VOLTAGE AND FREQUENCY RECOVERY IN POWER SYSTEM AND MICROGRIDS USING ARTIFICIAL INTELLIGENT ALGORITHMS

SOLEIMAN RAHMANI

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

GRADUATE PROGRAM IN ELECTRICAL AND COMPUTER ENGINEERING
YORK UNIVERSITY
TORONTO, ONTARIO

April 2019

ABSTRACT

The increase in demand and power transfer between utilities in the modern power system as well as penetration of distributed generations (DGs) which produce electricity sporadically in smart grids raise the voltage stability concerns in electric power networks. Noticeable challenges such as natural complexity and dynamics of the power system as well as sporadic generation of renewable energy sources (RESs) impose power grids to operate closer to the operating limit, which ultimately results in voltage instability issues. The driving force of voltage instability is generally loads and thus analyzing their effect on power system is of great importance in voltage stability assessments.

This thesis developed an advanced assessment tools to predict and recover the power system voltage margin to the acceptable values during the occurrence of a disturbance in the network. First, the effect of disturbance in islanded microgrids are analyzed using power factor-based power-voltage curves and a comprehensive under voltage-frequency load shedding (UVFLS) method is proposed as a last resort in order to restore the system voltage and frequency. It is shown that by considering power factor of busbars and implementing proposed UVFLS technique, optimum load shedding can be achieved.

Thereafter, the effect of disturbance in conventional power system is investigated by introducing a phenomenon called fault induced delayed voltage recovery (FIDVR) and comprehensive real-time FIDVR assessments are proposed to employ appropriate emergency control approaches i.e. load shedding as fast as possible to maintain the system voltage margins within the desired range. In the first step, linear and polynomial regression techniques have been used for predicting the FIDVR duration. In the next step, advanced FIDVR assessment is implemented which simultaneously predicts whether the event can be classified as FIDVR or not and also predicts the duration of FIDVR with high accuracy.

ACKNOWLEDGMENTS

I would first like to express my deep sense of gratitude to my mentor and supervisor, Dr. Afshin Rezaei-Zare for his keen interest and dedication to help me. The door to Prof. Rezaei office was always open whenever I ran into a trouble spot or had a question about my research or writing. His timely advice and scientific approach have helped me a lot to accomplish this chapter of my life.

I must express my very profound gratitude to my beloved parents and to my little lovely sister for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

I would like to thank Dr. John Lam and Dr. Solomon Boakye-Yiadom for their supports and also for their valuable comments which enhanced the quality of this thesis.

And finally, last but by no means least, also to my friends in the Bergeron Centre for Engineering Excellence… it was great sharing laboratory with all of you during last two years. Thanks for all your encouragement!

Author
Soleiman Rahmani

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABBREVIATION

UVLS	Under voltage load shedding
UFLS	Under frequency load shedding
UVFLS	Under voltage-frequency load shedding
LSA	Load shedding amount
UVA	Under voltage load shedding amount
UFA	Under frequency load shedding amount
OFA	Over frequency load shedding amount
P-V curve	Power-voltage curve
PF-based P-V curve	Power factor-based power-voltage curve
IBMG	Inverter-based microgrid
DG	Distributed generation
CPL	Constant power load
CCL	Constant current load
CIL	Constant impedance load
PF	Power factor
PLN	Piecewise linear nose curve
FIDVR	Fault induced delayed voltage recovery
ML	Machine learning
PG	Pair grid plot
LR	Linear regression
MSE	Mean squared error

RMSE	Root mean squared error
MAPE	Mean absolute percentage error
R^2	R squared
R^{*2}	Adjusted R squared
SST	Total sum of squares
PDI	Active power deviation index
QDI	Reactive power deviation index
PISI	Active power increment slope index
QISI	Reactive power increment slope index
RF	Random forest
RFR	Random forest regression
RFC	Random forest classification
DT	Decision tree
OOB	Out of bag score

Chapter 1 : THESIS OVERVIEW

1.1. Thesis Overview

The second chapter of this thesis deals with the effect of disturbances in microgrids and introduces simple and complete under voltage-frequency Load Shedding (UVFLS) schemes to achieve proper load shedding amounts (LSAs) in an islanded Inverter-Based MicroGrids (IBMGs) considering power and power factor (PF) of diverse load models such as constant power, constant current and constant impedance loads. The proposed method employs a fast time-step simulation approach, based on the complete state-space model of IBMG in order to obtain the transient and steady state responses of the system. PF-based power-voltage (P-V) curves are obtained considering loads PF. Ultimately, LSA is obtained according to the under-voltage LS (UVLS) and under frequency LS (UFLS) values. To demonstrate the effectiveness of the proposed simple and complete LS methods, the performances of the techniques are analyzed in a test islanded IBMG using MATLAB. The simulation results clearly show that the proposed LS methods are more accurate than the previous studies and they can provide optimum LSA according to the loads PF and enhance the power system stability considerably.

The third chapter discusses the stalling of constant torque induction motor loads caused by system faults that may lead to a phenomenon known as fault induced delayed voltage recovery (FIDVR) which is a critical threat to modern power systems. FIDVR can cause significantly depressed local voltage for several seconds after the fault clearance and in

more severe situations can also lead to widely cascaded system failure. Hence, predicting the FIDVR duration after the system fault can play an essential role in maintaining the power system voltage stability. To have the real-time assessment and prediction of FIDVR phenomenon, this study presents a data-driven multi-variable machine learning (ML)-based Decision-making method. Before building the model, a comprehensive feature analysis is accomplished and important features with high correlation with FIDVR duration are selected. The data of power quality recording devices (PQubes) are used for online ML model developing to rapidly predict the FIDVR duration following a system disturbance. To do so, linear regression and polynomial regression have been implemented in Python software along with power system data for online model developing. Using the developed model, FIDVR duration can accurately be obtained in real-time without using complicated load models. Aforementioned regression models result in simple programming and fast prediction and on the other hand, using multiple features for model development enhance the model accuracy. As a result, the FIDVR prediction can be accomplished very fast, yet with acceptable accuracy. Accurate and fast prediction provides the opportunity for consecutive last resort controls to restore the system voltage to the acceptable range.

The forth chapter proposes a probabilistic time-series data-driven multi-variable simultaneous classification and regression decision-making to assess the event and categorize it as FIDVR or non-FIDVR and predict the FIDVR duration if the event is classified as FIDVR. In real world, applying inaccurate prediction may result in false alarm which lead to unnecessary emergency control i.e. load shedding or missing an alarm which in severe cases may lead to voltage collapse. In order to enhance the machine

learning assessment accuracy, a more advanced method is designed to improve the real-time FIDVR assessment accuracy and speed. A robust validation technique is used to enhance the real-time prediction accuracy. In this chapter, an advanced feature analysis is accomplished and important features with high correlation with FIDVR duration are selected. The real-time data of power quality recording devices (PQubes) are used for online ML model developing to rapidly predict the FIDVR duration following a system disturbance. The proposed technique also tested in a test system modeled in EMTP-RV. Note that artificial intelligence algorithms have been used in different manners in chapter 2 compared to chapter 3 and 4. In chapter 2, a automatic algorithm is proposed which can select the best PF-based P-V curve according to the characteristics of the system. In chapter 3 and 4, machine learning algorithms have been used for FIDVR prediction. However, all these algorithms are under the umbrella of artificial intelligence.

Ultimately, the fifth chapter summarizes the thesis contributions and identified future works.

Chapter 2 : VOLTAGE AND FREQUENCY RECOVERY IN AN ISLANDED INVERTER-BASED MICROGRID CONSIDERING LOAD TYPE AND POWER FACTOR

2.1. Introduction

Ever growing electrical demand and penetration of distributed generations (DGs) which produce electricity sporadically, bring power systems closer to their stability limit. These issues result in an increasing risk of voltage and frequency instability in power system [1] -[2]. Aforementioned challenges are even of paramount importance in islanded microgrids (MGs) [3]. If disturbances such as sudden load increment and outages of DGs occur in such a way that the load demand exceeds the total generation and the available DGs operate at maximum power, the system will observe noticeable instabilities. The resulting active and reactive power imbalance can let the system voltages and frequency to deviate from their desired operating range. This issue may cause system collapse in such a way that standard system controls fail to stop this deviation. After applying all of the control measures, under voltage load shedding (UVLS) and under frequency load shedding (UFLS) would be the last inevitable efforts to maintain the stability of the system [4].

According to [5] and [6], traditional load shedding (LS) schemes are not capable of dealing with the combined instabilities. However, the majority of the existing studies are based on either frequency or voltage LS schemes. In UFLS approach, merely the frequency information is considered which may have adverse consequences on other

characteristics of the system. Similarly, UVLS utilizes local bus voltage values and may have unanticipated effects on the system characteristics [4]. In [7], [8], [9], frequency and its rate of change were used for UFLS calculation. Nonlinear mathematical programming and discretized differential-algebraic power system equations were combined to estimate the optimal LSA in [10]. In [11], a centralized adaptive UFLS scheme for smart grid has been investigated using synchronous phase measurement units. In [12], a distributed load shedding algorithm based on sub-gradient method via a wireless network was proposed for balancing the supply– demand and reducing the LS amount. [13] has proposed new scheme for frequency regulation by coordinating the operation of fast-responding inverter-based distributed energy resources (DERs) with the slow-acting gensets. The unequal transient LS between gensets and inverter-based DERs is redistributed to prevent system collapse as well as achieve desired frequency regulation. In [14], a two-level control method is presented to facilitate decentralized management of active power deficiencies in remote sustainable MGs. The primary level is responsible for fulfilling an UFLS action based on a developed multilayer droop structure. [15] has proposed a novel strategy for stand-alone multi-MGs operation considering flexible frequency operation in order to reduce the operation cost and the amount of LS.

There are research works about UVLS in which the voltage information is used for calculating LSA [16], [17]. A UVLS algorithm based on the rate of voltage recovery and predicted time was proposed in [18]. [19] has come up with a response-based system integrity protection scheme for adaptive UVLS in large interconnected systems.

While various papers utilized UFLS or UVLS for voltage and frequency recovery, there exist few studies such as [20], [21], [22], [23], [24], [25] which have proposed LS

for simultaneous voltage and frequency recovery. A multi-step hybrid islanding methodology was proposed in [20] for emergency operation of an industrial park. An under frequency/under voltage load shedding scheme is initiated, where a predetermined set of lower priority loads are shed in order to stabilize the individual islands. In [21], the magnitude of sub-transmission bus voltages and also static voltage stability margins of the buses were utilized in order to select the load shedding amount (LSA). The method of [22] employs voltage information available through a synchrophasor-based wide area monitoring and control system so as to find proper LS locations. In [23], through the dynamic simulations of power system network, the sensitivity of power system voltage and frequency response to the disturbance location is incorporated. Reference [24] proposed decentralized LS in which the instantaneous voltage deviation of load buses was used to determine the frequency thresholds of LS relays. In [25], an under-voltage frequency LS (UVFLS) scheme was proposed based on the piecewise linear nose (PLN) curve. However, these methods [21], [22], [23], [24], [25] neither consider the effects of loads reactive power and PF nor take into account the impacts of load types on the calculated LSA. Moreover, these adaptive strategies are based on pre-disturbance sensitivity factors and, hence, do not take into account the dynamic changes due to a contingency or system configuration change. If the load power factor changes during the disturbance, no updated LSA is proposed and the load shedding is based on the pre-disturbance LSA values.

In view of the shortcomings of the existing LS schemes, this study proposes a new UVFLS strategy which simultaneously regulates frequency and bus voltages within the permissible values in an IBMG consisting of various load types and PFs. In this approach,

voltage and frequency are measured and LSA is determined based on the developed PF-based P-V curves. In parallel, frequency droop characteristics are used for calculating the LSA obtained from UFLS. In the first stage, a fast time-step simulation technique is utilized which provide online transient and steady-state responses of the IBMG before, during, and after the disturbance. Using the proposed UVFLS method with the fast time-step simulation, the LSA computation time following contingencies will be decreased [26]. During the system operation, PF-based P-V curves are calculated at pre-specified time intervals, e.g. every minute, according to the characteristic of the load on each bus and the system configuration. In the simple LS method, PF-based P-V curves are derived before the disturbance and are used for UVLS. However, in the complete UVFLS, several PF-based P-V curves with different PFs are obtained, two closest PF-based P-V curves to loads PF are selected, and the LSA will be calculated according to the proposed algorithm.

2.2. **IBMG Modeling**

This chapter studies the autonomous voltage source inverter-based microgrid (IBMG). In islanded operation of IBMGs, merely distributed generations (DGs) carry responsibility of maintaining voltage and frequency of the system within the desired ranges. During the islanded operation of IBMG, DGs must meet the demand without over-loading of inverters, and any load change must be controlled by DGs. This operating mode requires crucial control techniques since any malfunction in the system can deteriorate the overall performance of the IBMG [27]. This study employs real power/frequency and reactive

power/voltage droop control which is an extensively popular power sharing method and shares demand between DGs according to DGs droop characteristic. A complete state-space of the IBMG is obtained using the method of [27]. A linear state-space model for the islanded IBMG and a time-step simulation are utilized [26], in order to obtain operating point as well as transient response of the system at any time for making LS decisions. According to [27], state-space model of IBMG divides components of the system into three sub-systems which are inverters, network and loads. Each inverter is composed of power sharing controller, voltage and current controller, output filter and coupling inductor. Fig. 2-1 shows the block diagram of the complete model of a VSI-based DG and its controllers.

