

METHODS OF HANDLING MISSING DATA IN  
ONE SHOT RESPONSE BASED POWER SYSTEM CONTROL

A Thesis

Submitted to the Faculty

of

Purdue University

by

Niraj Dahal

In Partial Fulfillment of the

Requirements for the Degree

of

Masters of Science in Electrical and Computer Engineering

August 2019

Purdue University

Indianapolis, Indiana

**THE PURDUE UNIVERSITY GRADUATE SCHOOL**  
**STATEMENT OF THESIS APPROVAL**

Dr. Steven Rovnyak, Chair

Department of Electrical and Computer Engineering

Dr. Lingxi Li

Department of Electrical and Computer Engineering

Dr. Euzeli Dos Santos

Department of Electrical and Computer Engineering

**Approved by:**

Dr. Brian King

Head of Graduate Program

## ACKNOWLEDGMENTS

First of all, I would heartily like to express appreciation and gratitude to my thesis advisor Dr. Steven Rovnyak. He has been more than an advisor to me who has always stood up and advocated for my academic well being at IUPUI. Since my enrollment in MSECE, Dr. Rovnyak has supported me with graduate assistantship giving an opportunity to work in the research about response based one shot control in power system. This opportunity has apparently uplifted my knowledge, skills and capacity and sowed a new aspiration within me.

I am also grateful to my advisory committee members Dr. Dos Santos and Dr. Lingxi Li for their time and support throughout the thesis completion process.

I am thankful to Sherrie Tucker for her help though-out my graduate career, and for all the reminders to help me stay on top of deadlines.

I am thankful to Shahrzad Iranmanesh for providing a new set of one shot control combination from her unpublished work to be used in the thesis.

Lastly, I would like to thank my wife Chahana and entire family who have supported me throughout my academic pursuits and are committed to support further for my PhD at IUPUI.

## TABLE OF CONTENTS

	Page
LIST OF TABLES . . . . .	vi
LIST OF FIGURES . . . . .	vii
ABBREVIATIONS . . . . .	ix
ABSTRACT . . . . .	x
1 INTRODUCTION . . . . .	1
1.1 Problem Statement . . . . .	1
1.2 Power System Stability and Control . . . . .	2
1.2.1 Power System Transient Stability . . . . .	3
1.2.2 Improvement of Transient Stability . . . . .	6
1.3 Tools used for the Thesis . . . . .	6
1.3.1 Transient Security Assessment Tool (TSAT) . . . . .	7
1.3.2 PowerFlow and Short Circuit Analysis Tool (PSAT) . . . . .	7
1.3.3 MATLAB . . . . .	7
2 RESPONSE BASED EVENT DETECTION FOR ONE SHOT STABILITY CONTROLS . . . . .	8
2.1 Overview . . . . .	8
2.2 Power System Model . . . . .	9
2.3 Phasor Measurements . . . . .	10
2.4 One Shot Response Based Control Combination . . . . .	12
2.5 Training and Test Data Sets for Decision Trees (DTs) . . . . .	14
2.6 Indices Used for Event Detection and One Shot Control . . . . .	15
2.6.1 Integral Square Generator Angle (ISGA) . . . . .	15
2.6.2 Square Bus Angle (SBA) Index . . . . .	16
2.6.3 Bus Magnitude Index . . . . .	16

	Page	
2.7	Classification of Event as ‘Stable’ or ‘Unstable’ . . . . .	17
2.8	DTs for Response based Control . . . . .	18
2.8.1	Training DT for One Shot Control . . . . .	19
2.8.2	Training DT for Event Detection . . . . .	20
2.9	Control Methodology for Test sets . . . . .	22
2.10	Performance Metrics for One Shot Control . . . . .	23
2.11	Simulation Results with Event Detection . . . . .	24
3	PROPOSED METHODS FOR MISSING DATA HANDLING . . . . .	25
3.1	Overview . . . . .	25
3.2	Missing Data Handling Techniques . . . . .	26
3.2.1	Data Hold . . . . .	27
3.2.2	Lagrange Interpolating Polynomial Method . . . . .	30
3.2.3	Bootstrap Averaging . . . . .	33
3.3	Generation of Pseudo-random Numbers . . . . .	34
4	RESULTS AND ANALYSIS . . . . .	36
4.1	Overview . . . . .	36
4.2	Performance using Zero Order Hold (ZOH) . . . . .	37
4.3	Performance using First Order Hold (FOH) . . . . .	39
4.4	Performance using Lagrange Polynomial Method . . . . .	41
4.4.1	Extrapolation technique for missing data recovery . . . . .	41
4.4.2	Interpolation technique for missing data recovery . . . . .	42
4.5	Performance using Bootstrap Averaging Method . . . . .	47
4.6	Results with Control Combination CC2 . . . . .	49
5	CONCLUSION AND FUTURE ENHANCEMENT . . . . .	51
	REFERENCES . . . . .	53
	APPENDIX 1 . . . . .	55

## LIST OF TABLES

Table	Page
2.1 Elements of Power System Model under consideration . . . . .	9
2.2 Information on Test Set . . . . .	14
2.3 Control Performance with Event Detection . . . . .	22
2.4 Decision trees used for response-based one-shot control . . . . .	24
2.5 Simulation results with event detection . . . . .	24
3.1 FO-Lag Method . . . . .	31
3.2 SO-Lag Method . . . . .	31
3.3 TO-Lag Method . . . . .	32
3.4 Strict Lagrange Extrapolation (SE-Lag) . . . . .	32
3.5 Lagrange Coefficients . . . . .	32
3.6 Missing data recovery by Lagrange methods . . . . .	33
4.1 Performance of CC1 and CC2 with zero percentage of missing data . . . . .	37
4.2 Events controlled using ZOH . . . . .	38
4.3 Events stabilized using ZOH . . . . .	38
4.4 Unnecessary controls . . . . .	39
4.5 Nstab using data holds and SE-Lag . . . . .	41
4.6 Performance of FO-Lag with consecutive missing data handling . . . . .	44
4.7 Events destabilized by bootstrap averaging . . . . .	48
4.8 Events destabilized by Mean Value method . . . . .	49

## LIST OF FIGURES

Figure	Page
1.1 SLD of system under consideration . . . . .	3
1.2 Equal area criterion . . . . .	4
2.1 Transmission line retained in the 176 WECC bus model . . . . .	10
2.2 Bus angle for 4 cycle SLG fault near Malin . . . . .	11
2.3 Reconstructed bus angles for 4 cycle SLG fault near Malin . . . . .	11
2.4 Effect of one shot controls on bus angles . . . . .	13
2.5 Loss of synchronism in one of the 29 generators . . . . .	18
2.6 Stable events with all 29 generators in synchronism . . . . .	18
2.7 Using derivative of bus angles index for Control DT . . . . .	19
2.8 Using derivative of bus magnitude for event detection . . . . .	20
2.9 Flow chart for control with event detection . . . . .	21
3.1 Different missing data handling techniques used in the thesis . . . . .	26
3.2 Input(left) and Output(right) from Zero Order Hold . . . . .	28
3.3 Input(blue) and Output(red) from First Order Hold . . . . .	29
3.4 Bootstrap Averaging Method . . . . .	34
4.1 Bus voltage magnitude and derivative of its variance . . . . .	36
4.2 Events controlled using ZOH . . . . .	38
4.3 Events stabilized using ZOH . . . . .	39
4.4 Number of unnecessary controls . . . . .	40
4.5 Comparison of Success rates for ZOH and FOH . . . . .	40
4.6 Number of events stabilized for extrapolation methods . . . . .	41
4.7 Success rates of extrapolation methods . . . . .	42
4.8 Events stabilized by FO-Lag without handling consecutive missing samples	43
4.9 Events stabilized by FO-Lag handling consecutive missing data . . . . .	45

Figure	Page
4.10 Events stabilized with consecutive missing data handling . . . . .	45
4.11 Success rates with consecutive missing data handling . . . . .	46
4.12 Mean control time with consecutive missing data handling . . . . .	47
4.13 Higher unnecessary controls for Bootstrap Averaging . . . . .	48
4.14 Number of events stabilized with CC2 . . . . .	49
4.15 Success rates with CC2 . . . . .	50



## ABBREVIATIONS

CC	Control Combination
DC	Direct Current
DT	Decision Tree
FOH	First Order Hold
FO-Lag	First Order Lagrange
HVDC	High Voltage Direct Current
ISGA	Integral Square Generator Angle
MW	Mega-Watt
PACI	Pacific AC Intertie
PMU	Phasor Measurement Unit
PSAT	Powerflow and Short Circuit Analysis Tool
RVM	Relevance Vector Machine
SBA	Square Bus Angle
SE-Lag	Strict Lagrange Extrapolation
SLD	Single Line Diagram
SO-Lag	Second Order Lagrange
SLG	Single Line to Ground
SVM	Support Vector Machine
TO-Lag	Third Order Lagrange
TSAT	Transient Security Assessment Tool
WAMS	Wide Area Monitoring System
WECC	Western Electricity Coordinating Council
ZOH	Zero Order Hold

## ABSTRACT

Dahal, Niraj. M.S.E.C.E., Purdue University, August 2019. Methods of Handling Missing Data in One Shot Response Based Power System Control. Major Professor: Steven Rovnyak.

The thesis extends the work done in [1] [2] by Rovnyak, et al. where the authors have described about transient event prediction and response based one shot control using decision trees trained and tested in a 176 bus model of WECC power system network. This thesis contains results from rigorous simulations performed to measure robustness of the existing one shot control subjected to missing PMU's data ranging from 0-10%.

We can divide the thesis into two parts in which the first part includes understanding of the work done in [2] using another set of one-shot control combinations labelled as CC2 and the second part includes measuring their robustness while assuming missing PMU's data.

Previous work from [2] involves use of decision trees for event detection based on different indices to classify a contingency as a 'Fault' or 'No fault' and another set of decision trees that decides either to actuate 'Control' or 'No control'. The actuation of control here means application of one-shot control combination to possibly bring the system to a new equilibrium point which would otherwise attain loss of synchronism. The work done in [2] also includes assessing performance of the one shot control without event detection.

The thesis is organized as follows-

Chapter 1 of the thesis highlights the effect of missing PMUs' data in a power system network and the need to address them appropriately. It also provides a general idea of transient stability and response of a transient fault in a power system.

Chapter 2 forms the foundation of the thesis as it describes the work done in [1] [2] in detail. It describes the power system model used, contingencies set, and different indices used for decision trees. It also describes about the one shot control combination (CC1) deduced by Rovnyak, et.al. of which performance is later tested in this thesis assuming different missing data scenarios. In addition to CC1, the chapter also describes another set of control combination (CC2) whose performance is also tested assuming the same missing data scenarios. This chapter also explains about the control methodology used in [2]. Finally the performance metrics of the DTs are explained at the end of the chapter. These are the same performance metrics used in [2] to measure the robustness of the one shot control. Chapter 2 is thus more a literature review of previous work plus inclusion of few simulation results obtained from CC2 using exactly the same model and same control methodology.

Chapter 3 describes different techniques of handling missing data from PMUs most of which have been used in and referred from different previous papers. Finally Chapter 4 presents the results and analysis of the simulation. The thesis is wrapped up explaining future enhancements and room for improvements.

# 1. INTRODUCTION

## 1.1 Problem Statement

In a wide area monitoring system (WAMS) technology, phasor measurement units (PMUs) are employed to acquire electrical signals over large areas in nearly real time. PMU calculates synchronized phasor measurements and communicates the data to a central location. These data can be used in different applications such as monitoring a power system network and state estimation. PMUs can also be used to better detect short circuit faults, small signal instabilities, rotor angle instabilities, etc. Reference [3] describes the concept of phasor measurements and highlights the application of synchronized phasor measurements to improve monitoring, protection and control of a power system.

The advent of real-time PMU data acquisition technology has opened the possibility of solutions to various power system instability problems, however, there is a risk of data quality issues [4] that might affect the performance of PMUs. Data quality issues refer to recording bad data or missing data due to synchronization and timing signal loss, device errors, equipment malfunction, communication infrastructure limits and many more [5]. On an average, 5-10% of missing points in some historical PMU data sets are recorded according to [6].

Missing data has significant effect that usually reduces the quality of the result and therefore should be handled appropriately. The simplest method popularly known as deletion technique can be used to eliminate missing values, however, in many cases a large amount of data elimination causes drastic influence on the results. Some papers discuss about mean value method that takes whole data average as the missing value, maximum frequency method that takes the most frequent data as the missing value, and so on. These methods are general and are unlikely to yield acceptable perfor-

mance for real-time control systems because they do not involve extrapolation of the most recent measurements. Other better methods that use extrapolation techniques are zero order hold that takes very recent available data as the missing value, first order hold that imputes missing value as weighted sum of two recent data and so on. There is a nice explanation and comparison between first order hold and zero order hold methods in [7] [8]. There are still other methods to recover PMU data such as Lagrange Polynomial Method, Bootstrapping, etc. Few of these methods are discussed separately in Chapter 3 of the document.

The thesis evaluates different methods of handling missing PMUs data in a response based one shot control scheme such as explained in [1] [2]. Chapter 2 of the thesis highlights about response based control scheme and describes the power system model and methodologies used in [2]. The thesis focuses simulations for different amounts of missing data ranging from 1-10% and recommends the method that gives relatively better control performance.

## **1.2 Power System Stability and Control**

Before delving into different missing data handling techniques, it is necessary to understand the platform on which these techniques are actually simulated in the research. As described in [1] [2], all the simulations are performed in a one-shot response based power system control against a wide variety of contingencies that will further be discussed in Chapter 2. The contingencies used include single line to ground faults and three phase faults to produce transient disturbances to the system. An overview of power system stability against transient disturbances is described below followed by the study of the system-response in case of a short circuit fault near a bus (equivalently in a transmission line).

### 1.2.1 Power System Transient Stability

The stability concept in power system has been divided into different categories among which the thesis mainly focuses on transient stability. The book [9] classifies power system stability into two parts which are angle stability and voltage stability. The angle stability has been sub-divided into small-signal stability and transient stability. Transient stability is the ability of the power system to maintain synchronism after being subjected to a severe transient disturbance. The disturbances could be either a short circuit on a transmission line(s), loss of load, loss of generator, or loss of transmission lines. The system response to these disturbances include large deviation in generator rotor angles, bus voltages, bus frequency, etc. In our case, the system response includes loss of synchronism for a substantial number of disturbances in the training and test data.

