967.

VITA

Alireza Ghassemian was born in Abadan, Iran on July 5, 1962.

He received his B.S. and M.S. degrees in Electrical Engineering from University of Maryland at College Park, MD. and Virginia Polytechnic Institute and State University at Blacksburg, VA. in 1989 and 1995 respectively.