By developing a linear model for each subsystem of IBMG and utilizing Taylor expansion, the general state-space equation of the system can be linearized as follows [26]:

$$\dot{x}(t+\Delta t) = A(t)x(\Delta t) + B(t)U \tag{2-1}$$

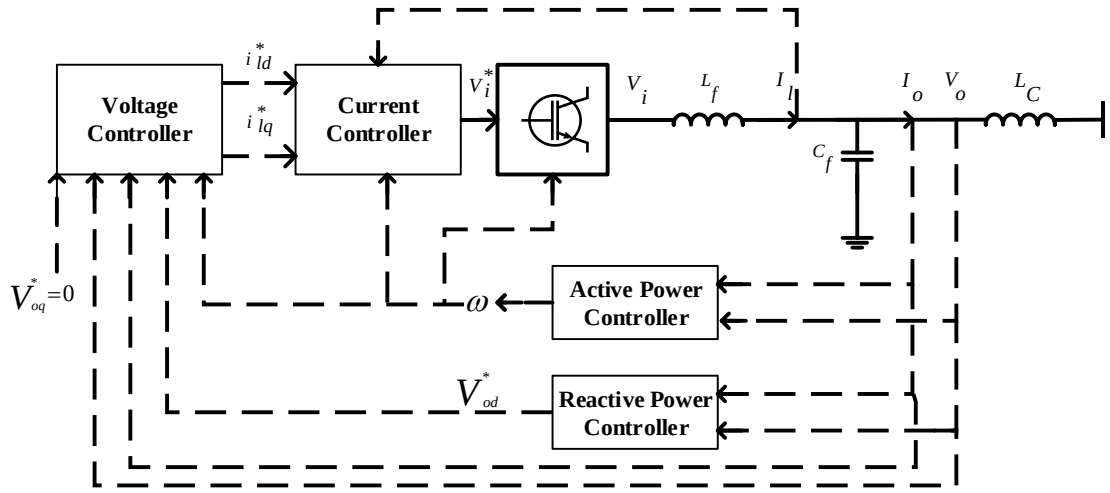

Fig. 2-1. Block diagram of VSI connected to an Islanded IBMG

where x and U are state and input vectors and A and B are system and input matrices. For n variables, matrix x is $n\times1$, U is $m\times1$, matrix A is $n\times n$ and B is $n\times m$ matrix. Matrix M, which is $(n+m)\times(n+m)$, is generated from A and B in order to create matrix exponential.

$$M = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}_{(n+m)\times(n+m)} \tag{2-2}$$

Assuming $M^*\Delta t$ has a full set of eigenvectors V with corresponding eigenvalues D, then $[V, D] = eig(M^*\Delta t)$ where function eig returns diagonal matrix D of eigenvalues and matrix V whose columns are the corresponding right eigenvectors, so that $M^*\Delta t^*V = V^*D$. Δt is time-step value. By deriving eigenvectors and eigenvalues of M, matrix exponential of M can be exploited by:

$$ExmP\ \ (\Delta t * M) = V * diag(\exp(diag(D)))/V \tag{2-3}$$

where V is the eigenvectors with the corresponding eigenvalues D. $diag\ (D)$ returns a

square diagonal matrix with the elements of vector *D* on the main diagonal. Matrix A_P will be generated from column *1* to *n* and row *1* to *n* of matrix exponential of $M^*\Delta t$ and Matrix B_P will be obtained from row *1* to *n* and column *n+1* to *n+m*. Finally, the response of the system is obtained at each time from:

$$x(t+\Delta t) = A_P(t)x(t) + B_P(t)U \tag{2-4}$$

Where A_P is the system matrix and matrix B_P is the input matrix. Using (2-4), time domain simulation is performed and all variables are obtained in each time instant. In other words, using this method not only the load flow solutions such as Newton-Raphson methods are not needed any more, but also transient response of the system can be derived at any time [28].

For having more realistic voltage stability investigation, it is necessary to have sufficiently proper models of the loads. In [26], [27] the relation between load active and reactive powers and the bus voltage are not included in the IBMG modeling. In this chapter, IBMG model contains various load models such as constant power loads (CPLs), constant current loads (CCLs) as well as constant impedance loads (CILs) with specified power factors (PFs) [2]. In CPLs, the active and reactive powers are independent of variations in the voltage magnitude. The voltage-power equation for this model can be written as (5) while n=0.

$$\frac{P}{P_0} = \left(\frac{V}{V_0}\right)^n \quad , \quad \frac{Q}{Q_0} = \left(\frac{V}{V_0}\right)^n \tag{2-5}$$

where *P* and *Q* are active and reactive power demand respectively, and P_0 and Q_0 are respectively active and reactive power demand at rated voltage V_0. In CCL model, the active and reactive powers vary directly with the voltage magnitude (n=1 in (2-5)).

Finally, by putting n=2 in (5), the powers vary with the square of the voltage magnitude which is the characteristic of the CILs. In this study, aforementioned models are included in linear state-space model of the IBMG and the proposed UVFLS technique.

To illustrate more note that the proposed load shedding method utilizes state-space model and time-domain methods to derive responses of the system i.e. bus voltages [26]. Based on the system characteristics, state-space equation can be updated and system responses change accordingly.

The complete state-space equation of the microgrid can be obtained according to the equation (1):

$$\dot{x} = A_{sys}\, x + B_{sys}\, U \tag{2-6}$$

Where A_{SYS}, B_{SYS}, X and U can be obtained as follows:

$$A_{sys} = \begin{bmatrix} A_{INV} + C_{INV} R_N M_{INV} D_{INV} & C_{INV} R_N M_{NET} & C_{INV} R_N M_{LOAD} \\ C_{NET} R_N M_{INV} D_{INV} & A_{NET} + C_{NET} R_N M_{NET} & C_{NET} R_N M_{LOAD} \\ C_{LOAD} R_N M_{INV} D_{INV} & C_{LOAD} R_N M_{NET} & A_{LOAD} + C_{LOAD} R_N M_{LOAD} \end{bmatrix} \tag{2-7}$$

$$B_{sys} = \begin{bmatrix} B_{INV} \\ 0 \\ 0 \end{bmatrix} \quad X = \begin{bmatrix} X_{INV} \\ i_{lineDQ} \\ i_{loadDQ} \end{bmatrix} \quad U = \begin{bmatrix} V_n \\ \omega_n \end{bmatrix} \tag{2-8}$$

Where A_{INVi}, B_{INVi}, C_{INVi}, D_{INVi}, A_{NETi}, C_{NETi}, A_{LOADi} and A_{LOADi} are shown below:

$$A_{INV} = \begin{bmatrix} -\omega_c a_1 & 0_{2\times2} & 0_{2\times2} & 0_{2\times1} & 0_{2\times2} & 0_{2\times2} & \omega_c a_2 \\ a_3 & 0_{2\times2} & 0_{2\times2} & 0_{2\times1} & 0_{2\times2} & -a_1 & 0_{2\times2} \\ K_{pv} a_3 & K_{iv} a_1 & 0_{2\times2} & 0_{2\times1} & -a_1 & a_4 & F a_1 \\ a_5 & 0_{2\times2} & 0_{2\times2} & 0_{2\times1} & 0_{2\times2} & 0_{2\times2} & 0_{2\times2} \\ K_1 a_3 & K_2 a_1 & K_{ic} a_1 & 0_{2\times1} & a_6 & a_7 & K_{pc} F a_1 \\ 0_{2\times2} & 0_{2\times2} & 0_{2\times2} & 0_{2\times1} & K_3 a_1 & a_8 & -K_3 a_1 \\ 0_{2\times2} & 0_{2\times2} & 0_{2\times2} & 0_{2\times1} & 0_{2\times2} & K_4 a_1 & a_9 \end{bmatrix}_{13\times13} \tag{2-9}$$

$$B_{INV} = \begin{bmatrix} 0 & 0 & 1 & 0 & K_{pv} & 0 & 0 & K_1 / L_f & 0_{1\times5} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0_{1\times5} \end{bmatrix}^T_{13\times2} \tag{2-10}$$

$$C_{INV} = \begin{bmatrix} 0_{6\times 1} & 0_{6\times 2} \\ m_p & 0_{1\times 2} \\ 0_{4\times 1} & 0_{4\times 2} \\ 0_{2\times 1} & \frac{-1}{L_c} a_5 \end{bmatrix}_{13\times 3} \tag{2-11}$$

$$D_{INV} = \begin{bmatrix} 0_{1\times 6} & -i_{oq} & 0_{1\times 4} & 1 & 0 \\ 0_{1\times 6} & i_{od} & 0_{1\times 4} & 0 & 1 \end{bmatrix}_{2\times 13} \tag{2-12}$$

$$A_{NETi} = \begin{bmatrix} -r_{linei} \diagup L_{linei} & \omega_1 \\ -\omega_1 & -r_{linei} \diagup L_{linei} \end{bmatrix} \tag{2-13}$$

$$C_{NETi} = \frac{1}{L_{linei}} \begin{bmatrix} \dots & 1 & 0 & \dots & -1 & 0 & \dots \\ \dots & 0 & 1 & \dots & 0 & -1 & \dots \end{bmatrix} \tag{2-14}$$

$$A_{LOADi} = \begin{bmatrix} -r_{loadi} \diagup L_{loadi} & \omega_1 \\ -\omega_1 & -r_{loadi} \diagup L_{loadi} \end{bmatrix} \tag{2-15}$$

$$C_{LOADi} = \begin{bmatrix} \dots & 1 & 0 & \dots \\ \dots & 0 & 1 & \dots \end{bmatrix} \tag{2-16}$$

The parameters used in (2-7)-(2-11) are shown in Table 2-3. All the matrixes are described and can be found in [26], [27]. MINV, MLOAD and MNET show the links among inverters, loads and lines with network buses, respectively. By applying a simple KCL in bus i, the elements of row i of these matrixes can be easily specified. Matrix MINV is of size 2m×2s which maps the (s) inverter connection points onto (m) network nodes. For example, if i-th inverter is connected at j-th node, the element MINV (j,i) will be 1 and all the other elements in that row will be zero. Similarly, MLOAD is of size 2m×2p maps (p) load connection points onto the (m) network nodes with 1. Matrix MNET of size 2m×2n maps the (n) connecting lines onto the (m) network nodes [26], [27]. Hence, if any

changes occur such as network size increase, variations on location of DGs/loads or scalability, the arrays of corresponding matrix will be changed and as a result, any connection between lines, loads and DGs will be implemented and the responses of the system will be updated accordingly.

2.3. **Fundamentals of P-V Curve**

The relationship between the voltage and the active power is described via power voltage (P-V) nose curve which is widely used as a tool for voltage stability and calculation of network loading margin. The equation for P-V curves can be achieved by using power flow equations for the receiving end in the system [2], [28]:

$$V = \sqrt{0.5 - q \pm \sqrt{0.25 - q - p^2}} \tag{2-17}$$

where p and q are active and reactive power demands respectively. According to (6), bus voltage is related to both active and reactive powers. PV-curves illustrate the dependency of the voltage on real power of a composite load for different PFs. Fig. 2-2 demonstrates family of P-V curves with different PF values highlighting significant impact of PF on the deriving P-V curve. To investigate the impact of various load types on voltage stability, the load characteristics (5) should be considered along with the P-V curve.

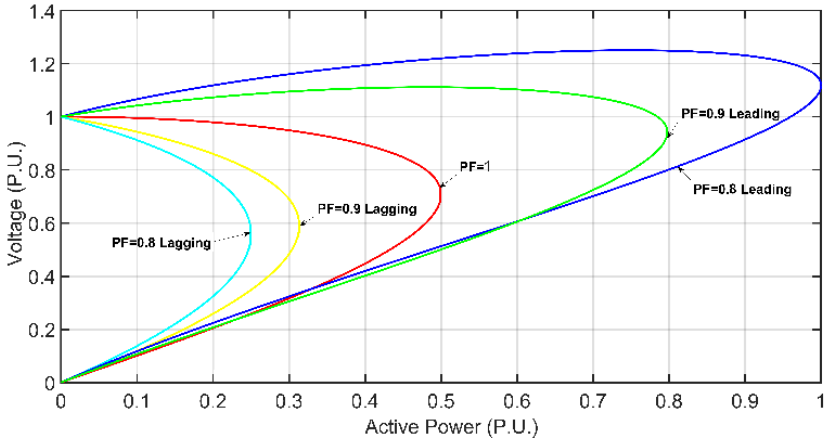

Fig. 2-2. P-V curve for PF=0.9 and PF=0.8 leading and lagging and PF=1

In the event of small disturbances, deviation of the load PF in post-disturbance and pre-disturbance may be negligible. As a result, post- and pre-disturbance P-V characteristics do not differ, noticeably. In this case, LSA can be calculated by analyzing proposed PF-based P-V curve in the pre-disturbance. However, in the case of large disturbance, load PF in post-disturbance may deviate from its value in pre-disturbance operating condition. However, P-V curve in the post-disturbance regime may be different from the one in pre-disturbance due to PF deviation. As a result, a precise procedure is required to derive accurate P-V curve.

2.4. **Proposed Load Shedding Technique**

This study proposes an under voltage-frequency load shedding (UVFLS) approach to restore voltage and frequency of the islanded IBMG. In the first step, active power demand sensitivity to bus voltage is obtained considering loads PF. The result is utilized for UVLS. In parallel manner, droop characteristics are employed for calculating UFLS. In the final step, the load shedding amount (LSA) will be calculated.

2.4.1. Deriving P-V Curve considering Loads PF

PLN curve algorithm was proposed in [25] in order to obtain P-V curve faster than conventional methods. In this algorithm adjustable value of the voltage difference between every two consecutive linear sections ($\Delta V_{desired}$) is used. The deficiency of PLN method is that, it is based on the resistive loads and hence without taking into account the reactive power and PF. Moreover, the PLN based approach employs PLN curve of the pre-disturbance condition. Since a large disturbance may change P-V characteristic of a bus, the PLN method may overshed the loads.