The response of the system to a short circuit fault is described below.

Considering a simple power system network as shown in figure 1.1, let us try to find the response of the system to a three phase fault at location F that is at some distance away from the generator end. The distance of fault occurrence from the generator adds some resistance across the generator terminals and the active power ( $P_e$ ) is not zero. Had the fault occurred right near the sending end, and considering machine winding resistances to be zero, the short circuit current would have flowed through pure reactance resulting zero active power flow [9].

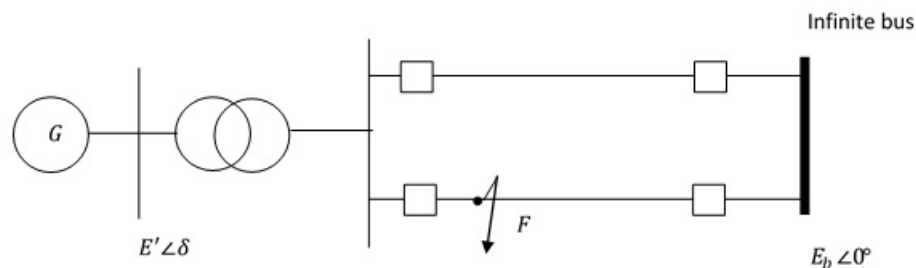


Fig. 1.1. SLD of system under consideration

The electrical power transmitted from sending end to the infinite bus will be,

$$P_e = \frac{(E' E_b)}{X_{eq}} * \sin\delta \quad (1.1)$$

where,  $X_{eq}$  equals the transformer reactance plus the reactance of the parallel transmission lines. The maximum electrical power that a generator can transfer is when  $\delta = 90$  degrees and equals  $E' E_b / X_{eq}$ . As evident from the Equation 1.1, the electrical power ( $P_e$ ) versus the rotor angle ( $\delta$ ) plot is a sinusoidal waveform as also seen from Figure 1.2.

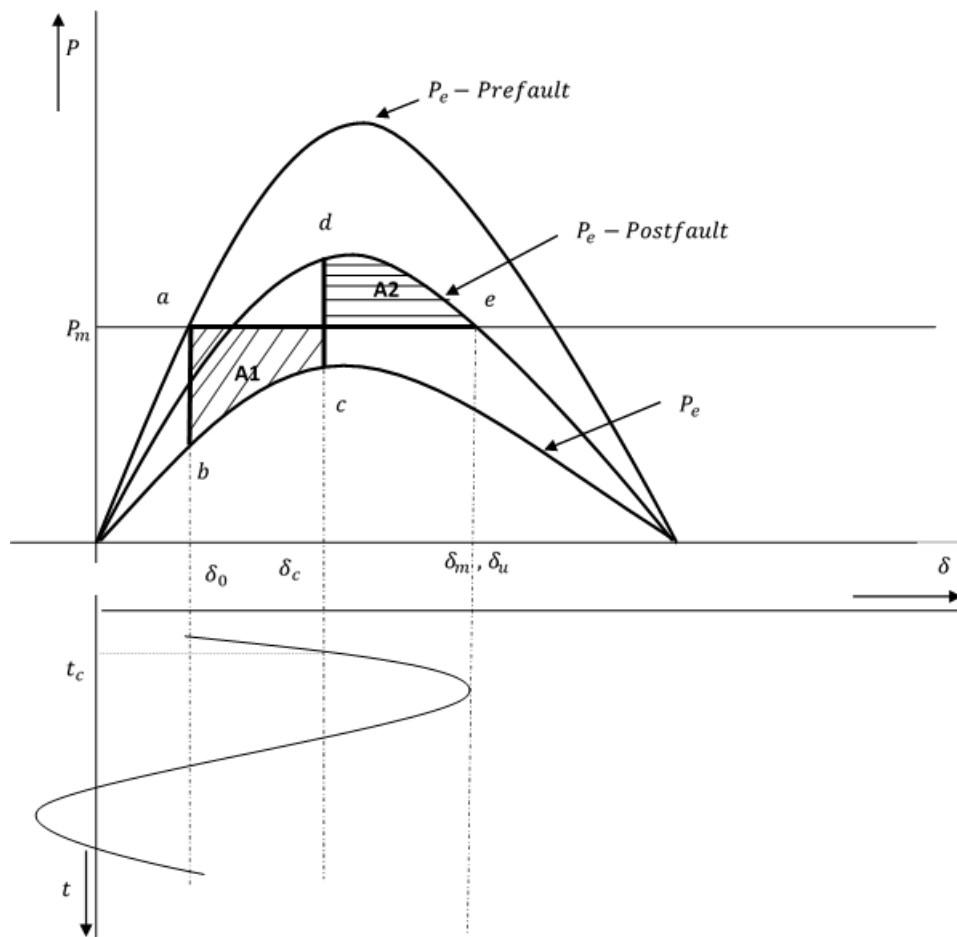


Fig. 1.2. Equal area criterion

Initially before the fault, the system operates at an operating point  $a$  that corresponds to  $\delta = \delta_0$  and  $P_e = P_m$  where  $P_m$  being the mechanical power input to the prime mover. As soon as the three phase fault occurs at  $F$ , the operating point shifts to point  $b$  because the electrical power delivered by the generator decreases substantially. At point  $b$  the mechanical power input is higher than the electrical power output, therefore, the rotor starts accelerating towards point  $c$  along the power angle curve during fault. The point  $c$  corresponds to a critical clearing angle  $\delta_c$  below which the fault should already be cleared to prevent the maximum rotor angle ( $\delta_m$ ) swing beyond the unstable equilibrium point  $\delta_u$ . The time corresponding to  $\delta_c$  is called critical clearing time and is denoted by  $t_c$  in the Figure 1.2. The Figure 1.2 shows the case where the fault is cleared exactly at the critical clearing time.

As soon as the circuit breaker opens and the faulty line gets removed, the power curve takes the post-fault path and reaches point  $d$ . The post-fault path follows lower locus because after the removal of a parallel transmission line, the maximum power transferred decreases due to increase in equivalent reactance.

When the fault gets cleared, the rotor starts decelerating but its phase angle continue to increase towards point  $e$  until the area  $A_2$  becomes equal to the area  $A_1$ . The angle at point  $e$  is the safest limit for maximum swing angle  $\delta_m$  for the rotor. If the generator angle continues swinging beyond this unstable equilibrium point, the system becomes transiently unstable.

If the maximum angle  $\delta_m$  is less than the unstable equilibrium point angle  $\delta_u$ , the mechanical input power is lower than the electrical output at  $\delta = \delta_m$ , so the rotor angle starts decreasing and goes past point  $d$  on the post-fault power angle curve. With some damping the rotor angle oscillates around and approaches the post fault stable equilibrium point.



The kinetic energy gained during rotor acceleration represented by area  $A_1$  is transferred to the system during rotor deceleration represented by area  $A_2$  such that  $A_1 = A_2$ . This is termed as equal area criterion. From Figure 1.2,

$$\begin{aligned} \text{Area } A1 &= \int_{\delta_0}^{\delta_c} (P_m - P_e) d\delta \\ \text{Area } A2 &= \int_{\delta_c}^{\delta_m} (P_m - P_e) d\delta \end{aligned} \tag{1.2}$$

### 1.2.2 Improvement of Transient Stability

There are different methods to improve transient stability of a system following severe short circuit faults. Book [9] describes methods like high-speed fault clearing, regulated shunt compensation, generator tripping, load shedding, dynamic braking, reduction of transmission system reactance, single-pole switching, etc. Discontinuous methods like generator tripping, load shedding, and fast HVDC power changes (DC power modulation) are considered to be promising and have been simulated in the simplified 176 bus model of the western North American synchronous interconnection as in [10]. The paper [11] explains about performance indices used for fast power changes on the DC line to improve stability implementing decision tree control using phasor measurements.

The one shot control combination used in this thesis is based on fast power changes at buses. The TSAT software uses ADD ADMITTANCE command that performs fast power changes at the DC terminal buses as explained in [11]. The one shot control combination used here is further described in Section 2.3 in Chapter 2.

### 1.3 Tools used for the Thesis

The following software tools were used for the thesis.

### **1.3.1 Transient Security Assessment Tool (TSAT)**

TSAT is a simulation tool established by Powertech Labs Inc. designed for assessment of dynamic behavior of power system. It is an easy-to-use graphical interface for processing large number of contingencies. TSAT is best suited for studies involving rotor angle, voltage, or frequency stability [12].

### **1.3.2 PowerFlow and Short Circuit Analysis Tool (PSAT)**

PSAT is mainly used for creation, examination and modification of power-flow models and cases. It supports wide range of models and is a robust power-flow solver [13].

### **1.3.3 MATLAB**

MATLAB is a numerical computing environment used for matrix manipulations, plotting of function and data, implementation of algorithms and for interfacing different programs written in C, C++, FORTRAN, etc. [14]. In the thesis, MATLAB is used to call TSAT to carry out simulations. In addition, overall programming to detect event and actuate control was written in MATLAB.

## 2. RESPONSE BASED EVENT DETECTION FOR ONE SHOT STABILITY CONTROLS

### 2.1 Overview

Early methods to protect a power system from impending instability generally employed out-of-step relaying using apparent resistance ( $R$ ) measured around the electrical centre of a transmission line. The paper [15] published in early 1980's explained a better out-of-step relay that involved augmentation of apparent resistance control with computed rate of change of apparent resistance ( $R_{dot}$ ). For these out-of-step relaying techniques, only local measurements and simple computation were required and were able enough to avoid tripping of any power system element during stable swings while protect the system during out-of-step conditions.

Real time transient stability prediction and control in early days solely relied on event based methods. Event based methods use the status of a particular circuit breaker to decide when to actuate control. The event-based methods were common in practice, however, the challenge to employ these methods in real time was to know when and where a fault had actually occurred in the network. Monitoring the status of circuit breakers would give an idea of faults in a particular region so that the measurements could then be taken to predict loss of synchronism. Since event based control relied on direct detection of equipment outage in order to operate, these methods were somewhat limited. Modern method that researchers have tested for real time wide-area monitoring system (WAMS) control include response based event detection and control methods.

Response based event detection does not rely on direct detection of equipment outage but involves monitoring and analysis of continuous phasor measurements from different synchronized PMUs from different locations to detect a fault. This method,

unlike event based method, detects events not only associated to a particular faulty component but throughout the system. For the past few years, real time transient stability prediction and control has been extensively studied and implemented using machine learning algorithms. Reference [16] describes the work from early 1990's that involves real time transient stability prediction based on decision trees approach tested on New England 39 bus model. References [17] [18] uses  $R - R_{dot}$  scheme measured near the electrical center of Pacific AC Intertie (PACI) as an input to response based decision trees (DTs) to trigger one-shot stabilizing control. The method of using DTs and one-shot controls are extended in a 176 bus model of WECC system using synchronized phasor measurements for wide-area response based control in [1].

The following sections in this chapter mostly take references from [2] to describe about power system model and response based one shot control scheme used in this thesis.

## 2.2 Power System Model

The power system used in this thesis is a 176-bus model from Western Electricity Coordinating Council (WECC) consisting of 29 generators and 71 load buses. PMUs in this study are assumed to be installed at 17 buses. This is the same model used in [1] [2] and consists of following elements.

Table 2.1.  
Elements of Power System Model under consideration

Total Buses	176
Generator Buses	29
Load Buses	71
Fixed Transformers	60
Fixed Shunts	39
Lines	199
Total PMUs	17

Figure 2.1 extracted from [2] shows the transmission line retained in the 176-bus model of the WECC.

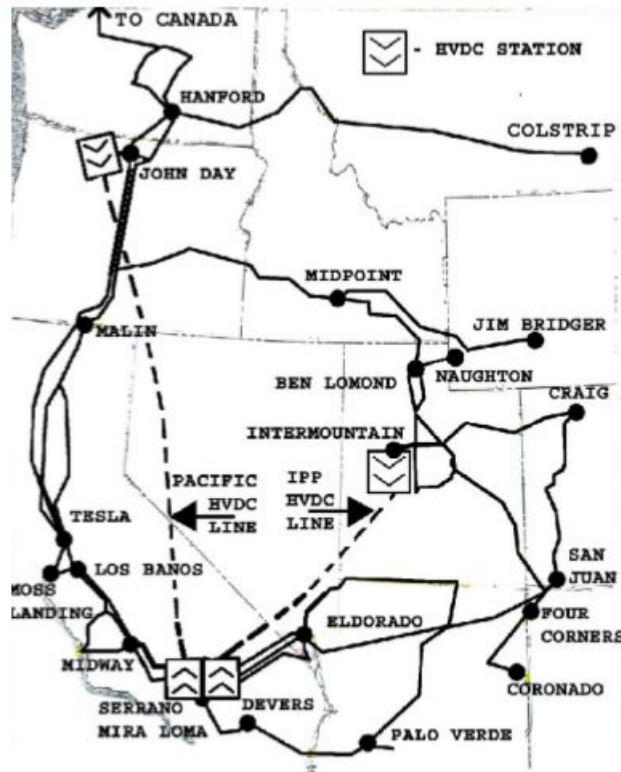


Fig. 2.1. Transmission line retained in the 176 WECC bus model

### 2.3 Phasor Measurements

The phasor measurements used in the thesis are collected from 17 PMU buses at a rate of 30 samples per second like in [2]. The bus voltage angles, unlike generator angles wrap around  $-180$  to  $+180$  degrees, therefore needs to be reconstructed further for continuous trajectories. Reference [19] more clearly explains how these angles are reconstructed. Figure 2.2 involves 12-second simulation to show wrapped bus angles for a 4-cycle SLG fault on a transmission line connected to Malin. It is seen that the bus angles do not go outside  $-180$  to  $180$  degrees. Figure 2.3 shows the unwrapped bus angles for the same event.