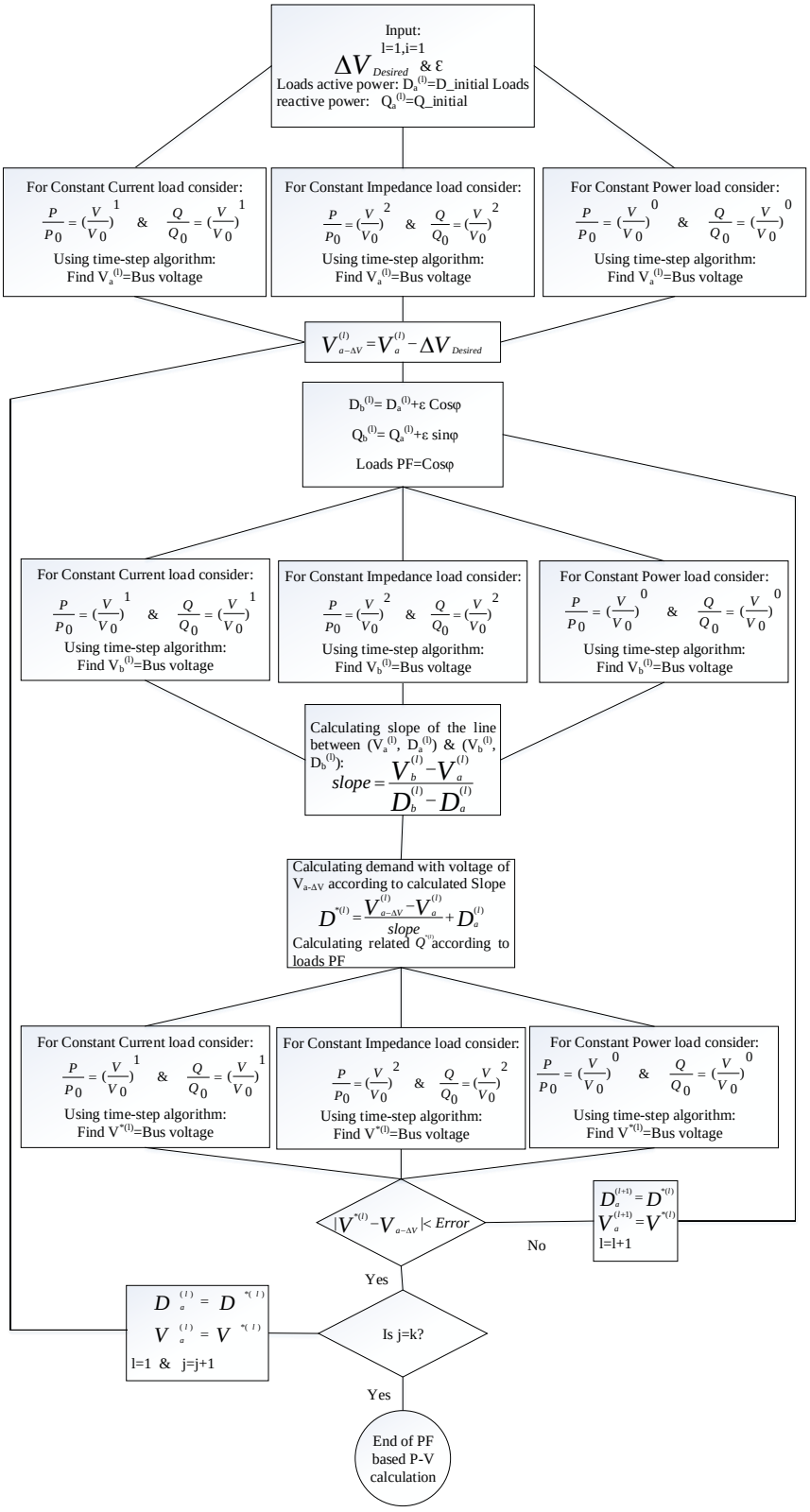

Fig. 2-3. Flowchart of the proposed algorithm for deriving PF based P-V curve

In the proposed algorithm of this chapter, loads PF are considered and P-V curves are obtained with regard to reactive and active demands. The flowchart of the proposed algorithm including various load models is demonstrated in Fig. 2-3 which is the updated version of the one presented in [25]. In the proposed algorithm, the load active demand (which is written as D) and reactive powers (which is written as Q change in each step according to the load PF. The voltage of each bus is computed by solving the time-domain equations incorporating the load model at that bus. If the number of operating DG units and loads or their power demands are changed, the computed bus voltages will be updated, accordingly. Furthermore, the PF-based P-V curves are updated based on the new configuration of the MG, if required.

As shown in Fig. 2-3, the PF-based P-V curve is obtained based on the load types. Due to relatively small size of MG compared with a bulk power system, it is possible to have a good estimation of the load types in the MG based on analyzing loads current, voltage and power behavior. According to the load's behavior, one of the mentioned models on Fig. 2-3 which is closer to loads behavior will be selected for load shedding. By using an estimation of load types and the proposed method in Fig. 2-3, a more optimal amount of load shedding can be obtained compared to using one load model for whole system loading.

2.4.2. Simple LS method

After employing proposed algorithm for deriving PF-based P-V curve in the IBMG system, the LSA can be calculated by computing under-voltage load shedding amount (UVA), under-frequency load shedding amount (UFA), and over-frequency load shedding

amount (OFA) [25]. Assuming i is the number of buses with the voltages under the desirable voltage level, UVA for bus number i can be calculated as follows:

$$UVA_i = S_i.\cos\varphi_i - S_d.\cos\varphi_d \qquad (2\text{-}18)$$

where S_i is the load power at bus i which can be derived from load voltage using PF-based P-V curve. $\cos\varphi_i$ is PF of the load after disturbance. In (2-7), S_d is the load power which can be obtained on the basis of lower limit of acceptable voltage and PF-based P-V curve. The value of $\cos\varphi_d$ is PF of the bus i. It should be noted that PF-based P-V curve has been obtained before disturbance. Hence, $\cos\varphi_d$ is the PF of the bus before disturbance. The complete UVA matrix can be written as follows which consists of UVA of all i buses experiencing voltage decline:

$$\left[UVA_{total}\right]_{i\times 1} = [UVA_1 \quad \cdots \quad UVA_i]^T \qquad (2\text{-}19)$$

UVA_{total} consists of i values, which one of them should be selected as LSA. By accomplishing the following steps, one element of UVA_{total} will be selected.

The voltage variation followed by UVA (ΔV_{UVA}) can be calculated by subtracting the load voltage and the lower and upper limits of the desired voltage. For a system with b buses, by calculating UVA_i for one of i buses and implementing LS according to measured UVA_i, voltage variation ΔV_{UVA} of all b buses can be calculated. By employing this approach to all i buses, an $i\times b$ matrix of voltage variations will be obtained. Then, by selecting maximum voltage variation in each row an $i\times 1$ matrix ($\Delta V_{UVAtotal}$) will be

derived. Finally, in $\Delta V_{UVAtotal}$ one row will be selected as LSA if its corresponding row in the obtained $i \times 1$ matrix has the minimum value of whole matrix.

UFA is designed to restore the frequency of the system to lower limit of acceptable frequency. UFA can be obtained by using droop coefficient which is used as power control of the system [27].

By subtracting system frequency from lower limit of acceptable frequency and dividing it by droop coefficient, UFA will be obtained:

$$UFA_i = \frac{f_l - f_i}{m} \tag{2-20}$$

where m is droop coefficient and f_l is the lower limit of the desired frequency. UFA_i is the power change needed to restore frequency from fi to f_l.

On the other hand, OFA is designed to prevent the occurrence of over frequency after LS, which is the maximum LSA to recover frequency of the system to upper limit of acceptable frequency. OFA can be derived by subtracting higher limit of acceptable frequency of the system from the system frequency and dividing it by droop coefficient.

$$OFA_i = \frac{f_u - f_i}{m} \tag{2-21}$$

Where m is droop coefficient and f_u is the upper limit of the desired frequency. OFA_i is the power change needed to restore frequency from fi to f_u. The voltage variations followed by OFA and UFA are ΔV_{OFA} and ΔV_{UFA}, respectively. Finally, LSA will be equal to one of UVA, UFA, or OFA values which can be obtained by analyzing one of the nine modes *M1-M9* presented in Table 2-1. Table 2-1 lists the possible modes of calculating LSA using the PF-based P-V curve [25].

Table 2-1
Obtaining LSA while using one PF-based P-V curve [26]

MODES	MIN(ΔV_{OFA},ΔV_{UFA},ΔV_{UVA})	1ST CONDITION	2ND CONDITION	LSA
M1	$\Delta V_{OFA}=\Delta V_{UFA}=\Delta V_{UVA}$	-	-	UFA
M2	$\Delta V_{UFA}=\Delta V_{UVA}$	-	-	UFA
M3	$\Delta V_{OFA}=\Delta V_{UVA}$	-	-	Min(UVA,OFA)
M4	$\Delta V_{OFA}=\Delta V_{UFA}$	-	-	UFA
M5	ΔV_{UFA}	-	-	UFA
M6	ΔV_{OFA}	-	-	OFA
M7	ΔV_{UVA}	UFA<UVA<OFA	-	UVA
M8	ΔV_{UVA}	OFA<UVA or UVA<UFA	$\Delta V_{UFA}<\Delta V_{OFA}$	UFA
M9	ΔV_{UVA}	OFA<UVA or UVA<UFA	$\Delta V_{UFA}>\Delta V_{OFA}$	OFA

2.4.3. Complete LS method

The process of updating PF-based P-V curve is treated as an online approach since each P-V calculation can take a few minutes, depending on the size of the IBMG system. According to [25], if a contingency occurs between two P-V curve updates, the only solution is to exploit the P-V curve computed at the last online computation period. In [25], the PLN curves are assumed unchanged during post-disturbance and pre-disturbance periods since IBMG loads were resistive. Based on such an unrealistic assumption, LSA calculation using Table 2-1 is acceptable as PF equals one and the post-disturbance P-V curve does not deviate from its pre-disturbance curve. However, depending on the load type, the load active and reactive powers and therefore its PF can change following the

disturbance. As a result, the corresponding PLN curve may change, and the approach of PLN may lose credibility.

In the proposed complete LS algorithm, for a load with the specified PF of PF_L, several PF-based P-V curves with pre-specified margins such as $\pm 0.1+PF_L$ and $\pm 0.2+PF_L$, are derived. By changing the number or margin of the curves, there is a trade-off between online processing time and the LS accuracy. The more the number of curves with pre-specified margins, the higher the accuracy of the method; but the online processing time will be increased. However, margins of pre-specified curves have reverse relation with the method accuracy. The process of updating PF-based P-V curve is done following each system change and the algorithm is then ready to act on subsequent disturbances. Thus, the LS operating time is not affected by the curve updates and the algorithm can quickly respond in real-time operation based on simple calculations using lookup table interpolations.

After the disturbance occurrence, the PF of the bus will be calculated. Afterwards, by analyzing the pre-calculated PF-based P-V curves, two closest curves which have the nearest upper and lower PF will be selected for obtaining the initial LSA value based on the 9 modes of Table 2-1. Thereafter, according to the mode of each condition, one of the 81 probable arrays in Table 2-2 will be selected for LSA. It should be noted that in Table 2-2, for instance M_{31} and M_{42} demonstrate mode M_3 in the 1^{st} and mode M_4 in the 2^{nd} selected PF-based P-V, respectively. By assuming that M_{31} and M_{42} are the modes obtained from the nearest PF-based PV curves to loads PF, the final LSA value will be M_{31} according to Table 2-2. By increasing the number of PF-based P-V curves, the method efficiency will be increased. However, generating higher number of PF-based P-V curves

will be more time consuming. Nevertheless, such P-V curves can be obtained offline and stored as lookup tables for fast calculation in real-time.

It should be mentioned that due to the massive number of probable contingencies, only credible disturbances such as DG unit outage and/or load drastic changes are considered based on techniques and historical records such as contingency ranking or bounding [29]. These disturbances can cause voltage and frequency instability as well as load PF deviation. It should be noted that distribution of the LSA among buses is proportional to their loading at pre-disturbance condition.

Table 2-2
Obtaining LSA using two upper and lower P-V curves

1st P-V	2nd P-V	LSA
$\begin{bmatrix} M_{11} \\ M_{21} \\ M_{31} \\ M_{41} \\ M_{51} \\ M_{61} \\ M_{61} \\ M_{71} \\ M_{81} \\ M_{91} \end{bmatrix} \times$	$\begin{bmatrix} M_{12} \\ M_{22} \\ M_{32} \\ M_{42} \\ M_{52} \\ M_{62} \\ M_{62} \\ M_{72} \\ M_{82} \\ M_{92} \end{bmatrix}^T$	$\begin{bmatrix} \min(M_{11},M_{12}) & \min(M_{11},M_{22}) & M_{32} & \min(M_{11},M_{42}) & \min(M_{11},M_{52}) & M_{11} & M_{72} & \min(M_{11},M_{82}) & M_{92} \\ \min(M_{21},M_{12}) & \min(M_{21},M_{22}) & M_{32} & \min(M_{21},M_{42}) & \min(M_{21},M_{52}) & M_{21} & M_{72} & \min(M_{21},M_{82}) & M_{92} \\ M_{31} & M_{31} & \min(M_{31},M_{32}) & M_{31} & M_{31} & M_{31} & M_{31} & M_{31} & M_{31} \\ \min(M_{41},M_{12}) & \min(M_{41},M_{22}) & M_{32} & \min(M_{41},M_{42}) & \min(M_{41},M_{52}) & M_{41} & M_{72} & \min(M_{41},M_{82}) & M_{92} \\ \min(M_{51},M_{12}) & \min(M_{51},M_{22}) & M_{32} & \min(M_{51},M_{42}) & \min(M_{51},M_{52}) & M_{51} & M_{72} & \min(M_{51},M_{82}) & M_{92} \\ M_{12} & M_{22} & M_{32} & M_{42} & M_{52} & \min(M_{61},M_{62}) & M_{72} & M_{82} & \min(M_{61},M_{92}) \\ M_{71} & M_{71} & M_{32} & M_{71} & M_{71} & M_{71} & \min(M_{71},M_{72}) & \min(M_{71},M_{82}) & \min(M_{71},M_{92}) \\ \min(M_{81},M_{12}) & \min(M_{81},M_{22}) & M_{32} & \min(M_{81},M_{42}) & \min(M_{81},M_{52}) & M_{81} & \min(M_{81},M_{72}) & \min(M_{81},M_{82}) & \min(M_{81},M_{92}) \\ M_{91} & M_{91} & M_{32} & M_{91} & M_{91} & \min(M_{91},M_{62}) & \min(M_{91},M_{72}) & \min(M_{91},M_{82}) & \min(M_{91},M_{92}) \end{bmatrix}$

2.5. **Simulation and Case Studies**

In grid-connected operation mode of MG, the main utility maintains the voltage and frequency of the system. In this case, any active and reactive power imbalance can be compensated by the main grid. On the contrary, in islanded mode of MG, DGs have to maintain the voltage and frequency of the islanded MG into their acceptable values. Since the MG is more vulnerable to disturbances under islanded mode of operation, this

operating mode is the focus of this study to determine the required LSA, under the most severe scenarios.

It does not matter if the MG operates in islanded mode or grid-connected mode, the sensitivity of bus voltage to the load power can be obtained based on the proposed PF-based P-V curve. In both grid-connected or islanded operation, following an event and dropping the bus voltages or frequency below the acceptable values, if generation has enough power to compensate the power imbalance, the bus voltages and frequency recover to the acceptable values and LS is not needed. However, if generation is not sufficient to compensate the power imbalance, voltage and frequency will not recover and the proposed LS will be applied to the system as the last firewall to maintain the voltage and frequency within the acceptable ranges.

In the proposed LS method, the voltage sensitivity to the system loading in each bus should be calculated, based on the load power factor. Such a sensitivity is demonstrated by the PF-based P-V curves. If any changes occur in the system parameters, topology, generation, or loads, the state-space equations will be updated and, as a result, the responses of the system and the PF-based P-V characteristics will be updated, accordingly. Based on the updated PF-based P-V curves and the frequency of the system, new LSA will be calculated.

The test islanded IBMG, shown in Fig. 2-4, consisting of two DGs and three various load models is used to demonstrate the application of the proposed LS method. The parameters of the DG unit controllers are listed in Table 2-3. The complete state-space equations of the islanded IBMG is implemented in MATLAB software. The acceptable range of frequency and voltage are considered to be (49.9-50.1) Hz and (0.95-1.05) pu,

respectively [30]. In addition, $\Delta V_{desired}$ in the process of obtaining PF-based P-V curve is chosen to be 0.001 pu.

Three different probable contingencies are considered as three scenarios. Three LS methods are exerted and compared to each other in each of three scenarios.

Table 2-3
Control Parameters of DGs [27]

Parameter		*Value*	*Parameter*		*Value*
Switching frequency (f_s)		8 kHz	Droop	m_p	10e-5
Filter	L_f	1.35 mH		n_q	10e-5
	C_f	50 µF	Voltage Controller	K_{pv}	0.05
	r_f	0.1Ω		K_{iv}	390
	L_c	0.35 mH		F	0.75
	r_{Lc}	0.01 Ω	Current Controller	K_{pc}	10.5
Grid frequency (f)		50 Hz		K_{ic}	16e3
Power controller	ω_c	10π rad/s	Inverter rated power		10 KVA

- **PLN based LS:** In the first LS technique, LSA is calculated using the PLN curves proposed in [25]. In this method P-V curves are derived by using only active power of the loads. It should be noted that by utilizing this LS approach, the reactive power will not be considered in the algorithm in Fig. 2-3. In addition, the P-V curve using for LS is obtained before the contingency.

- **Simple LS technique:** This method is based on the technique proposed in Fig. 2-3. In this analysis, PF-based P-V curves are derived considering load active power and PF. In this method, PF-based P-V curve is obtained from pre-disturbance condition for each bus and Table 2-1 is employed for calculating LSA.

- **Complete LS Technique:** The complete LS approach considers load power and PF

similar to simple LS method. In addition, several PF-based P-V curves with various PFs such as ±0.1+PF and ±0.2+PF are derived as well. At the end, both Table 2-1 and Table 2-2 are used for calculating accurate LSA.

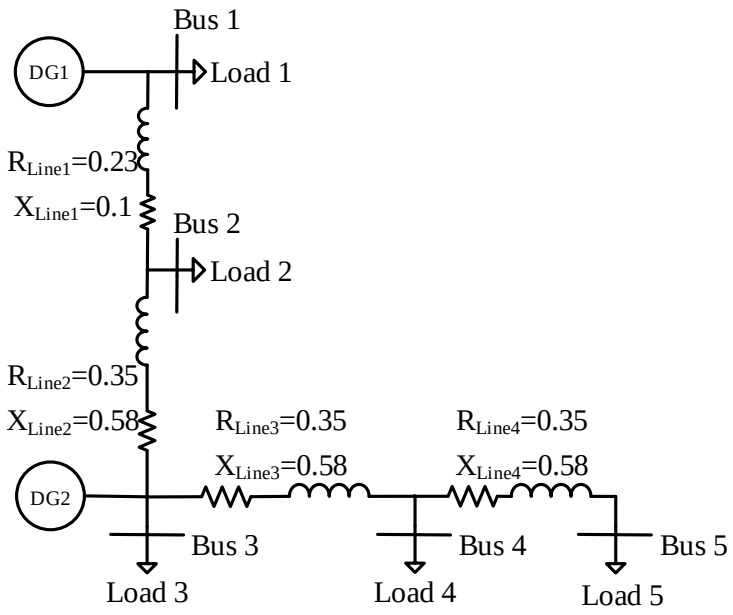

Fig. 2-4. Test Islanded IBMG for validation of proposed LS method

2.5.1. Scenario 1: Sudden Load Increment

In this case study, three CILs are increased drastically where CPL and CCL are constant. Both DGs are operating in islanded IBMG and equipped with droop control. The test islanded IBMG is demonstrated in Fig. 2-4. DG control structure is according to Fig. 2-1 and control parameter of voltage source inverters of the DGs are shown in Table 2-3. The pre-disturbance loading of the system is shown in Table 2-4. As it is demonstrated in Table 2-4, load increment occurs in such a way that the load PF deviates from its value in pre-disturbance condition. For instance, PF of load1 which was 0.8 lagging before the disturbance, is lagging 0.65 during the disturbance.

To recover voltage and frequency, three aforementioned LS methods are exerted and

analyzed separately. The value of LSA in each method is shown in scenario 1 of Table 2-8. It can be seen from Table 2-8 that voltage of buses 4 and 5 are below acceptable range. By using PLN method, LS is done considering active power of the loads. Hence, P-V curves are not classified according to the loads PF. Comparing the LSAs of Table 2-8 reveals that LSA obtained from PLN method is higher than the other ones which implies that an over LS occurs.

Simple LS technique considers load PF and derives PF-based P-V curve, according to the PF of the loads in the pre-disturbance conditions. As load PF in pre-disturbance is different from that of the post disturbance condition, the associated PF-based P-V curves are also different. Hence, LSA calculation is not accurate. Finally, by employing complete LS method, for each load having PF equal to PF_L in pre-disturbance, four PF-based P-V curves with $\pm 0.2+PF_L$ and $\pm 0.1+PF_L$ are derived. Then, two PF-based P-V curves which have the nearest PF value to the loads PF during disturbance are selected and LSA and mode for each are calculated using Table 2-1. Afterwards, Table 2-2 is employed in such a way that respective mode for the first selected mode is inserted in the vertical matrix and respective mode for the second selected mode is inserted in the horizontal matrix and other elements in both matrices are assumed zero. The non-zero value in the square matrix is the LSA. As it can be seen from Table 2-8, LSA using the third method is 6.4% lower than the PLN-based LS methods. In addition, voltage restoration of buses 5 and 6 are achieved better than the previous methods.

Fig. 2-5 demonstrates the difference between P-V curves obtaining from PLN method and the proposed PF-based PV curves. The dashed lines in Fig. 2-5 show PF-based P-V curves while the solid lines are P-V curves obtained from the PLN method considering

only the load active power. The PF of the curves associated with these PF-based P-V curves are 0.7, 0.6, 0.85, 0.7 and 0.8 for buses 1 to 5 respectively which are closest PF-based P-V curves to loads PF during disturbance. In this scenario, the voltages of buses 4 and 5 are more susceptible to load increment as illustrated in Fig. 2-5. Fig. 2-5 clearly demonstrates the effectiveness of using PF-based P-V curve in respect to PLN method. By considering a specific voltage drop at bus 5 for instance, the corresponding power change is lower with PF-based P-V curve than that of the non-PF-based P-V curves. This fact is illustrated in scenario 1 of Table 2-8 where LSA of PLN based LS method is higher than other LS methods.

In addition, as shown in Fig. 2-5, the weakest bus (lowest stability index) is found at bus 5. This is expected, since this is the farthest bus from the DGs. In Fig. 2-5, P-V curve of bus 5 has the highest slope which explains low voltage regulation of this bus. If, for instance, bus 4 becomes heavily loaded due to a specific spatial distribution, this bus may become the weakest bus of the system and in this condition the slope of P-V curve of bus 4 will increase which means this bus will become more vulnerable to disturbances.

Finally, the optimum LSA will be calculated according to the updated PF-based P-V curve to maintain each bus voltage to their lower acceptable limit in order to enhance the stability of the system.

Table 2-4
Loading of the System in Scenario 1

Condition		Load1 CPL	Load2 CIL	Load3 CCL	Load4 CIL	Load5 CIL
Pre-disturbance	P	3.39 kW	2.56 kW	8.1 kW	2.07 kW	3.35 kW
	PF	0.8	0.7	0.85	0.9	0.75
During disturbance	P	3.39 kW	5.12 kW	8.1 kW	4.14 kW	6.7 kW
	PF	0.65	0.61	0.85	0.72	0.78

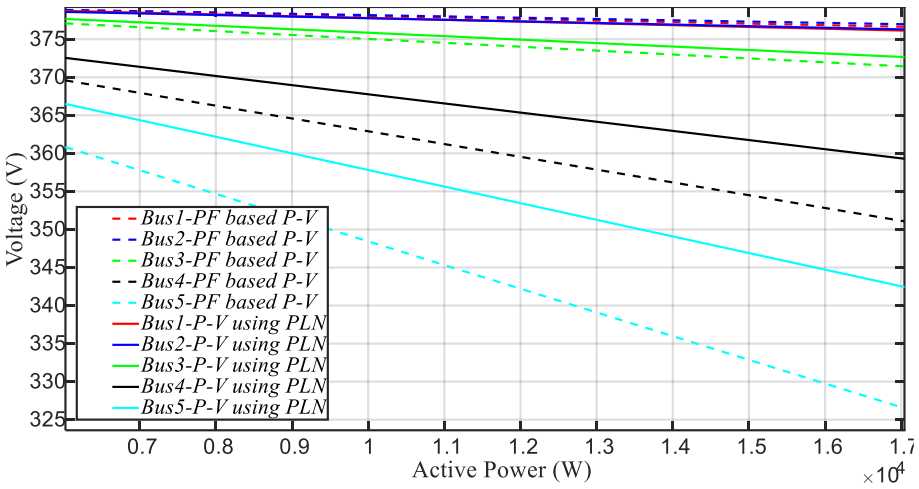

Fig. 2-5. Comparing PF-based P-V curve and P-V curve using PLN

2.5.2. Scenario 2: DG2 trip

In this analysis, it is assumed that DG2 in Fig. 2-4 is tripped and loads PFs are changed. Loads characteristics before and after the DG trip are shown in Table 2-5.

Due to DG2 trip, voltages of buses 3, 4 and 5 are not in the desirable range and frequency drops to 49.8692 Hz. According to Table 2-8, by utilizing the complete LS technique, not only the voltage recovery is carried out effectively, but also LSA is respectively 36.68% and 30.2% lower than the first and second LS methods.

Table 2-5
Loading of the System in Scenario 2

Condition		Load1 CPL	Load2 CIL	Load3 CCL	Load4 CIL	Load5 CIL
Pre-disturbance	P	0.4 kW	1 kW	2.6 kW	1.2 kW	1.5 kW
	PF	0.8	0.7	0.85	0.9	0.75
During disturbance	P	0.4 kW	1 kW	2.6 kW	1.2 kW	1.5 kW
	PF	0.65	0.61	0.85	0.72	0.78

2.5.3. Scenario 3: DG2 trip and loads increase

In this case study, it has been assumed that both CIL increment and DG2 trip occur

simultaneously. Table 2-6 presents the simulated loads active power and PF in pre-disturbance and during disturbance conditions. The effect of various LS methods on voltages and frequency are demonstrated in scenario 3 of Table 2-8. Simultaneous load increment and DG trip causes the voltage magnitudes of buses 3, 4, and 5 to decrease below 0.95 pu. In this case, PLN based LS method restores voltage of susceptible buses more effectively than the other methods. However, higher voltage magnitude in post-disturbance is due to over LS. On the other hand, simple LS and complete LS methods both restore voltage to the acceptable value and the value of LS in these two methods are respectively 37.28% and 32.6% lower than the first method. Complete LS technique however results in the lowest LSA since it employs the PF-based PV data during the disturbance.

Table 2-6
Loading of the System in Scenario 3

Condition		Load1 CPL	Load2 CIL	Load3 CCL	Load4 CIL	Load5 CIL
Pre-disturbance	P	0.8 kW	0.5 kW	2.5 kW	0.5 kW	0.6 kW
	PF	0.8	0.7	0.85	0.9	0.75
During disturbance	P	0.8 kW	1 kW	2.5 kW	1 kW	1.2 kW
	PF	0.65	0.61	0.85	0.72	0.78

Table 2-7
Loading of the System in Scenario 4

Condition		Load1 CPL	Load2 CIL	Load3 CCL	Load4 CIL	Load5 CIL
Pre-disturbance	P	0.6 kW	0.5 kW	2.1 kW	0.5 kW	0.6 kW
	PF	0.8	0.7	0.85	0.9	0.75
During disturbance	P	0.6 kW	1 kW	2.1 kW	2 kW	1.5 kW
	PF	0.7	0.61	0.7	0.92	0.78

Table 2-8
Comparison of LSAs obtained from three LS methods in Scenario 1, 2 and 3 – lowest bus voltages are shown

Scenarios	Indices	During disturbance	Method 1: Pervious LS method	Method 2: Simple LS Technique	Method 3: Complete LS Technique
			Post- disturbance	Post- disturbance	Post- disturbance
Scenario1	V_{B4}(pu)	0.9074	0.9743	0.9728	0.9757
	V_{B5}(pu)	0.8621	0.9658	0.9655	0.9665
	f (Hz)	49.7268	49.7765	49.8994	49.9155
	LSA(VA)		**1.4547e+04**	**1.3715e+04**	**1.3615e+04**
Scenario2	V_{B3}(pu)	0.9433	0.9638	0.9608	0.9618
	V_{B4}(pu)	0.9251	0.9562	0.9539	0.9524
	V_{B5}(pu)	0.9159	0.9519	0.9500	0.9500
	f(Hz)	49.8692	49.9157	49.9083	49.9072
	LSA(VA)		**1.5592e+03**	**1.4144e+03**	**987.1745**
Scenario3	V_{B3}(pu)	0.9444	0.9735	0.9643	0.9641
	V_{B4}(pu)	0.9265	0.9687	0.9573	0.9569
	V_{B5}(pu)	0.9169	0.9660	0.9534	0.9529
	f(Hz)	49.8711	49.9089	49.9106	49.9105
	LSA(VA)		**2.0479e+03**	**1.3801e+03**	**1.2843e+03**

2.5.4. Scenario 4: Line 4 resistance increase

In this case study, the effect of line resistance, as a sample of a possible change in the topology of the system and representing a distribution feeder is investigated. It has been assumed that the demand is increased and DG2 is tripped. To recover voltage and frequency of the system, complete LS technique is implemented.