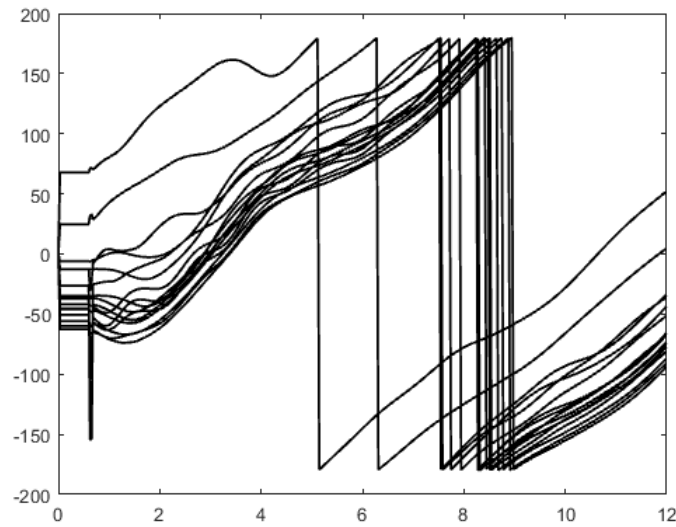


Fig. 2.2. Bus angle for 4 cycle SLG fault near Malin

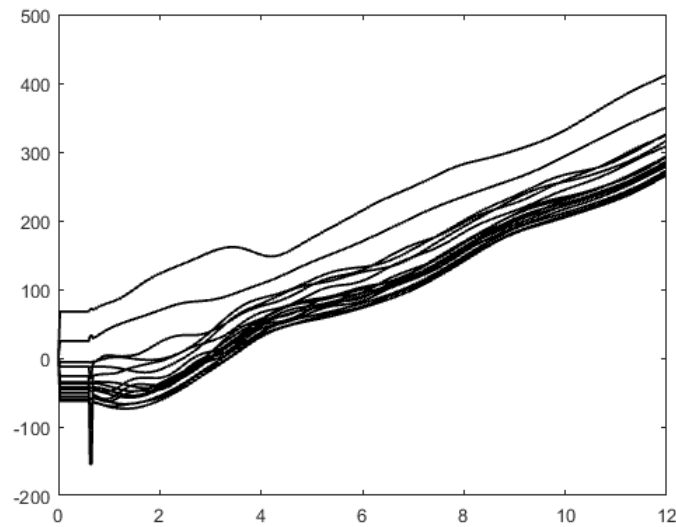


Fig. 2.3. Reconstructed bus angles for 4 cycle SLG fault near Malin

After reconstruction of the bus angles into continuous trajectories, bus angle of Westwing is subtracted from angles of each phasor. Westwing is labelled as the 17<sup>th</sup> bus in this model and is located in the southern part of the network. Reference [2]

explains the reason of selecting Westwing as the reference angle mentioning that the southern area has relatively strong transmission compared to other areas in the model. The choice of Westwing as the reference angle does not effect square bus angle (SBA) index as this index is calculated using center of angle coordinates (COA). More about SBA index can be found on Section 2.6.2.

## 2.4 One Shot Response Based Control Combination

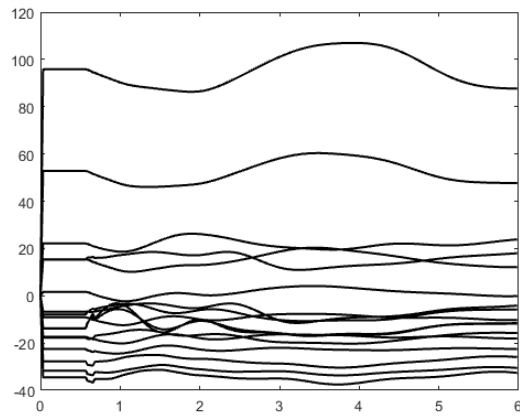
One shot control, as the name implies only actuates once during a simulation after the DT orders it and does not operate until it is reset. It is a single combination of controls that is selected to reduce the phase angle differences for preventing loss of synchronism. The one shot control used in [1] [2] is mainly real power injections at pairs of AC buses. The control combination used is 500 MW fast power increase in buses to Montana and Canada and corresponding 500 MW fast power decrease to two buses in the southern ends of two HVDC lines. We label this control combination as CC1 in this thesis.

In addition to CC1, simulation with another one-shot control combination is also tried by changing power on buses using ADD ADMITTANCE command in TSAT. Adding admittance with positive power increases the power demand whereas with negative power decreases it. Another one shot control combination comprises of 500 MW fast power increase to buses in Montana and Canada and corresponding 500 MW fast power decrease to buses MIDWAY 200, NAVAJO1 500, and MOHAVE 500. We label this control combination as CC2 in the thesis.

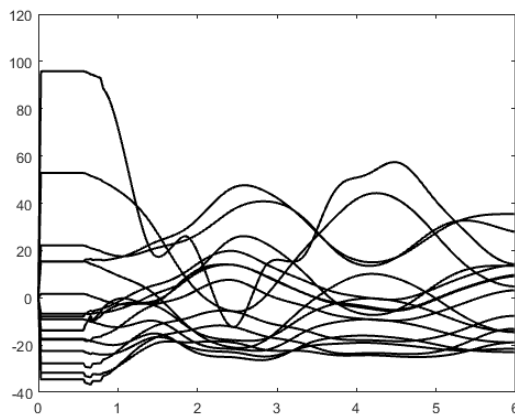
As already stated, these one shot controls are so selected to reduce phase angle differences among generators to prevent loss of synchronism. Lets see a simulation example that shows reduction of phase angle differences among generators after control application.

Figure 2.4(a) shows bus angles for a 5 cycle single line to ground fault near Palo Varde without applying one shot control. Figure 2.4(b) shows how the one shot control combination CC1 helps reduce the bus angle differences for the same fault. The figures are in center of angle (COA) coordinates.

It can be predicted from Figure 2.4(a) that the event is likely to be stable even without control. It is actually an unnecessary control, however the decision to order control occurs at 0.78 seconds.



(a) Bus angles for a 5 cycles SLG faults near Palo Varde



(b) Reduced bus angle differences after control application

Fig. 2.4. Effect of one shot controls on bus angles



## 2.5 Training and Test Data Sets for Decision Trees (DTs)

According to [2], about 30 transmission lines retained from 176-bus model in addition with 10 more lines from the Pacific AC Intertie (PACI) are used to develop training sets, making a final list of 40 lines. The training set contains 160 single outages contingencies (4 outages for each 40 lines), 210 double outages contingencies and 15 more single contingency events on the PACI. Four single outages contingencies for each 40 lines include (i) line removal after a momentary fault (ii) one-cycle three phase fault, (iii) four-cycle three phase fault, (iv) six-cycle one phase fault.

For test set, each of the 40 lines has 12 simulations with different fault duration for three phase short circuit faults and single line to ground faults making a total of 960 simulations. The duration for SLG fault varies from 4 cycles to 15 cycles, while for the three phase faults it is 0-11 cycles. All the fault clearance occurs at 40th cycle i.e. 0.67 seconds implying a different fault starting time for different fault lengths. Table 2.2 below provides information about the test set.

Table 2.2.  
Information on Test Set

<b>Fault Type</b>	<b>Single Phase</b>	<b>Three Phase</b>
Simulation in Test set	480	480
Simulation Length	6 seconds	6 seconds
Fault Duration	4-15 cycles	0-11 cycles

The thesis focuses simulation primarily on 480 single phase to ground fault contingencies. The contingencies are subjected to 6 seconds TSAT simulation with step size of 0.5 cycles. The following lines are extracted from a contingency file where a single line to ground fault occurs at bus MALIN5 at 0.6 seconds. The fault is cleared at 0.67 seconds when the faulted line connecting MALIN5 to MALIN6 is removed.

```

AT TIME 36 CYCLES /
LINE TO GROUND FAULT ;MALIN5 500 0.0002 0.0040 0.0001 0.0020 /
AT TIME 40 CYCLES /
CLEAR LINE TO GROUND FAULT ;MALIN5 500 /
REMOVE LINE ;MALIN5 500;MALIN6 500;1 /

```

Appendix 1 highlights some of the messages generated by a MATLAB program that runs two six-second TSAT simulation, one without control and another with control for the above contingency. The message shows that the control is applied at 0.883 seconds.

## 2.6 Indices Used for Event Detection and One Shot Control

The following indices have been used in [2] [20].

### 2.6.1 Integral Square Generator Angle (ISGA)

The Integral Square Generator Angles (ISGA) index in equation 2.1 can be used to evaluate the severity of transient events [21]. The ISGA can also be used for finding one shot control combinations to stabilize severely unstable transient events.

$$J = \int_0^T \sum_i M_i (\delta_i(t) - \delta_{coi}(t))^2 dt / \sum_i M_i \quad (2.1)$$

where  $M_i$  (constants) are called machine inertias,  $\delta_i(t)$  represent generator angles as a function of time and  $\delta_{coi}$  is the center of inertia (COI).

$$\delta_{coi}(t) = \sum_i M_i \delta_i(t) / \sum_i M_i \quad (2.2)$$

Generators farther from the COI are penalized more by the ISGA. Disconnected generators are not included in these calculations.

### 2.6.2 Square Bus Angle (SBA) Index

The Square Bus Angle (SBA) index and its derivative expressed as SBAdot can be used to detect event or actuate control. The SBA index is calculated in center of angle (COA) coordinates as shown in Equation 2.4.

$$SBA[k] = \frac{\sum_i M_i (\theta_i[k] - \theta_{coa}[k])^2}{\sum_i M_i} \quad (2.3)$$

$$\theta_{coa}[k] = \frac{\sum_i M_i \theta_i[k]}{\sum_i M_i} \quad (2.4)$$

Here, constants  $M_i$  have been chosen to weight the bus angles from different locations. SBAdot[k] is calculated from point to point differences 30 times per second as follows:

$$SBAdot[k] = 30(SBA[k] - SBA[k - 1]) \quad (2.5)$$

An Integrated Square Bus Angle (ISBA) similar to ISGA index can also be used to measure overall stress in terms of phase angle differences [20]. The bus angles used in ISBA calculation are taken from major buses all around the network. While the ISGA index used machine inertias for weights, it is possible to assign either equal weights for ISBA index or weights based on voltage levels of the buses. The ISBA can be obtained by integrating the SBA over a sliding window or by applying some other low pass filter to the SBA as explained in [20].

### 2.6.3 Bus Magnitude Index

Bus Magnitude Index is calculated from the bus voltage magnitudes. The average ( $BMavg$ ), variance ( $BMvar$ ) and their derivatives ( $BMavgdot$  or  $BMvardot$ ) can be used for actuating one shot control decision.

For 17 buses that are equally weighted,

$$BMavg[k] = \sum_i V_i[k]/17 \quad (2.6)$$

$$BMvar[k] = \sum_i (V_i[k] - BMavg[k])^2/17 \quad (2.7)$$

Their derivatives are calculated from point to point differences 30 times per second.

$$BMavgdot[k] = 30 * (BMavg[k] - BMavg[k - 1]) \quad (2.8)$$

$$BMvardot[k] = 30 * (BMvar[k] - BMvar[k - 1]) \quad (2.9)$$

## 2.7 Classification of Event as ‘Stable’ or ‘Unstable’

Generally the transient stability is associated with deviation of generator rotor angles. A transiently stable event has to have rotor angle oscillating around an equilibrium point. In our case, an event becomes unstable if one or more of the generator angles start to diverge as they move forward in the simulation. If the rotor angle keeps on diverging, the machine will lose synchronism with other generators making the system unstable. Figure 2.5 shows loss of synchronism of one generator in the 29-machine model of WECC. The synchronism is maintained after applying one-shot control as shown in Figure 2.6. It is therefore logical to classify any contingency as ‘stable’ or ‘unstable’ based on the analysis of its generator rotor angles.

In our case, the classification of a contingency as ‘unstable’ is done when the maximum difference between the angles of any pair of generators exceeds 300 degrees at any time during the simulation. This is the same criterion used for determining instability in the 176 bus network in [19] [20].

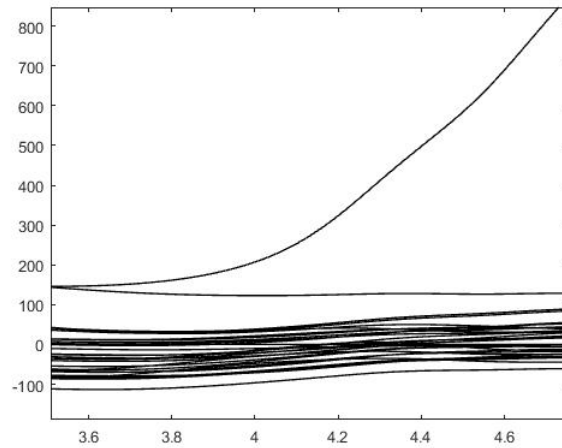


Fig. 2.5. Loss of synchronism in one of the 29 generators

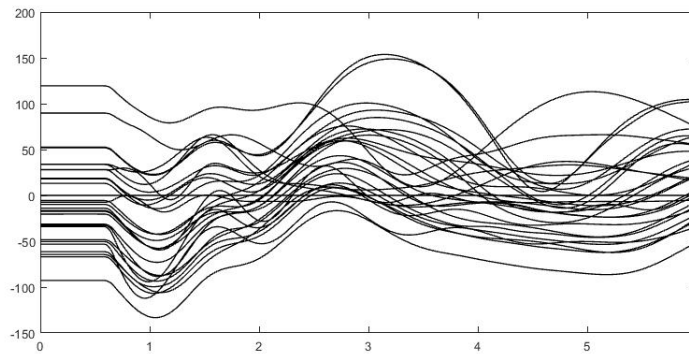


Fig. 2.6. Stable events with all 29 generators in synchronism

## 2.8 DTs for Response based Control

The DTs for response based control are trained using input-output vectors from each time-sample of PMU measurements recorded at 30 times per seconds. The input vector includes voltage magnitudes and angles from 17 PMUs and indices calculated from them. The target includes correct value of the output for every input vector. The DTs in [2] are trained using rpart function in Rattle.

### 2.8.1 Training DT for One Shot Control

The target for DT is set to ‘No control’ for every time-sample if a simulation is stable. If the simulation is not stable then the target is set to ‘Control’ for every time sample. The trained DT actuates control for any input vector the first time it outputs ‘Control’.

While training the DTs, there is a parameter called ‘misclassification cost’ that might be set higher like 10 or 100 times for misclassifying ‘No control’ compared to the cost of misclassifying ‘Control’ samples. The number of simulations having control action can be increased or decreased using different values of the relative misclassification cost. Similarly, there is another parameter called ‘complexity cost’ during the training that can sometimes be used to control the size of the DT generated. The DT trained using derivative of squared bus angle ( $SBA_{dot}$ ) as an input with misclassification cost set to 10:1 and complexity cost set to 0.1 gives ‘Control’ for any sample that has  $SBA_{dot} > 24$  (rounded figure). Using this control parameter, the number of events controlled is recorded to be 345 out of 480 single line to ground test faults with 29 events stabilized.

Figure 2.7 shows how an index like  $SBA_{dot}$  can be used for actuating control.

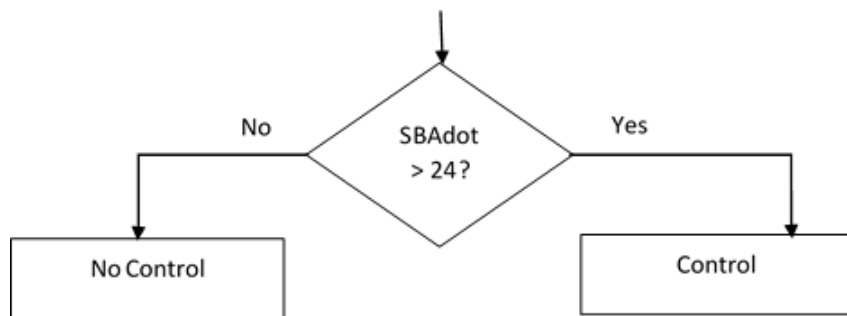


Fig. 2.7. Using derivative of bus angles index for Control DT

Previous work in [2] includes training the DTs for one shot control without event detection as well as with event detection. The addition of event detection to the one-shot control scheme helps to enable control only for a few samples after an event is detected. Thus, the number of unnecessary control actuations can be reduced with augmentation of event detection DT. The following section discuss how an event detection DT is trained and added to control DT for response based control.

### 2.8.2 Training DT for Event Detection

As stated in Section 2.5, all the simulations in the training set have a short circuit to ground fault that is cleared at 0.67 seconds. The target for DT is set to ‘Postfault’ for every time sample between 0.67 seconds to 0.72 seconds. The target is ‘Not Postfault’ for all other samples. There is generally one sample per simulation with the target set to ‘Postfault’.

The relative misclassification cost is set higher like 10 or 100 times for misclassifying ‘Postfault’ samples than the cost of misclassifying ‘Not Postfault’ samples.

The DT trained using derivative of bus voltage magnitude variance ( $bmvardot$ ) as an input with missclassification cost set to 1:100 and complexity cost set to 0.1 gives ‘Postfault’ for any sample that has  $bmvardot < -0.03$ . Figure 2.8 shows the use of derivative of bus voltage magnitude variance for event detection DT.

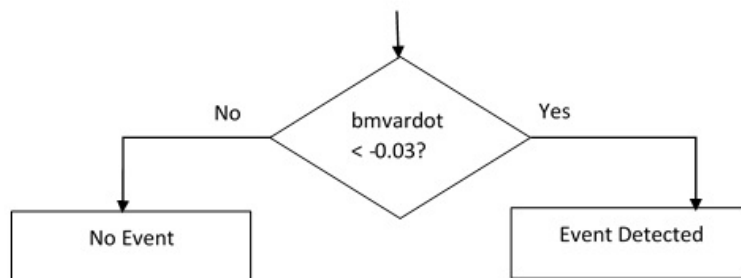


Fig. 2.8. Using derivative of bus magnitude for event detection

The DT that orders control with event detection is trained in similar manner as described in Section 2.8.1 except that the relative missclassification cost is set to 1:1 instead of 10:1. The control DT trained using angular velocity ( $\theta_{dot9}$ ) as an input with missclassification cost set to 1:1 gives ‘Control’ for any sample that has  $\theta_{dot9} > 65$ . Previous work from [2] states that a peak in success rate is obtained when  $\theta_{dot9} > 50$ . Therefore, for the control criteria of  $\theta_{dot9} > 50$  augmented with event detection criteria of  $bmvardot < -0.03$ , the number of events controlled is recorded to be 96 with 22 events stabilized and only 6 unnecessary controls.

Figure 2.9 shows the flow chart for response based control using event detection.

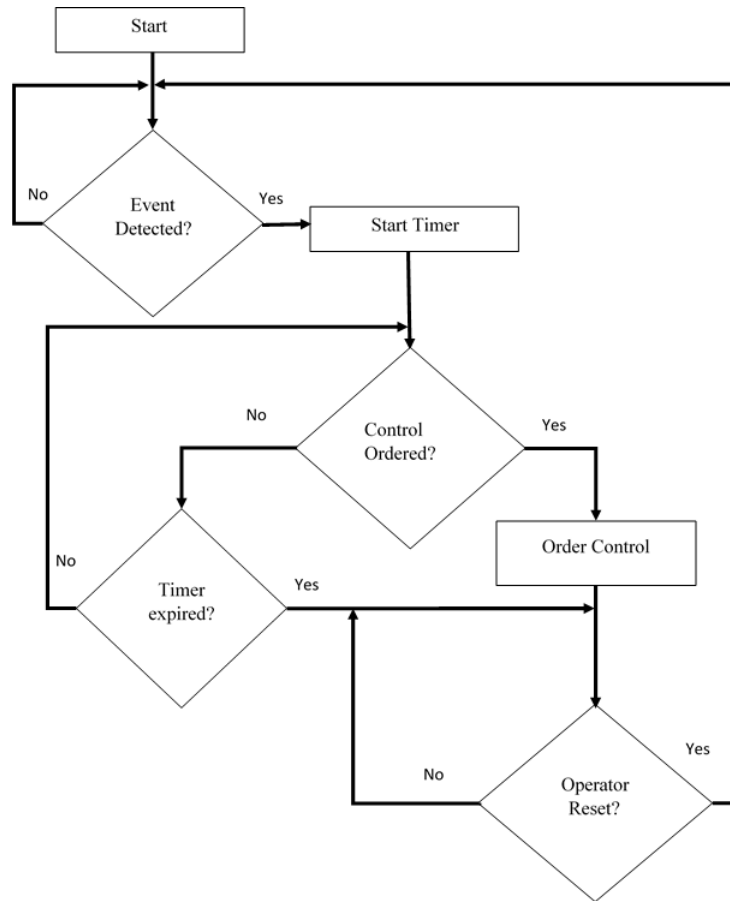


Fig. 2.9. Flow chart for control with event detection



As seen from the flowchart, for the case of one-shot control with event detection, the control DT is checked only for few samples until the end of timer to help reduce the number of unnecessary control actuations. The timer set in our case is maximum of 0.167 seconds or 5/30 seconds. Further, if the DT orders one shot control, the control is applied 100 milliseconds later in the simulation.

## 2.9 Control Methodology for Test sets

For every simulation in the test set that has control actuation, we run another six-second simulation with control actuation 100 milliseconds after the DT orders control. The generator rotor angles are analyzed before and after applying the one-shot control to see if the loss of synchronism, if any, is avoided. We first classify every contingencies as ‘stable’ or ‘unstable’ based on the analysis of generator rotor angles. For either a stable or an unstable event, we then check event detection DTs to every sample at 30 Hz until 6 seconds of simulation. If an event is detected, a timer of maximum 5/30 seconds is started and the control DT is checked. Even if an event is detected at any timestamp, it might not actuate control unless the control DT is satisfied within the specified timer. So, there could be any of the following possibilities to our simulation.

Table 2.3.  
Control Performance with Event Detection

Event Class	Event Detected?	Cntrl Actuated?	Result	Cntrl Remarks
Stable(S)	Yes	Yes	Stable	Unnecessary
Stable(S)	Yes	Yes	Unstable	Destabilized
Unstable(U)	Yes	Yes	Stable	Stabilized
Unstable(U)	Yes	Yes/No	Unstable	Kept Unstable
Stable(S)	Yes	Yes/No	Stable	Kept Stable
S/U	No	No	S/U	Not detected/kept as it is

The number of events kept stable also includes stable events where event is not detected. Similarly, the number of events kept unstable also includes unstable events where event is not detected. For the test set of data used in the thesis, there are no

events undetectable by the DTs. Similarly the test set does not consist of an event that is destabilized by the existing one shot control. The performance of the control is considered to be optimal when it can increase the number of events stabilized while decreasing number of unnecessary controls. Each of the control performance metrics is described below in Section 2.10.

## 2.10 Performance Metrics for One Shot Control

Performance of the one shot control is measured using the following parameters. The specifications for some of these parameters can be found in the previous section.

- i Number of events stabilized ( $N_{stab}$ )
- ii Number of events destabilized ( $N_{destab}$ )
- iii Number of events keep stable ( $N_{keepStab}$ )
- iv Number of events keep unstable ( $N_{keepUnst}$ )
- v Number of events controlled ( $N_{cntrl}$ )
- vi Number of control unnecessary ( $N_{cntrlUnne}$ )
- vii Number of events not detected ( $N_{notDet}$ )
- viii Mean control time ( $T_{avg}$ )
- ix Success rate ( $R_{success}$ )

The number of events controlled represents every simulations during which control is ordered and actuated. The number of events stabilized includes events that are stable with one shot control and unstable without the control. The number of events destabilized includes stable events that are unstable with the particular control scheme. This parameter is zero for all the test simulations in the thesis.

The number of control unnecessary includes events that are already stable without the control, however still undergoes control actuation and remains stable. Success rate is calculated as the number of events stabilized upon the number of events controlled.

The mean control time is calculated as the sum of control time divided by the number of events controlled.

## 2.11 Simulation Results with Event Detection

Reference [2] shows simulation results obtained from a number of different event detection and control criteria. The results obtained from the final DTs in [2] are only presented in this thesis. Their thresholds are shown in Table 2.4.

Table 2.4.  
Decision trees used for response-based one-shot control

Event detection DT	$bmvardot < -0.03$
Control DT	$\theta_{dot9} > 50$

Table 2.5 shows simulation results for controls CC1 and CC2 with event detection for 480 SLG faults.

Table 2.5.  
Simulation results with event detection

<b>Metrics</b>	<b>CC1</b>	<b>CC2</b>
Nstab	22	26
Ndestab	0	0
NkeepStab	276	276
NkeepUnst	182	178
Ncntrl	96	96
NcntrlUnne	6	6
NnotDet	0	0
Tavg	0.760766	0.760766
Rsuccess	0.229167	0.270833

### 3. PROPOSED METHODS FOR MISSING DATA HANDLING

#### 3.1 Overview

As described in Section 1.1, there can be different methods to handle missing data from PMUs. The thesis mainly focuses on methods that use extrapolation of previous data to recover missing ones. A simple extrapolation technique is zero order hold (ZOH) which holds a recent available data from a PMU channel until next sample appears. Another data hold method is first order hold (FOH) that uses two recent available data from the same channel to approximate the missing value. We are assuming that any value obtained from a PMU at a particular timestamp is correlated to previous samples from the same channel. Therefore, in general, we employ Lagrange polynomial method (discussed later) and particularly use up to three previous values for extrapolation. Lagrange polynomial method is also used for interpolation where future known values can be interpolated back to approximate a missing value at a certain timestamp. However, for a real time approximation of a missing value, the interpolation technique might be tricky as one has to wait until next samples arrive. In this thesis, we thus try to use only one succeeding sample value for interpolation in combination with extrapolation from previous samples.

The authors in [22] hypothesize that a single frequency event can be detected from many synchronized PMU stations and hence there is a certain correlation among many frequency data channels from these PMU stations. This allows to recover missing value of one PMU station by taking account of samples from different other PMU stations in the system. The thesis describes another method called bootstrap averaging and tries to approximate a missing value using data from other randomly selected PMU stations.

Figure 3.1 shows different missing data handling techniques used in the thesis.

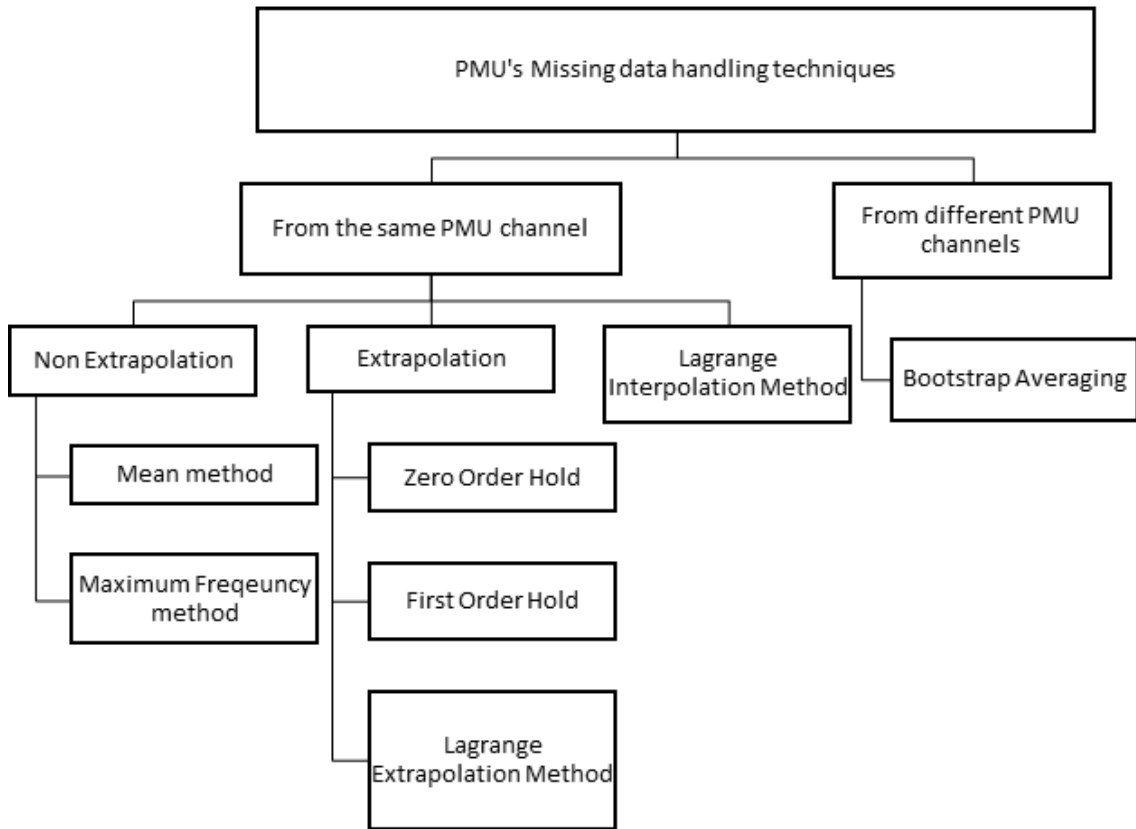


Fig. 3.1. Different missing data handling techniques used in the thesis

### 3.2 Missing Data Handling Techniques

Methods used in the thesis for handling missing data are discussed below.