In the first condition of this scenario, it is assumed that X/R ratio of line 4 is 1.67. In the second condition, X/R ratio of line 4 is assumed 0.89. In addition, loads PF are assumed to be different from previous scenarios. System loading in this scenario is demonstrated in Table 2-7. PF-based P-V curves for both conditions have been obtained according to the proposed algorithm in Fig. 2-3 and the curves are shown in Fig. 2-6. Fig. 2-6 shows that the slope of PF-based P-V curve related to bus 5 is higher for X/R ratio of 0.89. For a specific voltage drop, the higher the slope of P-V curve, the lower amount load shedding is needed to maintain the voltage to its acceptable range. Therefore, the amount of load

curtailment should be lower for the second condition compared to the first condition due to the higher slope to P-V curve in the second condition.

It should be noted that, since the characteristics of lines 1-3 are similar and loads PF during both conditions are unchanged, only PF-based P-V curve of bus 5 is changed and other PF-based P-V curves related to buses 1-4 do not have any deviation in the second condition compared with the first condition.

Table 2-9 demonstrates the effect of the feeder resistance resulting in the different X/R ratios on the LSA. Table 2-9 shows that the LSA in condition-2 is lower than that of condition-1. This is due to the fact that the obtained PF-based P-V curve of bus 5 has higher slope and thus, lower load shedding is needed to enhance the voltage magnitude to its acceptable value.

This scenario has been implemented in order to illustrate the applicability of the proposed LS technique to various systems with different feeder characteristics. The change in line resistance results in the change in PF-based P-V curve of the bus connected to the line. Therefore, PF-based P-V curve of prospective line has been updated and the LSA is changed accordingly. Any change in the topology of the system also results in the updated PF-based P-V curves and LSA will be calculated without dependency to the topology of the system.

Hence, no matter what the topology of the system is, responses of the system can be obtained and LSA will be updated.

Table 2-9
Comparison of LSA of two conditions with different line resistances

Indices	During disturbance	Complete load shedding technique	
		Condition 1: X/R of line 4=1.67	Condition 2: X/R of line 4=0.98

		Post- disturbance	Post- disturbance
V_{B3}(pu)	0.9422	0.9745	0.9745
V_{B4}(pu)	0.9237	0.9738	0.9732
V_{B5}(pu)	0.9141	0.9734	0.9712
f (Hz)	49.872	49.9432	49.9410
LSA (VA)		**1.3761e+03**	**1.1416e+03**

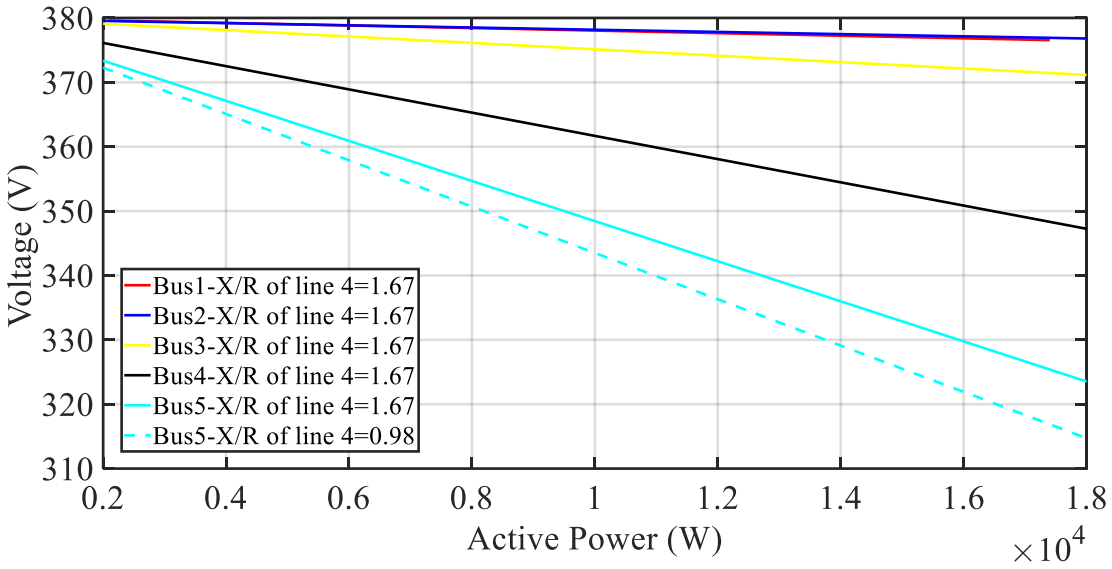

Fig. 2-6. Comparing PF-based P-V curves for different X/R ratios of line 4

2.5.5. Stability Analysis during disturbance and after load shedding

As it is mentioned in [26], by using the mentioned linear state-space model as well as the time-step simulation, eigenvalues and responses of the system can be obtained at each time. In addition, both small-signal stability and large-signal stability analysis is applicable [26].

The eigenvalues of the system in each iteration can be obtained using *A(t)* in (1). Fig. 2-7 shows the effective eigenvalues of the system during contingency and after the load shedding. Fig. 2-8 illustrates that after the contingency, the bus voltages have been

reduced and the system became closer to instability. In Fig. 2-7, effective system eigenvalues, 20 iterations after the occurrence of disturbance in scenario 3, are demonstrated in blue and effective system eigenvalues after load shedding and in steady-state are shown in red. The move of the eigenvalues to the left in Fig. 2-7 shows that the system stability is enhanced after the load shedding.

2.5.6. Transient response in pre-disturbance, during disturbance and post-disturbance

As mentioned before, this study is based on the time- domain simulation approach which provides transient response of the system before, during and after disturbance. Transient responses of the system in scenario 3 while using complete LS calculation are demonstrated in Figs. 2-8 to 13. At 0.45 *s*, DG2 trips and loads 2, 4 and 5 increase. Active power demand increment is demonstrated in Fig. 2-12 and reactive power is changed according to loads PF. Due to the increase in loads power and the DG2 trip, power of DG1 increases drastically to meet the demand, as shown in Fig. 2-11. Due to this disturbance, frequency is decreased and reaches below acceptable value in less than 20 *ms* as shown in Fig. 2-10. Based on Fig. 2-8, the voltage of buses 3 to 5 are lower than the acceptable value. In order to validate the proposed state-space-based method, scenario 3 has been simulated in PSCAD program and the results are shown in Fig. 2-9 which illustrate a close agreement with the bus voltages of Fig. 2-8. The PSCAD results are based on the 10 µsec time-step, whereas with the time step of 100 µsec (10 times larger), the proposed method results are almost identical to those of the PSCAD. Based on several case studies and the high-frequency transients which can be observed in IBMG [27], the time-step of 100 µsec

can ensure enough accuracy of the proposed simulation in order to have a fast response. However, there is a trade-off between time-step size and accuracy of transient response.

In pre-disturbance condition (before t=0.45 s), various PF-based P-V curves were derived for all the system buses. After t=0.45 s, two closest PF-based P-V curves to loads PF are selected and according to Table 2-1, two of 9 modes are selected and inserted in Table 2-2. Finally, LSA is calculated through one of 81 modes in Table 2-2. LSA is sent to relays and LS is accomplished after approximately 70 *ms*. As a result, the generated power of DG1 decreases as demonstrated in Fig. 2-11. Based on Figs. 2-8 and 2-10, the bus voltages and system frequency are restored to the desirable levels. Since load 1 is constant power load, its active power does not deviate from its value in this scenario (Fig. 2-12). In addition, as shown in Fig. 2-13, the current of load 3 is not changed during the events in this scenario as it is a CCL.

The main goal of the proposed load shedding method is to find the voltage sensitivity to the system loading at each bus according to the loads power factor. This sensitivity has been demonstrated by PF-based P-V curves. The sensitivity of bus voltages to the loads power can be obtained regardless of the impedances of the loads, location of DGs/loads, or size or type of the DGs. Based on the obtained PF-based P-V curves and the frequency of the system, load shedding amount can be obtained according to the technique proposed in the chapter. The proposed UVFLS method takes into account the system topology and it is based on simple mathematical computations such as maximum and minimum operators, algebraic operations and rule-based decision makings.

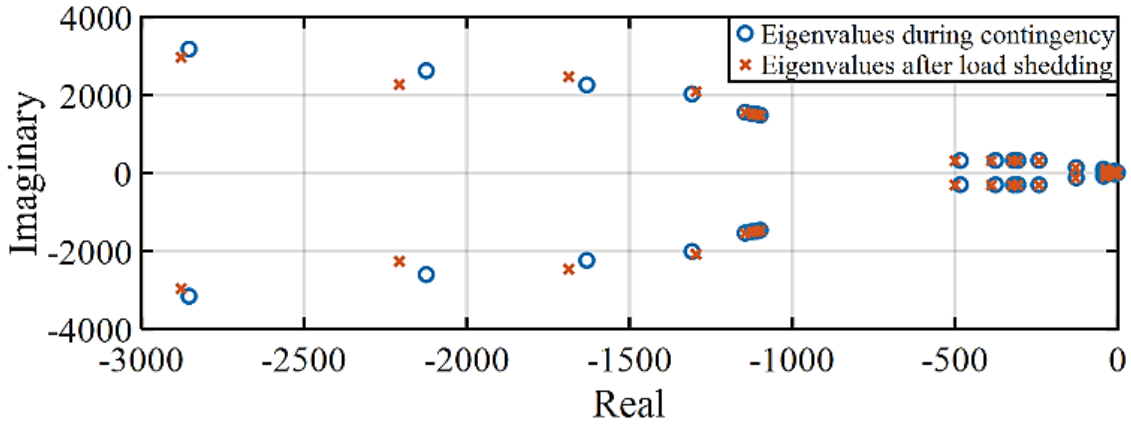

Fig. 2-7. Effective eigenvalues of the system in scenario 3 during the contingency and after the load shedding

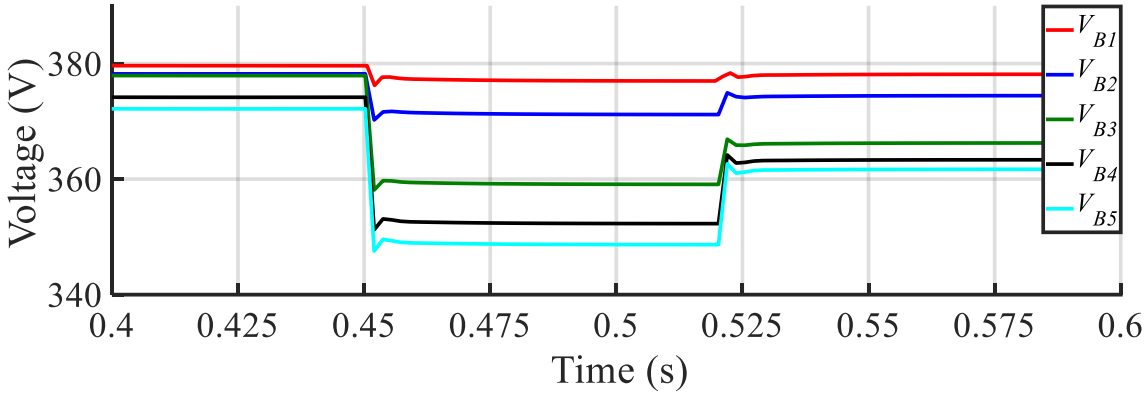

Fig. 2-8. Bus Voltages in pre-disturbance, during disturbance and post-disturbance in scenario 3 – obtained from the proposed state-space-based load shedding technique

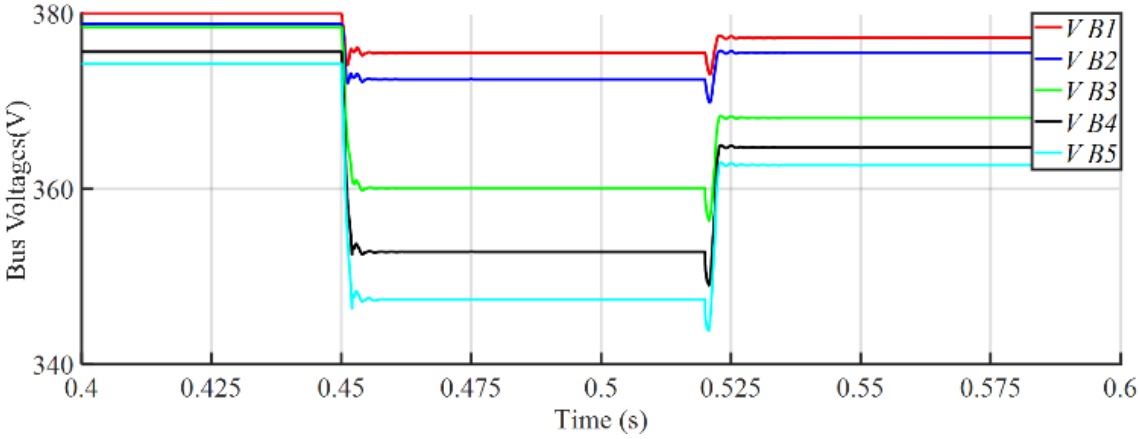

Fig. 2-9. Validation of bus voltages of Fig. 8 based on PSCAD time-domain simulation

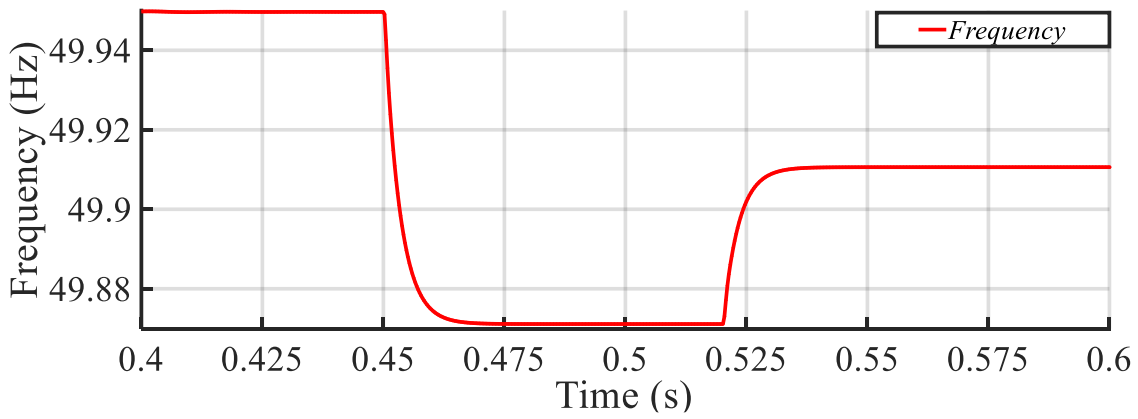

Fig. 2-10. Frequency in pre-disturbance, during disturbance and post-disturbance in scenario 3