### 3.2.1 Data Hold

Holding a data generates a continuous time signal  $g(t)$  from a discrete-time sequence  $x(kT)$  [8]. Let us consider a signal  $g(t)$  during a time interval  $t$  such that  $kT \leq t \leq (k+1)T$ , then  $g(t)$  can be approximated by a polynomial as-

$$\begin{aligned} g(kT + \tau) &= a_0 + a_1\tau + \dots + a_{n-1}\tau^{n-1} + a_n\tau^n \\ g(kT + \tau) &= x(kT) + a_1\tau + \dots + a_{n-1}\tau^{n-1} + a_n\tau^n \end{aligned} \quad (3.1)$$

where  $0 \leq \tau \leq T$ .

It can be seen that at  $\tau = 0$ ,  $g(kT) = x(kT)$ . An  $n^{\text{th}}$  order hold circuit uses  $(n+1)$  discrete data to generate  $g(kT + \tau)$ .

#### Zero Order Hold (ZOH)

Zero Order Hold (ZOH) performs signal reconstruction by holding a recent value until next sample arrives [8]. In the Equation 3.1, if  $n=0$ , a zero-order hold is obtained such that

$$g(kT + \tau) = x(kT) \quad (3.2)$$

where  $0 \leq \tau \leq T$  and  $k = 0, 1, 2, \dots$

**Example 3.1:** Let us consider an analog input voltage at 60 Hz sampled at a frequency of 720 Hz. The following matrix A is extracted from a file where the first column corresponds to sampled time and second column corresponds to measured voltages for a random positive half cycle of an ac signal.

A =  
 0.1389 2.4400  
 0.1403 5.7200  
 0.1416 6.9800  
 0.1430 6.5700  
 0.1444 4.3800  
 0.1458 0.8000

Figure 3.2 shows how a zero order hold circuit retains recent data until next sample arrives.

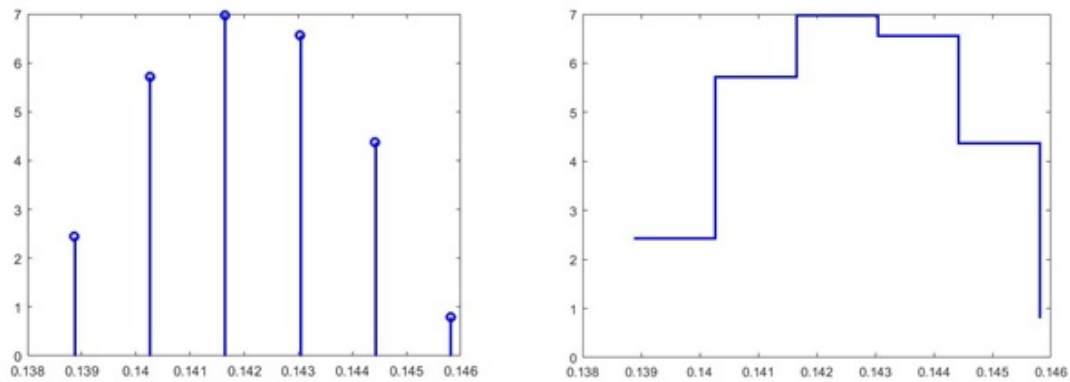


Fig. 3.2. Input(left) and Output(right) from Zero Order Hold

### First Order Hold (FOH)

First Order Hold (FOH) performs signal reconstruction by using  $(n+1) = 2$  recent discrete data [8]. Therefore, in Equation 3.1 for  $n=1$ ,

$$g(kT + \tau) = x(kT) + a_1\tau \quad (3.3)$$

where,  $0 \leq \tau \leq T$  and  $k = 0, 1, 2, \dots$

We have,

$$g((k-1)T) = x((k-1)T) \quad (3.4)$$

Therefore,

$$\begin{aligned} x(kT) - a_1T &= x((k-1)T) \\ a_1 &= \frac{x(kT) - x((k-1)T)}{T} \end{aligned} \quad (3.5)$$

This implies,

$$g(kT + \tau) = x(kT) + \frac{x(kT) - x((k-1)T)}{T} \tau \quad (3.6)$$

For the same example of a sampled ac signal dealt above, the first order hold would generate the following graph.

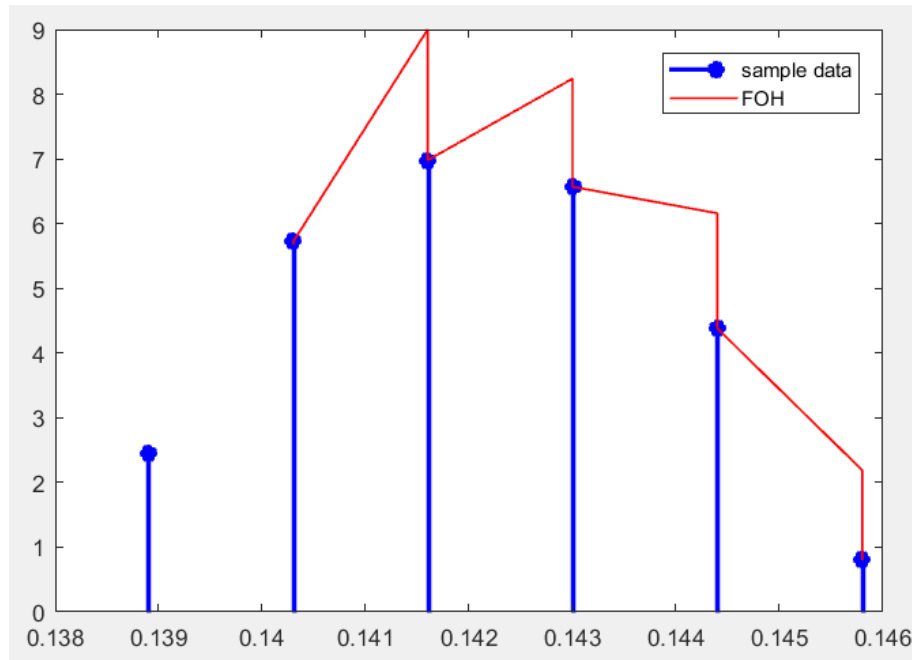


Fig. 3.3. Input(blue) and Output(red) from First Order Hold



### 3.2.2 Lagrange Interpolating Polynomial Method

Lagrange interpolating polynomial method gives the lowest degree polynomial passing through given set of points  $(x_j, y_j)$  where each  $x_j$  is different than the other such that  $1 \leq j \leq (n + 1)$ . For a polynomial approximation  $P$  of a function  $f(x) = y$  as follows-

$$P(n, x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x^1 + p_0 \quad (3.7)$$

there will be  $(n + 1)$  number of linear equations as follows:

$$\begin{aligned} p_n x_1^n + p_{n-1} x_1^{n-1} + \dots + p_1 x_1^1 + p_0 &= y_1 \\ p_n x_2^n + p_{n-1} x_2^{n-1} + \dots + p_1 x_2^1 + p_0 &= y_2 \\ &\dots \\ p_n x_{n+1}^n + p_{n-1} x_{n+1}^{n-1} + \dots + p_1 x_{n+1}^1 + p_0 &= y_{n+1} \end{aligned} \quad (3.8)$$

where  $p_0, p_1, \dots, p_{n+1}$  are the unknown coefficients of the polynomial  $P(n, x)$  and need to be determined. The square matrix formed by the above system of equations will be non-singular and hence will have a unique solution that can be obtained by solving the set of equations in 3.8. The Equation 3.7 is called interpolating polynomial.

Lagrange interpolating polynomial is described in [5] [22] [23] and can be represented as follows:

$$L(x) = \sum_{j=1}^n y_j l_j(x) = \sum_{j=1}^n y_j \prod_{k=1, k \neq j}^n \frac{x - x_k}{x_j - x_k} \quad (3.9)$$

where  $x_j \neq x_k$  and  $l_j(x)$  is the coefficient in the Lagrange polynomial.

Considering  $n=3$ , the Lagrange polynomial can be expanded as follows:

$$L(x) = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} \quad (3.10)$$

Here we choose  $n = 3$  to preserve accuracy while also considering the computation cost with increase in  $n$ .

Reference [23] classifies Lagrange polynomial method into different orders.

### First Order Lagrange Polynomial (FO-Lag)

FO-Lag involves polynomial approximation considering two sets of data samples. The missing sample can then be approximated from the polynomial obtained from points in Table 3.1.

Table 3.1.  
FO-Lag Method

Preceding data	Missing data	Succeeding data
$x_{k-1}$	$x_k$	$x_{k+1}$
$y_{k-1}$	$y_k$	$y_{k+1}$

In our case, the  $x_k$ 's are time domain samples and the  $y_k$ 's are phasor measurements corresponding to  $k^{th}$  sample.

### Second Order Lagrange Polynomial (SO-Lag)

SO-Lag involves polynomial approximation considering three sets of data samples.

Table 3.2.  
SO-Lag Method

Preceding data 2	Preceding data1	Missing data	Succeeding data
$x_{k-2}$	$x_{k-1}$	$x_k$	$x_{k+1}$
$y_{k-2}$	$y_{k-1}$	$y_k$	$y_{k+1}$

### Third Order Lagrange Polynomial (TO-Lag)

TO-Lag involves polynomial approximation considering four set of data samples.

Table 3.3.  
TO-Lag Method

Preceding data 3	Preceding data 2	Preceding data1	Missing data	Succeeding data
$x_{k-3}$	$x_{k-2}$	$x_{k-1}$	$x_k$	$x_{k+1}$
$y_{k-3}$	$y_{k-2}$	$y_{k-1}$	$y_k$	$y_{k+1}$

All the above Lagrange methods use one succeeding data for interpolating the missing value. If we were to use an strict extrapolation technique, we had to omit the succeeding data sample which would give us the Lagrange polynomial obtained from data points in Table 3.4. This is one of the methods used in [5] for recovering missing value.

Table 3.4.  
Strict Lagrange Extrapolation (SE-Lag)

Preceding data 3	Preceding data 2	Preceding data1	Missing data
$x_{k-3}$	$x_{k-2}$	$x_{k-1}$	$x_k$
$y_{k-3}$	$y_{k-2}$	$y_{k-1}$	$y_k$

We now determine Lagrange coefficients  $l_j(x)$  for each of first order, second order, third order and strict extrapolation polynomial to be used in the simulation. Since the PMUs used in [2] perform voltage measurements at 30 samples per second, the time domain difference ( $x_j - x_k$ ) for two consecutive samples is (1/30). Table 3.5 below provides the Lagrange coefficients calculated for the simulation.

Table 3.5.  
Lagrange Coefficients

Methods	$l_0$	$l_1$	$l_2$	$l_3$
FO-Lag	1/2	1/2	—	—
SO-Lag	-1/3	1	1/3	—
TO-Lag	1/4	-1	3/2	1/4
SE-Lag	1	-3	3	—

For the same figures in **Example 3.1** dealt earlier, the missing value  $y_4$  can be recovered by Lagrange polynomial method as-

Table 3.6.  
Missing data recovery by Lagrange methods

Methods	FO-Lag	SO-Lag	TO-Lag	SE-Lag
Value of $y_4$	5.68	6.533	6.455	6.22

### 3.2.3 Bootstrap Averaging

Bootstrap Method is a statistical re-sampling method that allows random sampling with replacement. This method helps to estimate quantities such as mean, standard deviation or variance using random sub samples of the available data set. Bootstrap Averaging (Bagged Averaging) method is explained and used in [22], where the authors hypothesize that a single frequency event can be detected from many PMU stations and hence there is a certain correlation among many frequency data channels from PMU stations. In our case, the missing value is estimated as the average of random  $m \times n$  subset matrix selected from the main data set, where,  $m$  represents the number of samples (size of each bootstrap) and  $n$  represents the number of buses included in the bootstrap. For example, let us consider  $60^{th}$  phasor sample from  $5^{th}$  PMU is missed. The missing value can be estimated as the average of randomly selected  $m \times n$  matrix, where, for example  $m = 5$  samples from  $n = 5$  PMUs like shown in the Figure 3.4. To make the selection of PMUs random with replacement, the program uses  $randi(imax)$  function where it can choose any five PMUs out of  $imax = 17$  total PMUs.

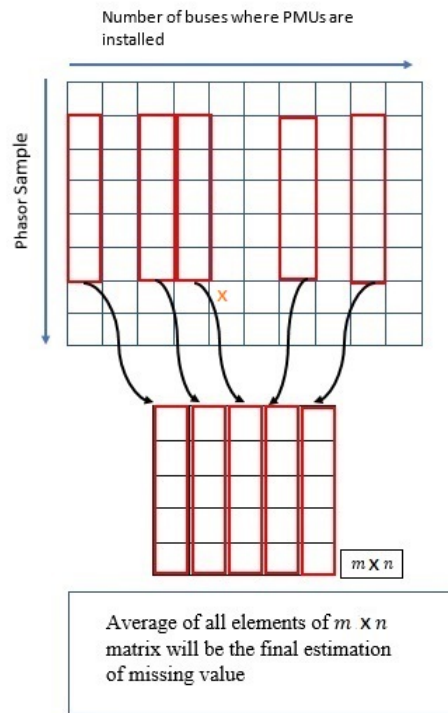


Fig. 3.4. Bootstrap Averaging Method

### 3.3 Generation of Pseudo-random Numbers

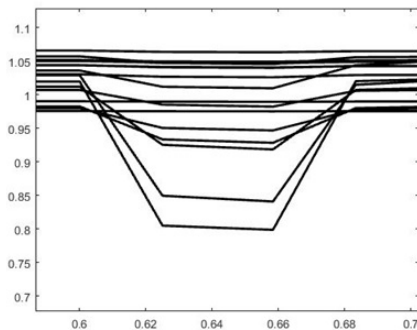
MATLAB creates a sequence of pseudorandom numbers using many inbuilt functions such as *rand*, *randi*, and *randn*. A pseudorandom sequence is the sequence of numbers that appears to be random but are not in actual. These sequence may be uniformly distributed, normally distributed or follow any other distribution. The *rand()* function in MATLAB returns a single uniformly distributed random number strictly in the open interval (0, 1). Therefore, *rand()* function serves as a useful tool to simulate percentage of missing data in our case. For example, if '*rand()*<0.03' is the criteria for which data points are missing, then the percent missing in a large sample would be around 3%. The thesis focuses simulation of missing data handling up to 10% only.

One thing to understand while using `rand()` function is that it returns the same sequence of numbers every time MATLAB is restarted. If we try to simulate our program using `rand()`, the results for same contingency may differ several times because of different values of `rand()` being generated by the system. MATLAB features another command ‘`rng default`’ that puts the pseudorandom generator back to its default settings.

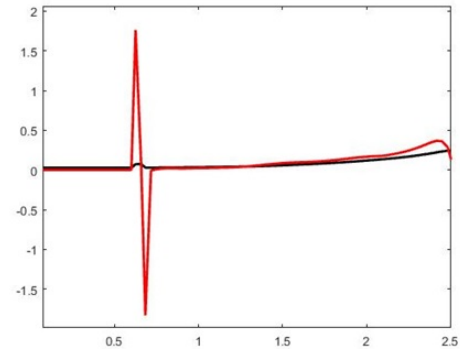
## 4. RESULTS AND ANALYSIS

### 4.1 Overview

Whenever a severe short circuit fault occurs in a transmission system, some of the bus voltage magnitudes reduce below 1 per unit (pu) as shown in Figure 4.1(a). This gives rise to large variance of bus voltage magnitudes. As soon as the faulty line is removed, the voltages tend to return to 1 pu and hence the variance decreases. The derivative of bus voltage magnitude variance (we refer it as *bmvardot*) thus has a large positive spike during the fault and large negative spike or dip right after the fault is cleared as shown in Figure 4.1(b).



(a) Reduction of bus voltage magnitudes (in per unit) during SLG fault



(b) Bmvardot spike and dip during and after a fault

Fig. 4.1. Bus voltage magnitude and derivative of its variance

Bus voltages are comparatively easier to measure and show prominent effects during short circuit faults. This is the reason why the authors in [2] have chosen bus voltage variance as an index for event detection.

The missing data handling techniques described in Chapter 3 are applied to both control combinations CC1 and CC2. As a baseline, simulations are performed for CC1 and CC2 assuming zero percentage of missing PMUs data. The simulation results tested for 480 single line to ground fault contingencies assuming zero percentage of missing data are tabulated below in 4.1. The number of events controlled only depends on the control DT so it is the same for both CC1 and CC2.

Table 4.1.  
Performance of CC1 and CC2 with zero percentage of missing data

<b>Metrics</b>	<b>CC1</b>	<b>CC2</b>
Nstab	22	26
Ndestab	0	0
NkeepStab	276	276
NkeepUnst	182	178
Ncntrl	96	96
NcntrlUnne	6	6
NnotDet	0	0
Tavg	0.760766	0.760766
Rsuccess	0.229167	0.270833

## 4.2 Performance using Zero Order Hold (ZOH)

With increase in percentage of missing data, the number of events controlled was expected to decrease resulting a lower overall performance. However, simulation results from Figure 4.2 shows a rise in number of events controlled after 2% of missing data. This is attributed to the fact that wherever a missing data is encountered from any channel, ZOH reserves the most recent value from that channel in its place. At the point of data recovery, the difference between recent consecutive values is zero. Further the very next difference after the recovery is almost twice as large as it should be. This large derivative value, which occurs frequently as missing percentage increases, causes the decision trees to actuate control for more number of events.



Table 4.2.  
Events controlled using ZOH

% missing	0	1	2	3	4	5	6	7	8	9	10
Nctrl	96	93	85	87	90	95	96	93	96	96	99

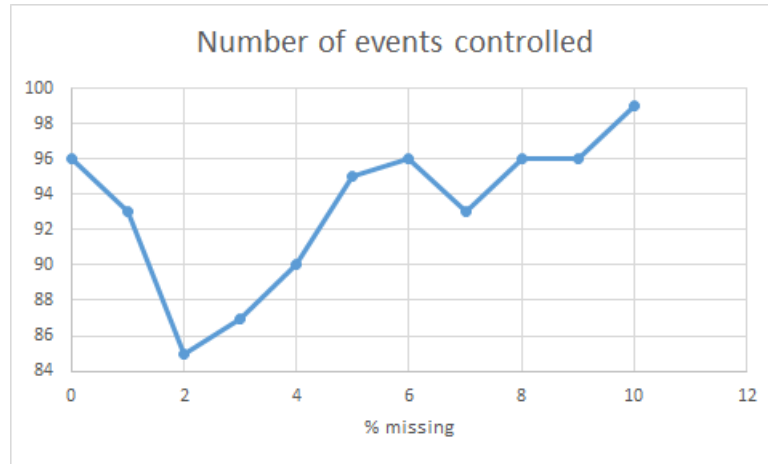


Fig. 4.2. Events controlled using ZOH

Figure 4.3 shows the number of events stabilized up to 10% of missing data using ZOH method for both control combinations CC1 and CC2.

Table 4.3.  
Events stabilized using ZOH

% missing	0	1	2	3	4	5	6	7	8	9	10
CC1	22	19	15	14	10	8	7	7	8	6	4
CC2	26	23	18	16	12	10	10	10	10	8	6

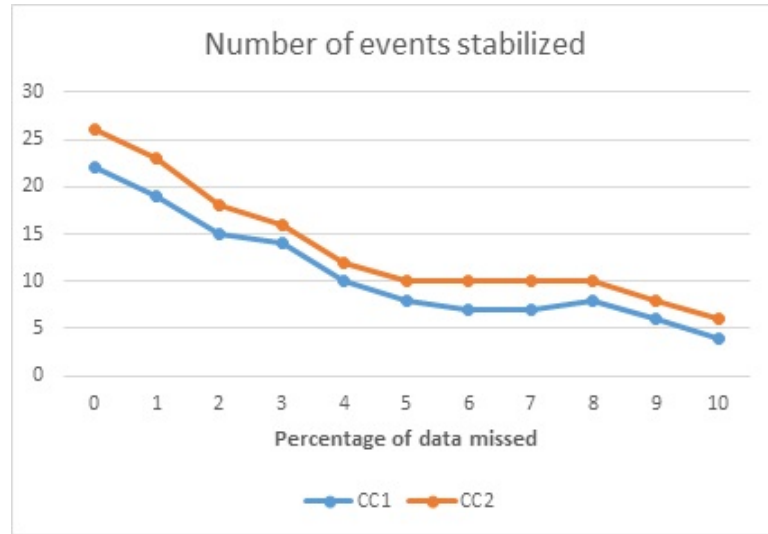


Fig. 4.3. Events stabilized using ZOH

### 4.3 Performance using First Order Hold (FOH)

Another important criteria to determine performance of the existing one shot control is the number of unnecessary controls. An unnecessary control happens when the control is applied to an event that would have been stable without the control. For the same percentage of data missed, first order hold reduces the number of unnecessary controls and proves to be better than its counterpart zero order hold as shown in the Figure 4.4.

Table 4.4.  
Unnecessary controls

% missing	0	1	2	3	4	5	6	7	8	9	10
ZOH	6	12	17	23	29	39	43	43	45	47	50
FOH	6	12	13	16	22	30	30	29	36	43	54

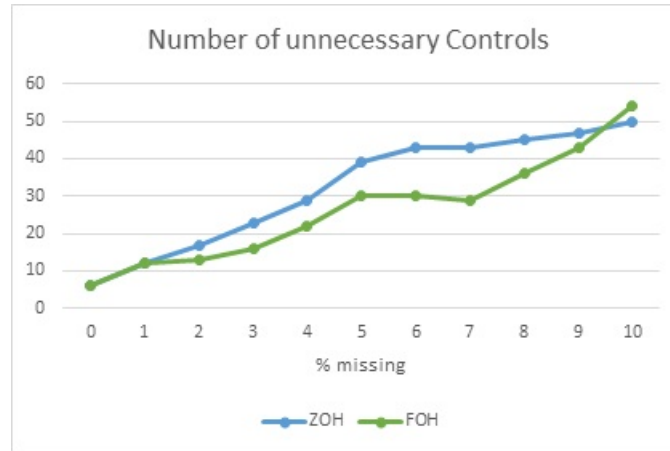
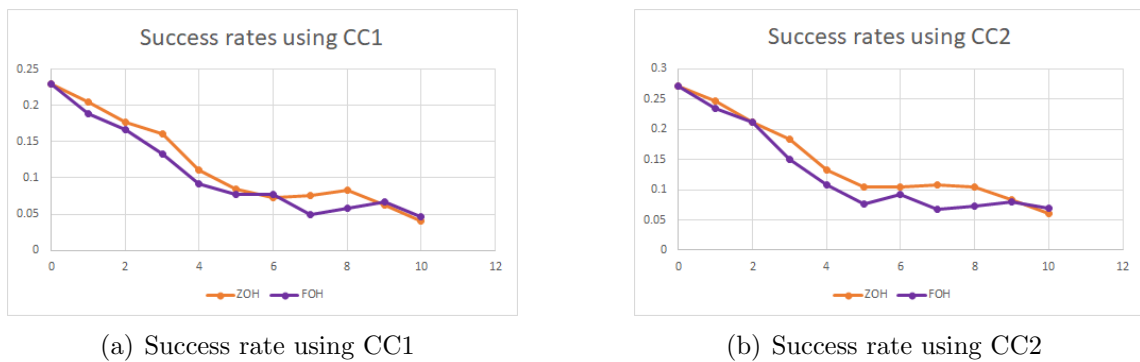


Fig. 4.4. Number of unnecessary controls

A FOH, as described in Section 3.2.1, uses last two available data from the same channel to recover a missing value. The recovered value does not exactly match previous value as in ZOH. This reduces the derivative of bus voltage magnitude variances that ultimately reduces the number of events controlled.

Figure 4.5 shows the comparison of success rate using ZOH and FOH. Higher success rate means the number of events stabilized is larger for same number of events controlled. The results show that irrespective of the one-shot control combinations used, ZOH has higher success rates and hence is better than FOH method.



(a) Success rate using CC1

(b) Success rate using CC2

Fig. 4.5. Comparison of Success rates for ZOH and FOH

#### 4.4 Performance using Lagrange Polynomial Method

Lagrange polynomial method has been divided into two parts viz. extrapolating polynomial method and interpolating polynomial method. The former method uses  $n$  number of preceding data (in our case  $n = 3$  at most) for recovering missing values while the latter also uses one succeeding sample from the missing point.

##### 4.4.1 Extrapolation technique for missing data recovery

We first present results obtained from strict Lagrange extrapolation (SE-Lag) method and compare them with data hold methods.

Table 4.5.  
Nstab using data holds and SE-Lag

%miss	0	1	2	3	4	5	6	7	8	9
ZOH	22	19	15	14	10	8	7	7	8	6
FOH	22	16	11	8	6	5	5	3	4	5
SE-Lag	22	15	12	10	6	6	7	7	10	12

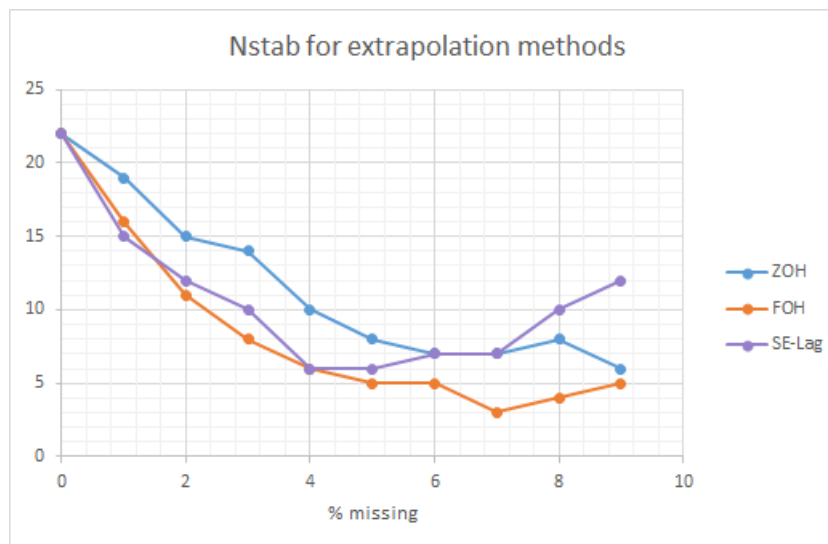


Fig. 4.6. Number of events stabilized for extrapolation methods

As seen from the Table 4.5 and its corresponding Graph 4.6, the SE-Lag method stabilizes less number of events than ZOH method below 7% of missing data. However, SE-Lag appears to be superior than FOH method. Comparing the success rates from the Graph 4.7 shows that the success rate of SE-Lag method is relatively lower than ZOH and FOH method.

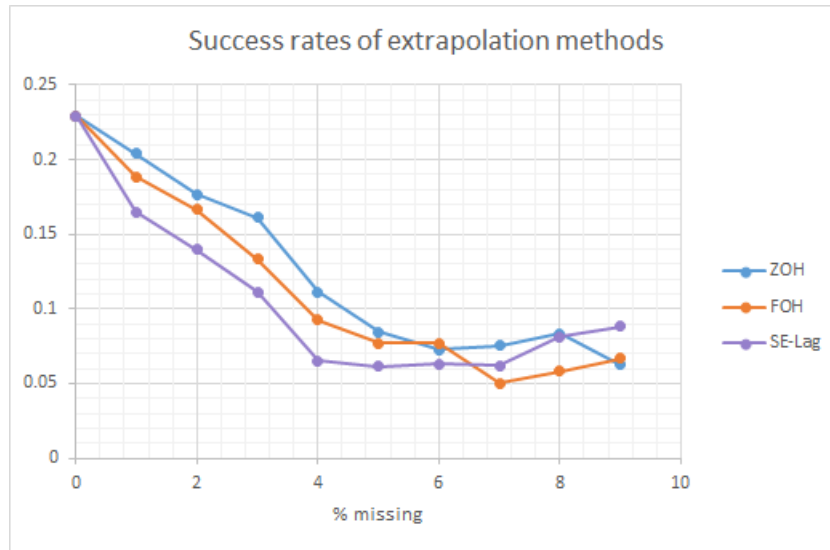


Fig. 4.7. Success rates of extrapolation methods

#### 4.4.2 Interpolation technique for missing data recovery

The methods discussed so far used only extrapolation technique to recover a missing value for which ZOH method proved to be the most advantageous one. Method like First Order Lagrange Interpolation (FO-Lag) discussed in this section additionally considers interpolation from one succeeding data point to recover a missing sample.