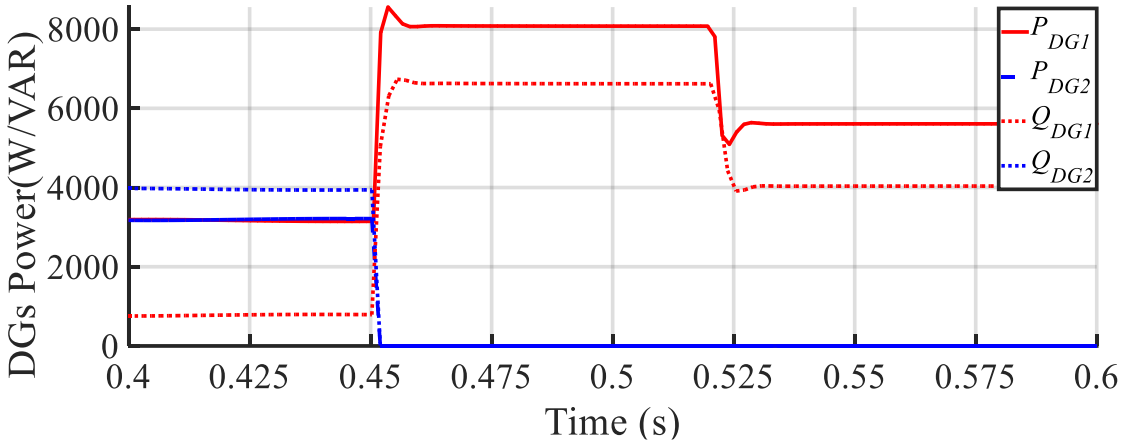

Fig. 2-11. DGs active and reactive power in pre-disturbance, during disturbance and post-disturbance in scenario 3

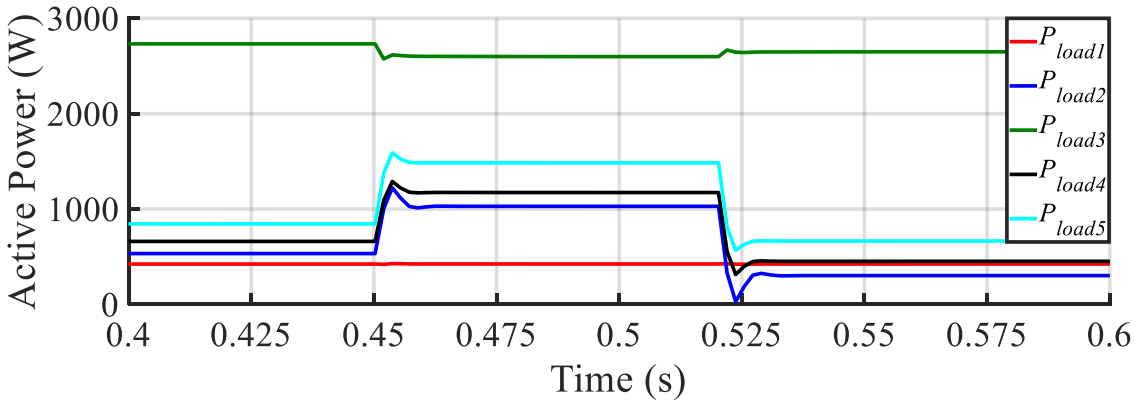

Fig. 2-12. Loads active power in pre-disturbance, during disturbance and post-disturbance in scenario 3

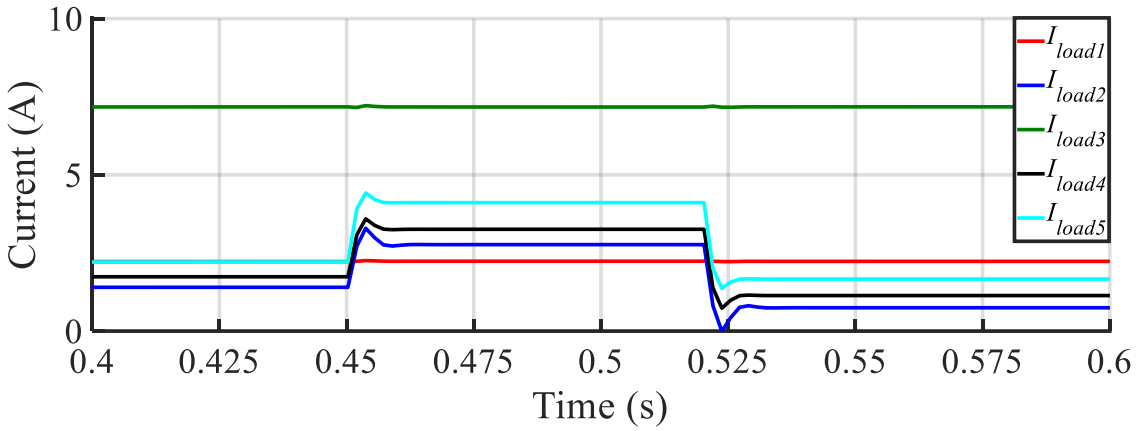

Fig. 2-13. Load currents in pre-disturbance, during disturbance and post-disturbance in scenario 3

2.6. **Conclusion**

Two LS techniques are developed in this chapter, i) a simplified LS technique and ii) a complete LS approach. The simplified LS technique derives PF-based P-V curve according to the type, size, and the power factor of the system loads. The obtained PF-based P-V curve as well as the frequency droop characteristics are used for calculating UVA, UFA and OFA and the final LSA is calculated. The complete LS method is based on the sensitivity factors in pre-disturbance conditions to address the dynamic changes of the operating conditions subsequent to the disturbance. In this method, several PF-based P-V curves are generated and two PF-based P-V curves, i.e., upper and lower PF values, closest to the loads are selected and the LSA is calculated. The proposed LS approach is validated based on the detailed time-domain simulation in the PSCAD program.

A comparative analysis in an islanded IBMG based on the time-domain simulation and taking into account various load models is carried out to investigate the effectiveness of the proposed LS techniques, compared with a previously published method. Three

possible disturbances are considered in the study system. The study reveals that the proposed complete LS technique results in the lowest LSA and effectively restores both voltage and frequency of the islanded IBMG to the desired values. Moreover, the system stability during disturbance and after the LS has been investigated which demonstrates the effectiveness of UVFLS in enhancing the system stability.

Chapter 3 : A DATA-DRIVEN MULTI-VARIABLE REGRESSION ANALYSIS FOR REAL-TIME FAULT-INDUCED DELAYED VOLTAGE RECOVERY PREDICTION

3.1. Introduction

3.1.1. Background and Motivation

In the modern power system, the increase in demand and power transfer between utilities raise the voltage stability concerns in electric power networks. Noticeable challenges such as natural complexity and dynamics of the power system as well as sporadic generation of renewable energy sources (RESs), transmission facilities in locations with a high concentration of industrial, commercial or residential induction motor loads and the environmental and political resistance to install or upgrade transmission lines are enforcing the power grid to rely on existing generation and transmission facilities. Aforementioned reasons impose power grids to operate closer to the operating limit, which ultimately results in voltage instability issues [31]. The ability of a power system to maintain acceptable bus voltages in the system under normal operating conditions, and after being subjected to a disturbance from a given initial operating state is called voltage stability [32]. The loss of load is a possible outcome of voltage instability where the reactive power imbalance, which in severe conditions may potentially lead to cascading outages. The driving force of voltage instability is generally loads and thus analyzing their

effect on power system is of great importance in voltage stability assessments [33].

A low voltages associated with system faults in transmission, sub-transmission or distribution feeders in a system with high penetration of induction motors may lead to significant depressed system voltage level for several seconds which is known as Fault Induced Delayed Voltage Recovery (FIDVR) [34]. FIDVR phenomena occurrs and is aggravated by the presence of large amounts of single-phase air conditioning (A/C) induction motor loads. This type of load plays main role in FIDVR events due to its low inertia and hence its predisposition to stall. These motors can stall for nearly all faults greater than five cycles as a result of fault voltages less than 60% to 70% nominal value [35]. Once a single-phase A/C compressor stalls, it will draw excessive current and reactive power from the grid that are in the range of four to six times its nominal steady-state value which further aggravating the initial fault voltage depression [34].

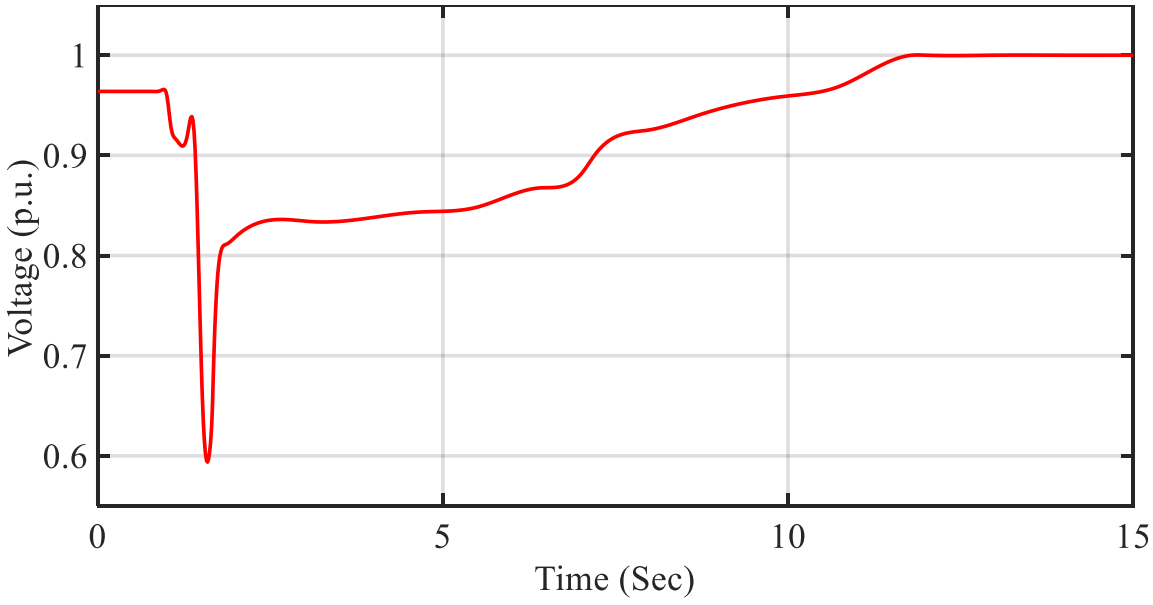

Fig. 3-1 A typical FIDVR event waveform

This situation makes it more difficult for the power system and its reactive power

resources to recover the voltages to acceptable levels in post-disturbance conditions [36]. As a result, the power system voltage remains at a significantly low level for a long time [34]. Eventually, the stalled motors will trip by thermal protection with an inverse time-overcurrent characteristic which can take from 3 to 30 seconds. A typical FIDVR event waveform is shown in Fig. 3-1. When FIDVR events affect the bulk power system, the situation becomes one of elevated risk, as it can trigger further losses of load and generation and even cascaded stalling of induction motors in the same or nearby feeders and potentially leading to system voltage collapse [34], [35].

The North American Electric Reliability Corporation (NERC) Transmission Issues Subcommittee defines FIDVR as a voltage condition initiated by a fault and characterized by stalling of induction motors where initial voltage recovery after the clearing of a fault is less than 90% of pre-contingency voltage and slow voltage recovery occur more than 2 seconds to expected post-contingency steady-state voltage level [34].

Tens of FIDVR cases were reported but not documented in available literature. Southern California Edison (SCE) experienced a great number of FIDVR cases since 1990. Some of the extreme FIDVR events happened due to the hot and humid weather conditions (which result in high increase of air-conditioning load) along with a large number of system faults [37]. In another case, the fault resulted in a voltage depression over hundreds of square miles, including the metro Atlanta, GA, area [38]. The initial fault-induced voltage sag was exacerbated by the presence of numerous large industrial and small residential induction motor loads. Note that the voltage recovery period was prolonged due to the reactive demand increase caused by slow down motor loads in response to a reduced supply voltage. Approximately 1900 MW of load was lost as a result of the

disturbance [39]. These observations demonstrate the significance of investigating and more importantly predicting FIDVR phenomenon. Aforementioned consequences of FIDVR could be mitigated if the duration of voltage sag could be predicted in the very first cycles after the fault.

3.1.2. Literature Review

Literature studying FIDVR can mainly be categorized as: 1) creation of time domain simulation models or dynamic stability tools to represent the motor load in a fashion that accurately replicates field-observed FIDVR phenomenon; However, the inherent uncertainty, complexity, and diversity of power system loads make dynamic load modeling very complicated and challenging and the user-defined parameters in the aggregate induction motor load model, makes these methods case sensitive [40]. 2) Proposing a method to prevent, and/or mitigate the impact of induction motor stalling and FIDVR.

To mitigate the impact of FIDVR and prevent cascading failure, NERC Transmission Issues Subcommittee [34], proposes two solutions, grid (Planning stage) level and unit (End user) level solutions. End user level solutions propose upgraded A/C units with equipment control devices to remove induction motor loads from the grid prior to stalling caused by under voltage conditions. This method is a long-term solution since A/C standards necessary to achieve this will not be enacted overnight. Grid level (Planning stage) solutions, on the other hand, can mainly be listed as two methods. One is reinforcing the power network with dynamic VAR compensation devices, such as SVC and STATCOM [41], [42], [43]. However, such devices remain very expensive, and their wide installation is usually limited by the investment budget and substation space. The second

solution at the grid-level planning stage is to deploy fast detection of FIDVR and emergency control e.g., load shedding (LS) [37], [44], [45] at post-fault control stage.

Among all aforementioned methods, deploying fast FIDVR detection and implementing appropriate LS is more applicable. Some of FIDVR detection methods are investigating the FIDVR via off-line time domain simulation, which can be accomplished for a pre-defined contingency set. These methods are usually computationally demanding and not suitable for real-time application; and it can only support event-based emergency control, which is less robust and accurate [46].