Preliminary results for FO-Lag method were obtained by never allowing consecutive missing samples. This was done by generating random sequence with a certain percentage missing and using interpolation to fill in some missing samples to avoid the occurrence of consecutive missing samples. Figure 4.8 compares the number of events stabilized by FO-Lag method assuming no consecutive missing data. Plots for extrapolation methods are also included for comparison.

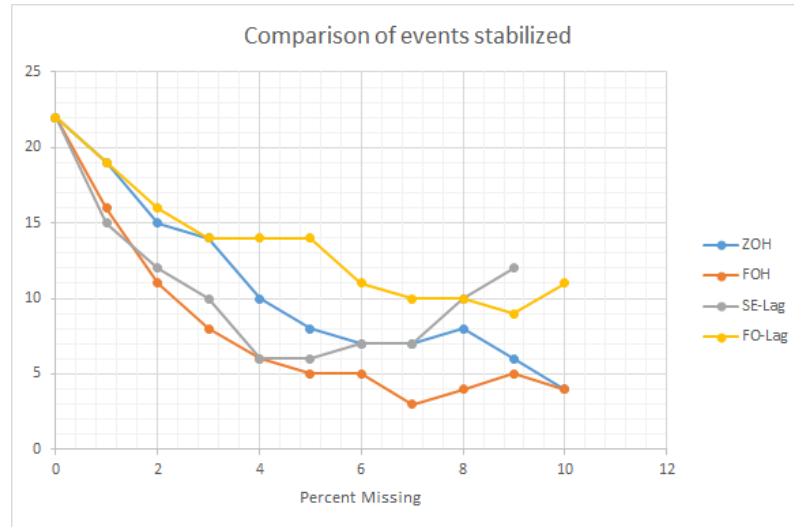


Fig. 4.8. Events stabilized by FO-Lag without handling consecutive missing samples

If two or more consecutive samples are assumed missing, it requires the controller to go farther for interpolation which increases the control actuation time. Thus, the next approach here is to modify the interpolation method to handle up to three consecutive missing samples only.

### Handling of consecutive missing samples in FO-Lag

For two consecutive missing samples within 6 seconds, a third sample is obtained for interpolation of the missing values. If three samples are missed in a row, the controller is set to ‘No control’ and the event is kept as it is. The algorithm checks for three consecutive missing samples within the first 1 seconds of simulation only.

As the simulation for test set data involves fault clearance at 0.67 seconds followed by control actuation within the next 0.167 seconds plus a short actuation delay, the control would already be applied before 1 second in a normal scenario. Setting the controller to ‘No control’ thus makes sense if applied before 1 second.

Table 4.6 shows different control performances for FO-Lag method where consecutive missing data handling technique is applied to 480 SLG contingencies up to 10% missing data. The success rate goes up as the percent missed increases from 5-10% because the rule about three consecutive samples makes the number of controlled events go down.

Table 4.6.  
Performance of FO-Lag with consecutive missing data handling

Percent missed	0	1	2	3	4	5	6	7	8	9	10
Events Stabilized	22	18	15	11	8	6	6	6	5	4	4
Unnecessary Controls	6	7	17	25	28	30	25	22	16	13	12
Success Rate	0.23	0.21	0.16	0.13	0.11	0.09	0.11	0.13	0.14	0.15	0.17

With consecutive missing data handling, the number of events stabilized reduced from the dashed line to the solid yellow line as shown in Figure 4.9. Plots for extrapolation methods are included for comparison.

### **Handling of consecutive missing samples for Extrapolation methods**

Even though the algorithm for extrapolation method automatically handled a few more than three consecutive missing samples, for a fair comparison, we also modified the ZOH, FOH and SE-Lag methods to respond the same way as FO-Lag to three consecutive missing samples. The controller output was set to ‘No control’ if three samples missed in a row. The ZOH method modified in this way was labelled as ‘MZOH’. Similar nomenclature was applied for every other methods.

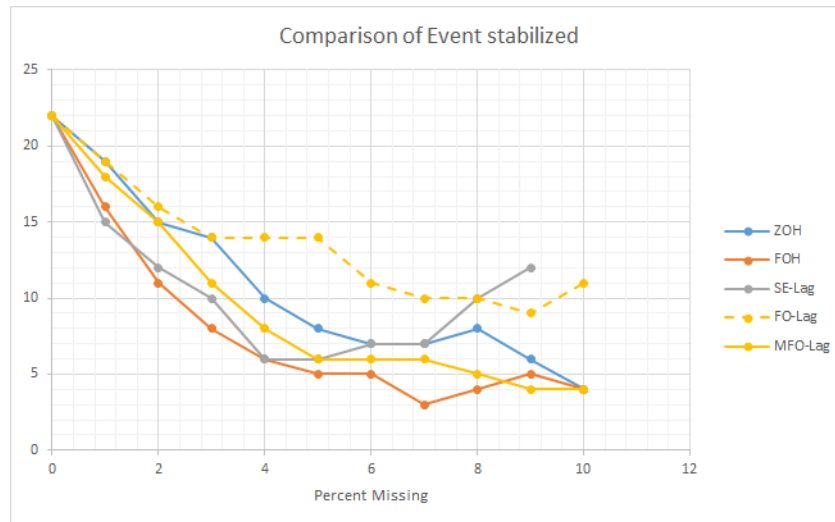


Fig. 4.9. Events stabilized by FO-Lag handling consecutive missing data

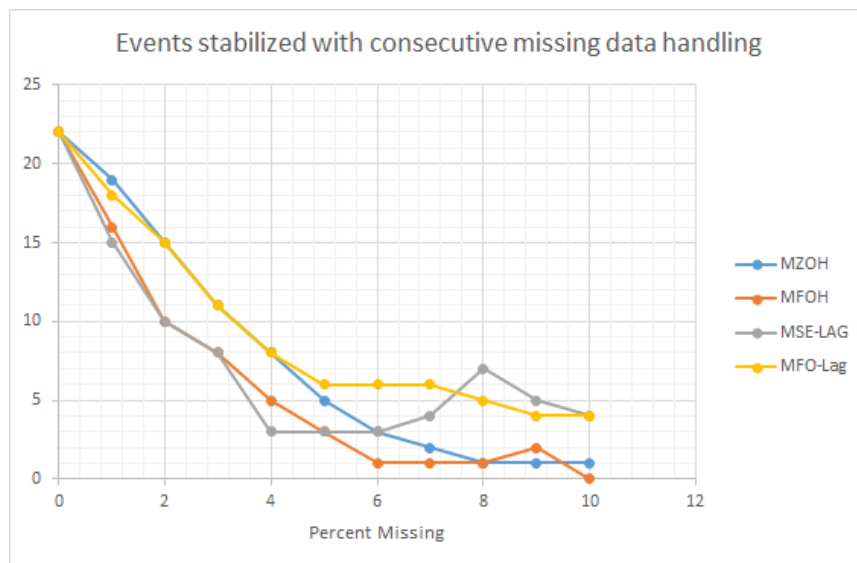


Fig. 4.10. Events stabilized with consecutive missing data handling



Figure 4.10 compares the number of events stabilized with consecutive missing data handling. The results show that MZOH method is as good as MFO-Lag until 5% missing data. Other methods like MFOH and MSE-Lag gives lower performance. Figure 4.11 compares the success rates of these methods.

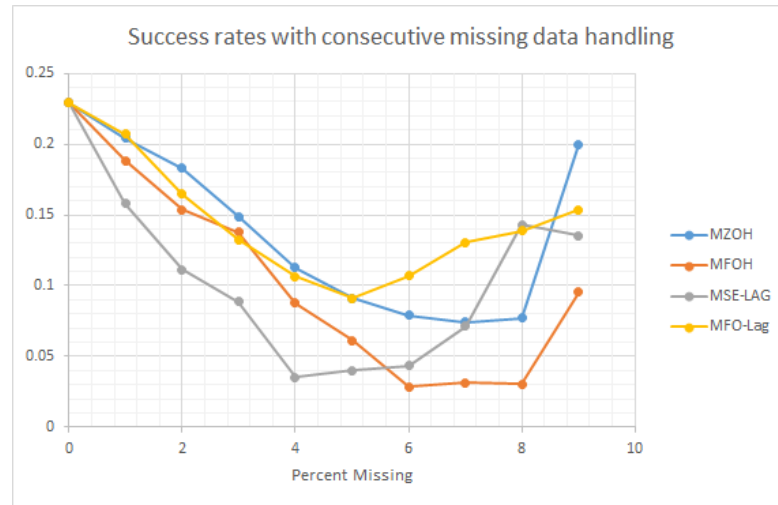


Fig. 4.11. Success rates with consecutive missing data handling

From Figure 4.11, the success rate of each method seems to be increasing after around 5% missing data. This is because after 5% missing samples, the rule about three consecutive missing samples reduces the number of events with control action. For the same number of events stabilized, lower number of events controlled gives higher success rates.

We stated in the beginning that for FO-Lag interpolation method, the effect of consecutive missing samples leads to increase in control actuation time. The plot of average control time as shown in Figure 4.12 instead shows a decrease in control actuation time. This indicated a problem with the simulations and in further examination we found that control was being ordered at the beginning of the simulation a large number of times when the percent missing was large. It appears that the problem arises from consecutive missing samples. The results in this thesis are probably more accurate for smaller percentages missing.

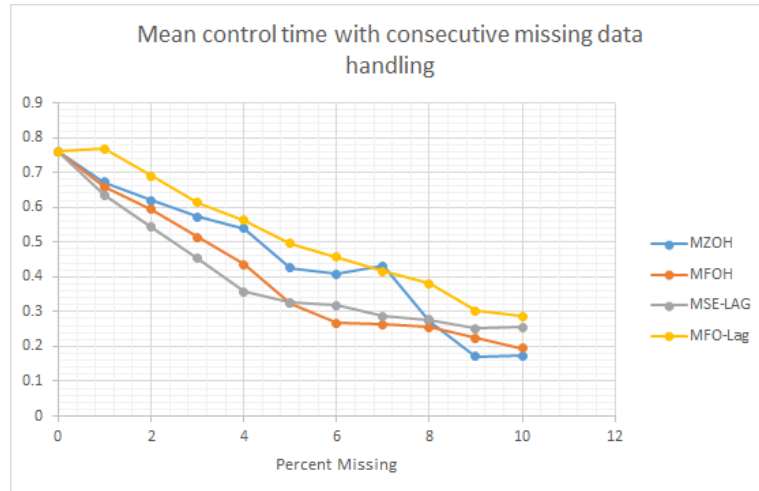


Fig. 4.12. Mean control time with consecutive missing data handling

#### 4.5 Performance using Bootstrap Averaging Method

Before we present results from bootstrap averaging method, it is worth recalling that the one-shot control combination used in [1] [2] do not destabilize any of the 480 SLG events. The data hold methods and Lagrange polynomial methods also do not destabilize any of the 480 SLG events. However, the results from bootstrap averaging led to a few events destabilized.

We had tried the method first by selecting 20 recent data samples from 5 randomly selected PMUs. The results showed up to 5 events destabilized and therefore had to be discarded. Looking at a pattern from Lagrange interpolating polynomial method, we next tried lowering the data samples substantially. Taking only 4 data samples each from 5 randomly selected PMUs, the number of stabilized events increased but at a cost of destabilizing up to 2 events as shown in the Table 4.7. Moreover, control actuation occurred for larger number of events including substantially increased unnecessary controls like in Figure 4.13. The bootstrap averaging method started off with lower success rates.

Table 4.7.  
Events destabilized by bootstrap averaging

%miss	0	1	2	3	4	5	6	7	8	9	10
Nstab	22	19	17	11	19	21	33	36	30	36	42
NDestab	0	0	0	0	0	0	1	0	2	1	0
Ncntrl	96	111	140	162	193	219	246	272	292	313	334
NUnne	6	32	58	89	113	126	125	145	162	164	189

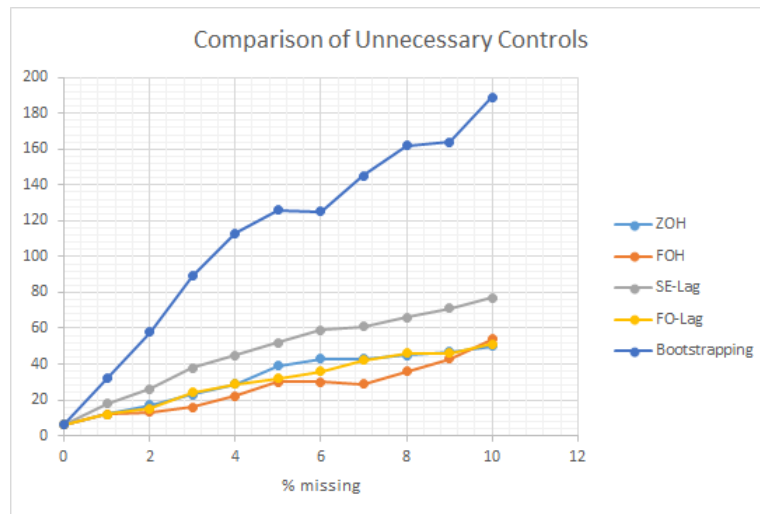


Fig. 4.13. Higher unnecessary controls for Bootstrap Averaging

One can try altering the size of bootstrap to see if the results improve. However, for this thesis, we assert that previous techniques like extrapolation or interpolation yield better results as they do not force the control to destabilize an already stable event.

Simulation results from methods such as mean value method that takes the average of all data samples and places it as the missing value, were fairly discarded as it increased the number of destabilized events with higher percentage of data missed as shown in Table 4.8 below.