Real-time post-fault system characteristics measurement can be achieved by employing wide area measurement systems and the consecutive data-driven control methods can be implemented using obtained data [47]. The data-driven methods proposed in the literature for voltage stability assessment can be grouped into analytical methods and machine learning methods. Some of the analytical methods in the literature are a slope-based method proposed in [17], [44], and [45] which assessed the recovery time of motor rotation speed, voltage instability predictor [37], Lyapunov exponent [48] which these methods only work for voltage collapse detection. A fixed post-fault time window has been used in [17]-[45] to assess FIDVR. The deficiency of analytical methods is that they are not fast enough for the subsequent LS [46]. On the contrary, in machine learning (ML)-based techniques, a fast real-time decision making can be achieved by implementing an online machine training using the system database. Less data requirement, less computation time, smart decision making and ability to upgrade the model based on updated dataset are some noticeable merits of ML-base methods [49] compared to analytical methods. ML-based techniques have been utilized for various

power system stability challenges by implementing data-driven stability assessment methods [50]- [22]. Decision tree [50], ensemble learning [51], fuzzy decision tree [52]-[53], shapelet classification [54] and imbalanced learning [55] techniques have been used in the literature to investigate steady-state voltage stability assessment, real-time rotor-angle stability assessment and real-time voltage collapse detection, respectively. However, very few works focused on ML-based FIDVR assessment methods. Some recent works which proposed valuable ML-based FIDVR prediction methods worked on linear regression [46], ensemble learning [56] and extreme learning machine [57] techniques for predicting FIDVR duration. However, these methods are either not accurate enough or too computationally complicated and do not use several features for ML model development.

3.1.3. Contribution of This Chapter

This chapter at first focuses on investigating FIDVR duration dependency on the system characteristic based on a comprehensive sensitivity analysis. The most important features are used for ML training in the next step. A comprehensive data-driven measurement-based ML-based technique for real-time FIDVR prediction is proposed which employs several power system data as an input for online ML design. The proposed method guarantees fast and accurate decision making for the consecutive LS methods. The novel contribution of the proposed method is summarized as follows:

1) A comprehensive feature selection to investigate the relation of each power system data with FIDVR prediction.

2) Instead of using complicated method which may reduce the real-time prediction speed and increase computational programming, simple regression methods are employed

to enhance the processing speed along with several important power system features to increase the accuracy of the model.

3) An online ML training is accomplished to build a model for real-time FIDVR assessment. Linear and polynomial regression models are used and compared.

4) Validation methods have been proposed and used for analysing the proposed FIDVR prediction accuracy.

5) The FIDVR duration is assessed in a continuous manner. The real-time FIDVR prediction is executed as a regression method.

6) Since the model is simple, the real-time FIDVR prediction processing is fast enough which provides enough time for subsequent LS.

3.2. Proposed FIDVR Assessment Features

Although the characteristics of induction motors are critical for developing an accurate A/C load model, detailed power system data provides the tools needed to investigate FIDVR event characteristics. In this section, the relation of several power system data with FIDVR duration is analyzed.

3.2.1. Feature Analysis for FIDVR Prediction

To investigate the detailed characteristics of FIDVR events in distribution circuits, Southern California Edison (SCE) installed 22 power quality recording devices (PQubes) on 17 of its Valley Substation's 24 sub-transmission circuits that serve the utility's

residential and commercial customers [58][59][38]. By recording these data, lots of information can be obtained such as the understanding of how FIDVR events evolve and impact local residential and commercial customers. The data saves by PQubes can be used for building, validating and tuning ML models for FIDVR studies.

According to the recorded data, the majority of FIDVR events occurred during the summer seasons in hot climate areas when the weather temperature is high and also there is high penetration of operating residential A/C units. In this condition, a fault in the distribution network can decay the system voltage below the acceptable threshold and as a result of the voltage sag, A/C motors may stall. Hence, the weather condition could be considered as a possible feature which can be recorded and employed for FIDVR prediction.

The heat accumulation in the machine is the main reason for the operation of A/C unit's thermal protections and the heat mainly comes from several factors such as the stalling currents of stalled A/C units, the weather condition, duration of A/C units use prior to the contingency, the settings of A/C manufacturers, etc. This is impossible to use element-based approach to calculate FIDVR duration, since there are lots of uncertainties such as the number of A/C units, the manufacturing models of the machines, the settings of A/C units, initial conditions when stalling, accessibility of A/C units for doing the analysis, etc. Therefore, a power system data-driven based method should be used for FIDVR assessment. By analyzing the recorded FIDVR events occurred in the SCE system, following important observations are achieved [58][59][38]:

- ✓ The deeper the initial voltage sag, the greater the increase in the motor load reactive power demand.

- ✓ Voltage recovery is affected by the ability of the system to supply reactive power to the area of depressed voltage.

The first point clearly suggests that the degree to which voltage deviates from normal must be considered as a factor in FIDVR analysis method developed. The second point indicates that the reactive power deviations of the system in the area of depressed voltage must also be considered in any model development [42].

As a result of the SCE observations and deficiency of element-based methods, a ML-based power system data-driven decision-making process is presented in this chapter to predict FIDVR duration. To do so, several system characteristics have been selected to analyze their impact on FIDVR phenomenon duration. The selected features are voltage magnitude at the fault, post-fault voltage magnitude, active power and reactive power values injected to the loads before the fault. Active power value is added to the features to analyze its relationship with FIDVR and compare it to reactive power value. The aforementioned features can be obtained in real-time from the system measurements installed in the distribution network. In addition to these power system data, local weather temperature has been added as a potential feature which affects FIDVR. Note that post-fault voltage magnitude refers to the voltage when the system goes to quasi-static state after the fault is cleared. This feature selection is consistent with the test results in [60]. Post-fault voltage and fault voltage indices can be calculated as follows:

$$V_{Post-f} = \frac{V_{Post\_f}}{V_0} \tag{3-1}$$

$$V_f = \frac{V_f}{V_0} \qquad (3\text{-}2)$$

Where V_{Post-f} is post-fault voltage, V_f is fault voltage and V_0 is pre-fault steady-state voltage magnitude.

3.2.2. Feature Sensitivity Analysis

To have an accurate prediction, analyzing the correlation between the selected features and FIDVR duration is necessary. A straightforward data visualization can highly help in having valuable data analysis. Thus, the following data visualization methods are used in this study to demonstrate data correlations.

3.2.2.1. Heatmap

Heatmap plots are used to show the data dependency on two or more variables as a color coded image plot. Heat maps can be obtained by using multivariate functional outlier detection based on halfspace depth. The halfspace depth method is used to measure the centrality of a point relative to a multivariate sample. The halfspace depth of any point $x \in R^p$ relative to P_Y is defined as the minimal probability mass contained in a closed halfspace with boundary through x [61]:

$$HD(x; P_Y) = \inf_{||v||=1} \quad P_Y\{v'Y \geq v'x\} \qquad (3\text{-}3)$$

Where Y is a random variable on R^p with distribution P_Y. Halfspace depth satisfies the requirements of a statistical depth function. For any statistical depth function D and for any $\alpha \in [0, 1]$, the α-depth region contains the points whose depth is at least α:

$$D_\alpha = \{x \in \Re^p \quad ; \quad D(x; P_Y) \geq \alpha\} \tag{3-4}$$

Where the α-depth contour is the boundary of D_α. The halfspace depth regions are closed, convex and nested for increasing α [61]. To construct a heatmap, the cross-sectional depth values should be obtained. Each cell of this map is colored according to *HD* values with the biggest positive value shown as white and the highest negative depth value colored dark in Fig. 3-2. Fig. 3-2 is the heatmap of the SCE FIDVR data [58][59][38]. Fig. 3-2 is a two-dimensional graphical representation which demonstrates the dependency of the data to the selected variables. As mentioned in previous section and shown in Fig.3- 2, the variables of features are weather temperature, fault voltage, post-fault voltage, pre-fault active power and pre-fault reactive power. These variables are the inputs of ML method and are shown in the first five columns and rows of the heatmap matrix. The last column and row is the FIDVR duration which is the output of the ML technique. The individual values that are contained in the matrix are the correlation values which are represented as color coded image. As shown in Fig 3-2, the heatmap is a symmetric matrix. Features relation with FIDVR duration can be seen by analyzing the last column. In terms of feature importance, post-fault voltage (-0.92), weather temperature (0.72), pre-fault reactive power (0.64), fault voltage (-0.43) and pre-fault active power (0.35) were sorted from highest to lowest correlation number as shown in Fig. 3-2. Note that the *seaborn* package has been used in this study to create the annotated heatmap.

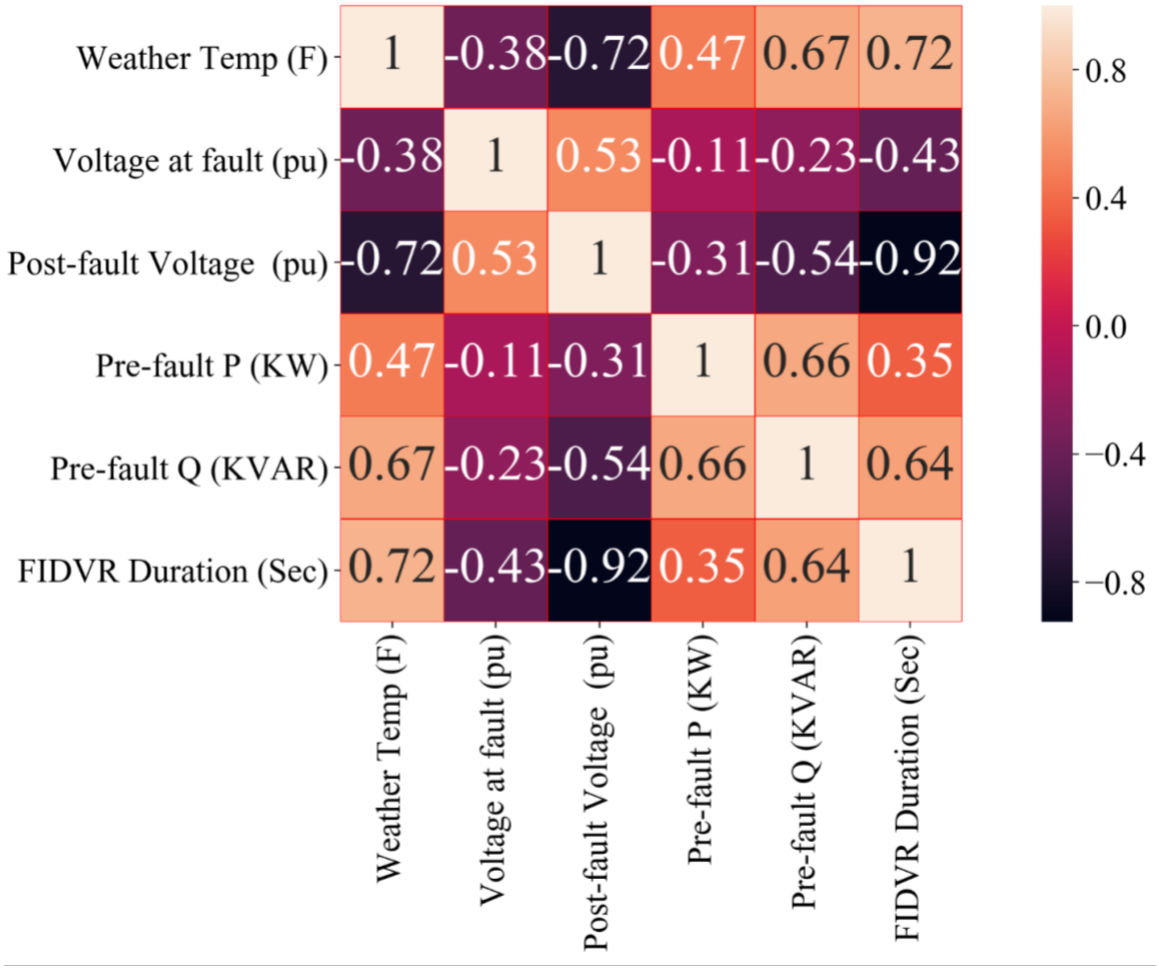

Fig. 3-2. Heat Map of the dataset

3.2.2.2. PairGrids

Pair grids are powerful tools to quickly explore distributions and relationships in a dataset. These plots use different pair of variables for each subplot and forms a matrix of sub-plots. Pair grid is used to understand the best set of features to explain a relationship between more than three variables in the dataset. In this study, five features are considered for FIDVR analysis and pair grid plots have been used to illustrate the relation between these features and FIDVR duration.

The pair grid shown in Fig. 3-3 maps each variable of the dataset onto a column and row in a grid of multiple axes. Each one of axes-level plotting functions are being used to draw bivariate (scatter) plots in the upper and lower triangles which provide the relationship between every two features of the dataset. These upper and lower triangles are the mirror image of each other. The diagonal plot showcases the histogram which demonstrates the distribution of a single variable.