Table 4.8.  
Events destabilized by Mean Value method

%miss	0	1	2	3	4	5	6	7	8	9	10
Nstab	22	19	17	22	20	25	26	34	30	39	39
NDestab	0	0	0	1	2	2	0	1	2	3	1

#### 4.6 Results with Control Combination CC2

Simulations using control combination CC2 generate similar plots. The number of events controlled, the number of unnecessary controls, and the mean control time are exactly same for both CC1 and CC2 for every percentage of missing data. The plots for number of events stabilized and success rates are shown in Figures 4.14 and 4.15.

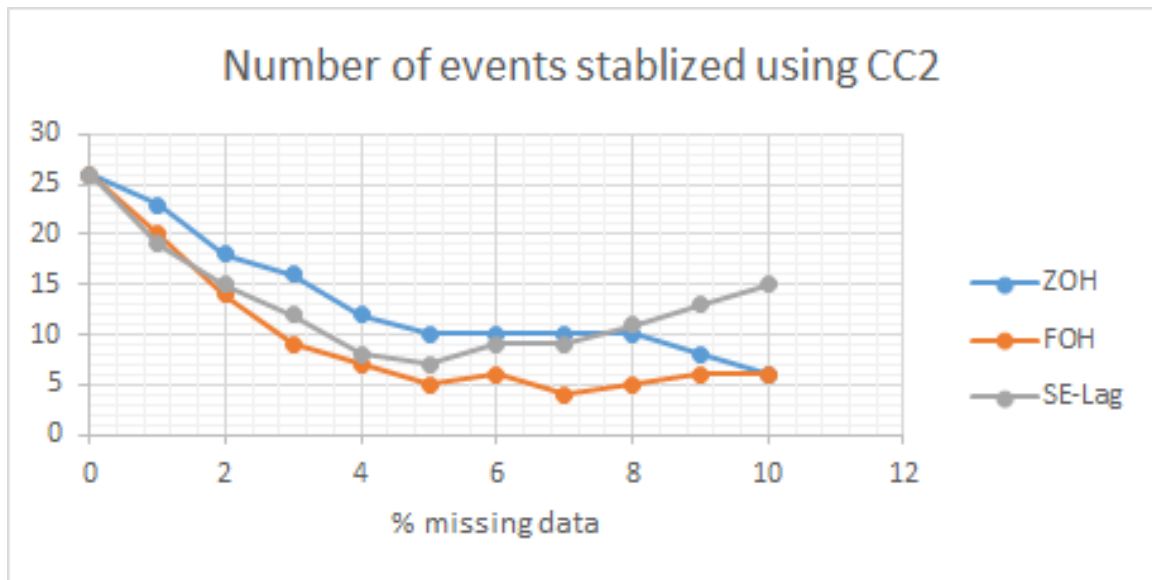


Fig. 4.14. Number of events stabilized with CC2

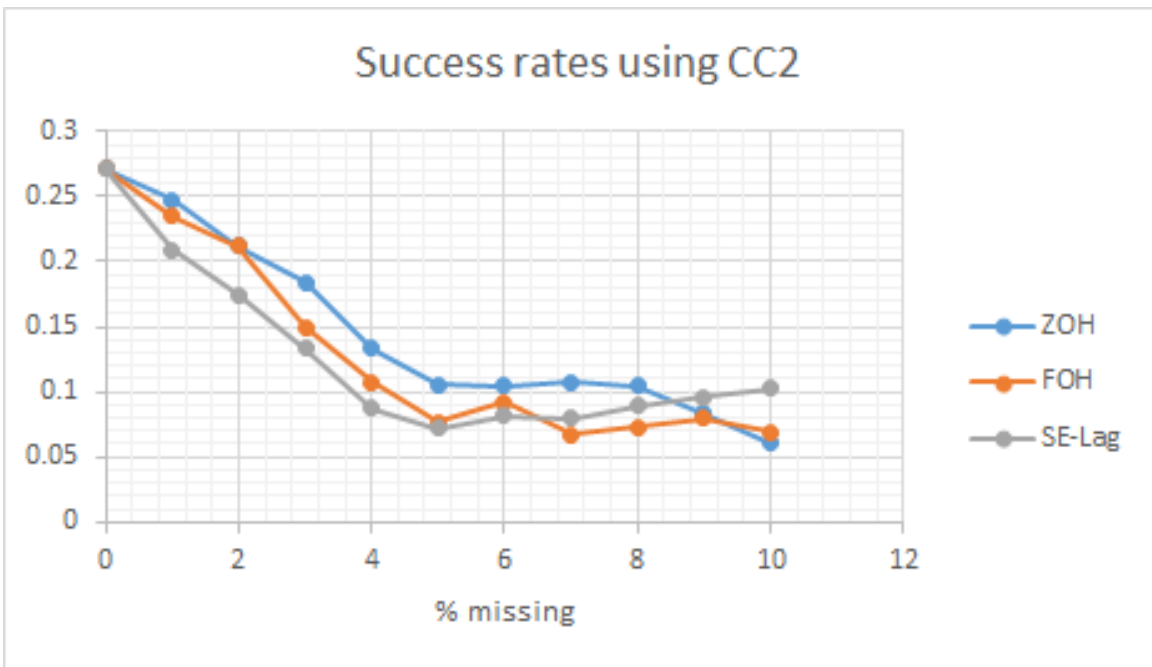


Fig. 4.15. Success rates with CC2

## 5. CONCLUSION AND FUTURE ENHANCEMENT

Analyzing a sufficient number of PMUs phase angle data retrieved from TSAT [12] shows the bus angles change relatively slow and hence a missing value can be recovered using a simple extrapolation technique. In our case, the zero order hold (ZOH) method proves to be advantageous to recover missing values in a response based scheme as in [1] [2].

The first order Lagrange interpolating polynomial method yielded as good performance as ZOH combined with a similar rule about three consecutive missing data. However, methods involving interpolation like this are prone to increase control actuation time with higher consecutive data missing. The plot of average control time instead shows a decrease in control actuation time. This indicated a problem with the simulations and we concluded that the results in this thesis are probably more accurate for smaller percent missing. Future work will include repeating the calculations in this thesis after resolving the problem with the simulations.

Methods like bootstrap averaging undoubtedly increased number of stable events but at the cost of significantly increasing unnecessary controls and destabilizing one or two events.

As a future enhancement of the thesis, one can research techniques of recovering missing data in a realistically noisy PMU environment. The task becomes more challenging as recovering a data sample in presence of noise does not guarantee a noise-free imputation. Therefore, one should come up with an adequate degree of filtering methods without losing the significant frequency events.

In this thesis, we have tried to use two different one-shot control combinations CC1 and CC2. Researchers can change one shot control combinations and see if the missing data handling techniques used in this thesis works well with other control combinations.

Another scope of extension of this thesis is to evaluate the methods used here in a power system with different operating conditions or in a relatively stressed network with more buses. In a stressed network, small variation of data may result in misclassification of a stable event as unstable. While we assume that the proposed zero order hold (ZOH) method works well in this situation, more precise techniques can still be investigated to impute the missing values.

## REFERENCES



## REFERENCES

- [1] Q. Gao and S. M. Rovnyak, "Decision trees using synchronized phasor measurements for wide-area response based control," *IEEE Transactions on Power System*, vol. 26, May 2011.
- [2] S. Rovnyak, D. W. Longbottom, D. C. Vasquez, and M. N. Nilchi, *Response Based Event Detection for One Shot Wide-area Stability Controls*. IET, 2017, ch. 8.
- [3] J. D. L. Ree, V. Centeno, J. Thorp, and A. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 20–27, June 2010.
- [4] PMU data quality. [https://www.naspi.org/sites/default/files/reference/\\_documents/PARTF\\\_WhitePaper\\\_20170314\\\_Final\\\_PNNL\\\_26313.pdf](https://www.naspi.org/sites/default/files/reference/_documents/PARTF\_WhitePaper\_20170314\_Final\_PNNL\_26313.pdf)?fileID=1689, Last accessed 1/10/2019.
- [5] C. Hunag, F. Li, and e. L. Zhan, "Data quality issues for synchrophasor applications part ii: Problem formulation and potential solutions," *J. Mod. Power System Clean Energy*, 2016.
- [6] M. Wang, "Data quality management of synchrophasor data in power system by exploiting low-dimensional models," *51st Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD*, pp. 1–2, 2017.
- [7] Z. Zhang and K. T. Chong, "Comparison between first-order hold with zero order hold in discretization of input-delay nonlinear systems," *International Conference on Control, Automation and Systems, Seoul*, pp. 2892–2896, 2007.
- [8] Data hold methods. <http://web.cecs.pdx.edu/~tymerski/ece452/Chapter3.pdf>, Last accessed 01/12/2019.
- [9] P. Kundur, *Power System Stability and Control*, 22nd ed. McGraw Hill Education (India) Private Limited, 2017.
- [10] J. T. S.M. Rovnyak, C.W. Taylor, "Real-time transient stability prediction- possibilities for online automatic database generation and classifier training," *IFAC Control of Power Plants and Power System*, Mexico 1995.
- [11] S. Rovnyak, C. Taylor, J. Mechenbier, and J. Thorp, "Plans to demonstrate decision control tree using phasor measurements for hvdc fast power changes," *Faults and Disturbance Analysis and Precise Measurements in Power Systems Conference*, November 1995.
- [12] Powertech labs, inc. TSAT: Transient security assessment tool. <http://www.dsatools.com/wp-content/uploads/2017/07/TSAT.pdf>, Last accessed 12/15/2018.

- [13] Powertech labs, inc. PSAT: Powerflow and short circuit analysis tool. <http://www.dsatools.com/wp-content/uploads/2017/07/PSAT.pdf>, Last accessed 12/15/2018.
- [14] MATLAB. <https://www.mathworks.com/products/matlab.html>, Last accessed 12/05/2018.
- [15] C. Taylor, J. Haner, L. Hill, W. Mittelstadt, and R. Cresap, "A new out-of-step relay with rate of change of apparent resistance augmentation," *IEEE Transactions on Power Apparatus and System*, vol. PAS-102, no. 3, pp. 631–639, March 1983.
- [16] S. Rovnyak, S. Kretsinger, J. Thorp, and D. Brown, "Decision trees for real-time transient stability prediction," *IEEE Transactions on Power Systems*, vol. 9, no. 3, pp. 1417–1426, August 1994.
- [17] K. Mei and S. M. Rovnyak, "Response based decision trees to trigger one shot stabilizing control," *IEEE Transactions on Power System*, vol. 19, pp. 531–537, February 2004.
- [18] S. M. Rovnyak, G. Li, and K. Mei, "One-shot controls for preventing loss of synchronism," *IEEE Power Engineering Society General Meeting*, 2003.
- [19] M. Nilchi, "Electric utility planning methods for the design of one shot stability controls," *MS Thesis, Purdue University, Indianapolis*, 2012.
- [20] S. Rovnyak, M. Nilchi, D. Longbottom, and D. Vasquez, "Angle stability predictive indices," *IEEE Power and Energy Society General Meeting, San Diego, CA*, pp. 1–6, 2012.
- [21] G. Li and S. M. Rovnyak, "Integral square generator angle index for stability ranking and control," *IEEE Transactions on Power System*, vol. 20, May 2004.
- [22] N. Le and W. Benjapolakul, "A data imputation model in phasor measurement units on bagged averaging of multiple linear regression," *IEEE Access*, vol. 6, pp. 39 324–39 333, 2018.
- [23] P. Manembu, A. Kewo, and B. Welang, "Missing data solution of electricity consumption based on langrange interpolation case study:intelligensia data monitoring," *2015 International Conference on Electrical Engineering and Informatics (ICEEI), Denpasar*, pp. 511–516, 2015.

## APPENDICES

## APPENDIX 1

### SOLUTION REPORT -- SWITCHING SPECIFICATION FILE

```
=====
TIME  BUS No.  STATION NAME  GID  CHANGE  BUS No.  BUS NAME  MISMATCH
---  - - - -  - - - -  - - -  - - - -  - - - -  - - - -
0.600  64  MONTA G120.0  1  0.00  75  NORTH G320.0  0.0000
```

**SWITCHING ACTION : AT TIME 0.600 SECONDS - LINE TO GROUND FAULT**

BUS : 166 MALIN5 500

ZERO SEQ. : R = 0.00020 X = 0.00400 P.U.

NEG. SEQ. : R = 0.00010 X = 0.00200 P.U.

**SWITCHING ACTION : AT TIME 0.667 SECONDS - LINE TO GROUND FAULT CLEARED**

BUS : 166 MALIN5 500

**SWITCHING ACTION : AT TIME 0.667 SECONDS - LINE REMOVED**

FROM BUS : 166 MALIN5 500

TO BUS : 167 MALIN6 500

**SWITCHING ACTION : AT TIME 0.883 SECONDS - ADMITTANCE ADDED**

BUS : 33 CA230 230

AMOUNT : G = 5.00000 B = 2.00000 P.U.

**SWITCHING ACTION : AT TIME 0.883 SECONDS - ADMITTANCE ADDED**

BUS : 148 VALLEY 500

AMOUNT : G = -5.00000 B = -2.00000 P.U.

**SWITCHING ACTION : AT TIME 0.883 SECONDS - ADMITTANCE ADDED**

BUS : 65 MONTANA 500

AMOUNT : G = 5.00000 B = 2.00000 P.U.

**SWITCHING ACTION : AT TIME 0.883 SECONDS - ADMITTANCE ADDED**

BUS : 47 RINALDI 230

AMOUNT : G = -5.00000 B = -2.00000 P.U.

```
TIME  BUS No.  STATION NAME  GID  CHANGE  BUS No.  BUS NAME  MISMATCH
---  - - - -  - - - -  - - -  - - - -  - - - -  - - - -
0.900  75  NORTH G320.0  1  0.43  12  NAVAJO 226.0  0.0024
0.917  75  NORTH G320.0  1  0.43  12  NAVAJO 226.0  0.0012
0.933  35  BRIDGER2 22  1  0.48  75  NORTH G320.0  0.0009
0.950  35  BRIDGER2 22  1  0.56  75  NORTH G320.0  0.0010
0.967  35  BRIDGER2 22  1  0.62  75  NORTH G320.0  0.0012
```