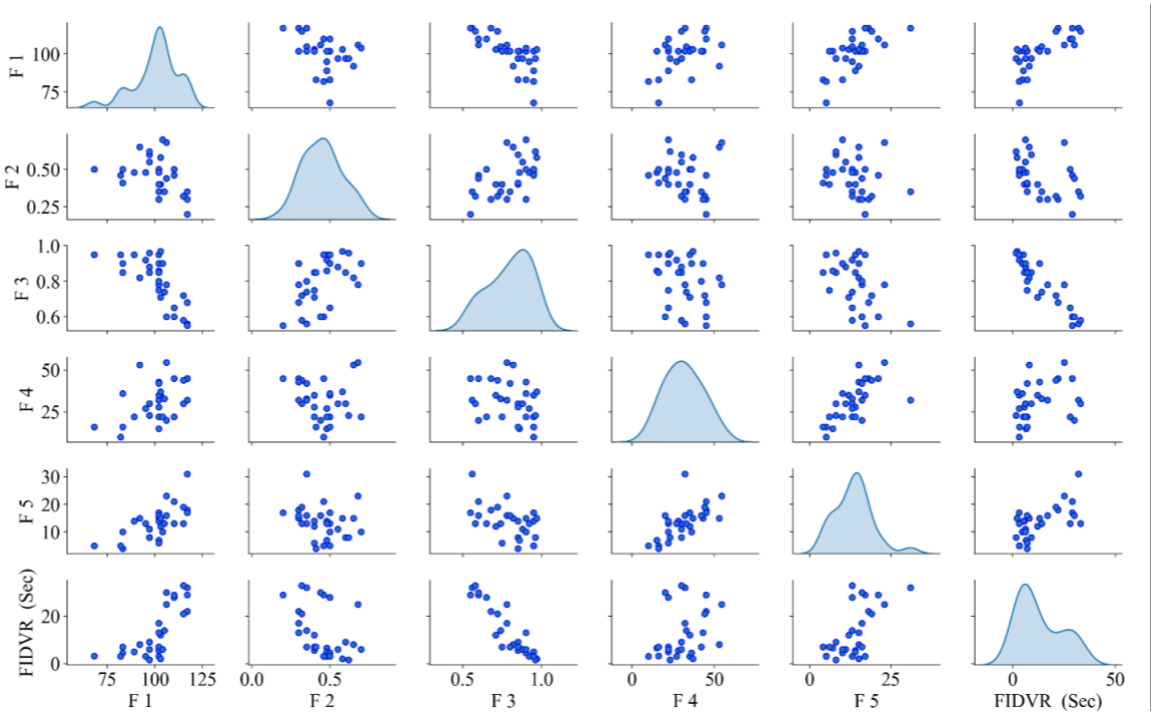

Fig. 3-3. Pair Grid plot of the features

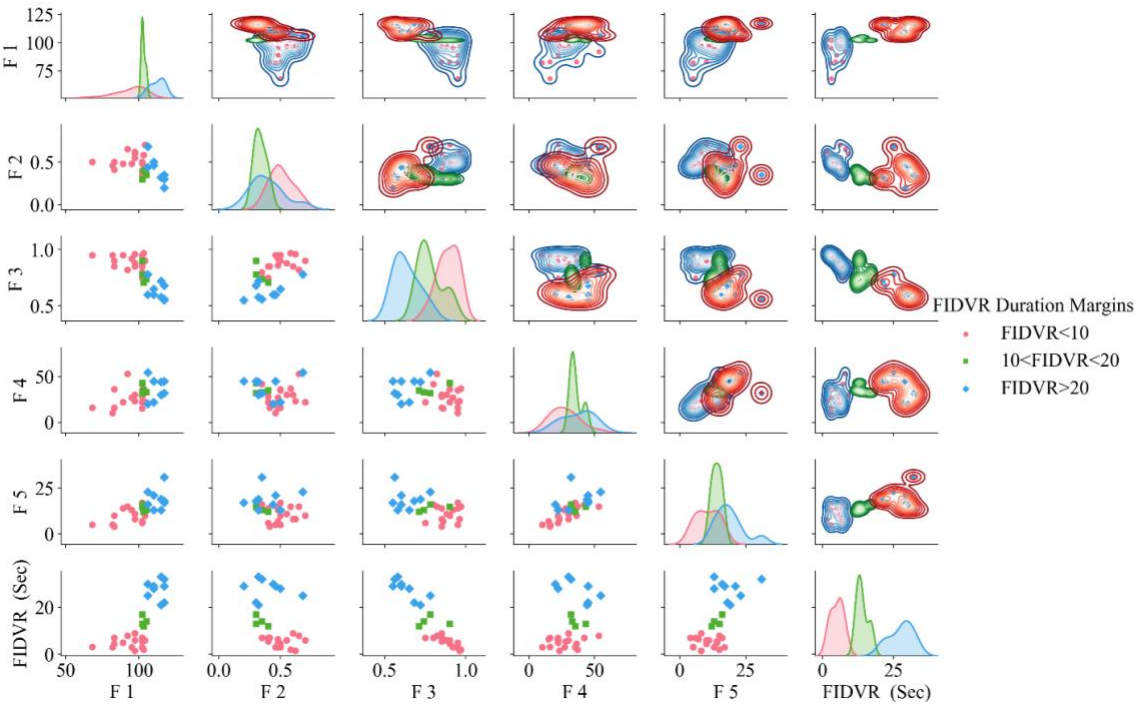

Fig. 3-4. Pair Grid plot of the features categorized according to the FIDVR duration. The higher triangle shows the bivariate kernel density estimation

In Fig. 3-4, a more detailed level of conditionalization has been demonstrated to differentiate various subsets of data in different colors. The *hue* parameter has been used for this purpose. By implementing this method, colors are being used to resolve elements on higher dimensions, but only draws subsets on top of each other. Valuable information is provided by coloring the data based on FIDVR duration margin as a categorical variable. The FIDVR cases which their duration are lower than 10 s, between 10 and 20 s and more than 20 s are colored red, green and blue respectively. Suppose pair grid matrix is called PG. As it can be seen from element PG_{61}, FIDVR cases with lower duration occur in lower weather temperatures. Element PG_{62} and PG_{63} show that lower voltage magnitude results in higher FIDVR duration. According to PG_{64}, the pre-fault active power index is not an appropriate feature for FIDVR duration analysis since the relation of FIDVR

duration margins and the values of the feature are indistinctable. Element PG_{65} shows that higher pre-fault reactive power result in higher FIDVR duration. On the top triangle, bivariate kernel density estimation of each two feature of the dataset has been demonstrated to visualize the distribution of the dataset. Elements of the 6^{th} column of PG demonstrate the kernel density estimation of each of the features and the FIDVR duration. By looking at this column, it can be seen that the fault-voltage and post-fault voltage have inverse relation with the FIDVR duration, while weather condition and pre-fault reactive power have direct relation. Pre-fault active power also has a slightly direct relation with the FIDVR duration.

After analyzing the features, it can be concluded that weather temperature, fault-voltage, post-fault voltage and pre-fault relative power could be selected as important features for assessment of FIDVR duration which can possibly enhance the prediction accuracy. More analysis will be accomplished in the following sections.

3.2.2.3. Feature Scaling

By analyzing the dataset, one can notice that the data consists of features highly varying in units, ranges and magnitudes, i.e. voltage magnitude in p.u., reactive power in kVAR and weather temperature in ^{o}F. ML algorithms use Euclidean distance method for training the data [62]. The measured length of the segment connecting two points in the space is the Euclidean distance. If only magnitude of features take into account (units neglected), the result will differ between various units and as a result, the weight of features with high magnitude will be more than low magnitude features in the distance calculations. To mitigate this issue, Min Max Scaler has been used in this study to bring the features to the

same magnitude level [63]. This method shrinks the range of data between -1 and 1. The data has been normalized according to the following equation:

$$x_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \qquad (3\text{-}5)$$

where x_i is the for the i-th case of the dependent variable X.

3.2.3. Partitioning data

An issue when fitting a ML model is its performance behavior while applying it to a new data. To address this issue, the data set is split to three partitions as shown in Fig. 3-5: a training partition which is used to train the model, a validation partition which is considered to test the performance of the model and test partition which is used for analyzing the efficiency of the model. Partitioning in this study is performed randomly to protect against a biased partition.

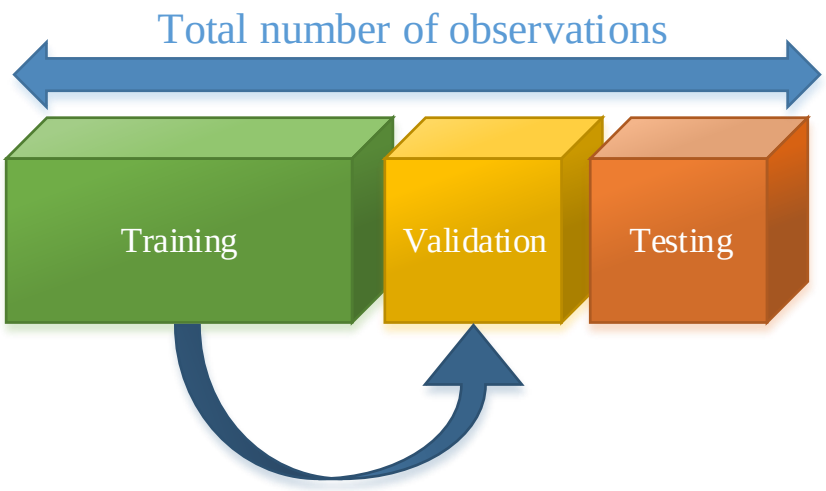

Fig. 3-5. Splitting the data for training, validation and testing

3.2.3.1. Training Set

The Training Set is used to build the ML model. The training set contains known outputs and the model learns on this data in order to be generalized to other data later on. In this chapter, the training set is used to fit the regression model and compute the regression coefficients. After fitting the model on the Training Set, the next step would be testing the performance of the model on the Validation Set.

3.2.3.2. Validation Set

The validation set is a new set of data which is used for validating the performance of the model, once a model is built using the training set. The validation set is a part of training data which are not used for model fitting. If the accuracy of the model computed using the training set, the result would be a highly optimistic estimate of model accuracy. It is due to the fact that the model is specifically suited to the training data and the fitting process is set based on the training data to guarantee that the accuracy of the model for training set is high enough. To overcome this issue, a part of original data is set aside as validation set and is not included in training process in order to obtain more realistic estimation of the model performance with unseen data. The discrepancy between the actual observed values and the predicted value of the observation are measured to validate the performance of the model. The discrepancy is known as the error in prediction and is used to measure the overall accuracy of the model.

Holding back a validation dataset is a valuable method in machine learning for estimating model accuracy on unseen data before implementing the model on test dataset.

3.2.3.3. Test Set

As mentioned above, the validation set is often used to fine-tune models. The test set is another portion of the data which is used to estimate the realistic performance of the model on completely unseen data. In other words, test set is used in order to evaluate model's prediction.

There are two approaches to partitioning: user-defined partitioning and random partitioning which is used for data splitting in this study.

3.2.3.4. Random Partitioning

In simple random sampling, every observation in the main data set has equal probability of being selected for the partition data set. If *P%* is set for train set, the *P%* of total observations are randomly selected for training set which means each observation has *P%* chance of being selected as training set. The same random selection is considered for Validation and Test sets.

3.3. Supervised Regression Algorithms

3.3.1. Linear Regression Algorithm

Regression analysis consists of identifying the relationship between a dependent variable and one or more independent variables. To implement the model, the relationship between input and output will be hypothesized, and the parameter values which are used to develop an estimated regression equation will be calculated. As discussed in the previous section, model validation is an important step in the modelling process which

helps in assessing the reliability of models before they can be used in decision making. Finally, various tests are employed to determine if the model is satisfactory.

3.3.2. Multiple Linear Regression Algorithm

Multiple linear regression method consists of more than one independent variables. The basic multiple regression model of a dependent variable Y on a set of k independent variables (x_k) can be expressed as [64]:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \cdots + \beta_k x_{1k} + e_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \cdots + \beta_k x_{2k} + e_2 \\ \qquad\qquad \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \cdots + \beta_k x_{nk} + e_n \end{cases} \tag{3-6}$$

Hence:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + e_i \quad for \;\; i = 1,2,\cdots,n \tag{3-7}$$

where y_i is the i-th case of the dependent variable Y , x_{ij} is the value of the j-th independent variable (X_j) for the i-th case of the dependent variable, β_0 is the Y-intercept of the regression surface, each β_j is the slope of the regression surface with respect to variable X_j and finally e_i is the random error component for the i-th case. In basic equations (1) we have n observations and k predictors (n> k+1). For each observation the errors (e_i) is distributed with mean zero and standard deviation σ ($e_i \sim N(0,\sigma^2)$) and is independent of the error terms associated with all other observations. The errors are uncorrelated with each other. The error is independent of other errors.

The variables (x_j) are fixed quantities which means the only randomness in Y comes from the error term. However in the context of correlation analysis, input variables are considered to be random variables. In any case, x_j are independent of the error term. In matrix notation, we can rewrite (1) as [64]:

$$Y = X\beta + e \tag{3-8}$$

Where:

$$Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T \tag{3-9}$$

$$\beta = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{k+1} \end{bmatrix}^T$$

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1(k+1)} \\ x_{21} & \ddots & & \\ \vdots & & \ddots & \\ x_{n1} & & & x_{n(k+1)} \end{bmatrix}$$

Where Y is the target vector and e is the error vector which is a column vector of length n, β is the vector of parameters which is a column vector of length $k+1$. Matrix X is the input matrix which is n by $k+1$ matrix. The first column of X has all elements equal to 1 and the rest of the columns are filled by the observed values of X_1, X_2, etc. to do prediction, β and e should be calculated.

3.3.3. Polynomial Regression Model

Polynomial regression is a special case of multiple regression in which dependent variables are regressed on powers of the independent variables. A polynomial regression model can be expressed as [64]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}^2 \cdots + \beta_k x_{ik}^k + e_i \quad for \;\; i = 1,2,\cdots,n \tag{3-10}$$

where k is the degree of the polynomial. Effectively, this is the same as having a multiple model with $X_k=X^k$ etc.

3.3.4. Accuracy Evaluation Indices

Least squares method is utilized to estimate the regression parameters. It measures the total deviation of the response values from the fit to the response values. The sum of the squared error or residuals sum of squares can be measured by:

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3-11}$$

where y_i are observed values and $y^{\wedge}_i$ are the fitted values of the dependent variable *Y* for the *i*-th case.

The mean squared error (MSE) is an unbiased estimator of the variance (σ^2) of the random error term and is defined as follows:

$$MSE = \frac{SSE}{n-(k+1)} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-(k+1)} \quad (3\text{-}12)$$

MSE is a measure of how well the regression fits the data.

By obtaining the square root of MSE, the standard deviation (σ) of the random error term can be estimated.

Root mean squared error ($RMSE = \sqrt{MSE}$) or also called the standard error of the regression is an estimate of the standard deviation of the random component in the data. The RMSE and MSE are in the range of the size of the regression errors and do not provide an indication about the explained component of the regression fit [65].

Mean absolute percentage error (MAPE) is a measure to compare the accuracy of predictions since it measures relative performance [66], [67]. MAPE can be measured as follows:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{\hat{y}_i}\right| \quad (3\text{-}13)$$

The smaller the MAPE, the better the prediction [66].

The R-squared coefficient of determination (R^2) is defined as:

$$R^2 = \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (3\text{-}14)$$

where $\bar{y}$ is the arithmetic mean of the *Y* variable which is used for calculating the total sum of squares (SST). R^2 is an important measure of how well the regression model fits the data since it measures the percentage of variation in the response variable *Y* explained

by the explanatory variable X. In other words, it is the square of the correlation between the response values and the predicted response values. Note that the value of R^2 is between zero and one ($0 \leq R^2 \leq 1$). The closer it is to 1, the better the prediction [65], [66], [67].

The advantage of R^2 over MSE and RMSE is that it is scale-free. However, R^2 has some problems that Adjusted R-squared is designed to overcome these issues. The first issue is that by adding predictor to a model, the R^2 increases. Consequently, a complex model may fallaciously appear to have a better result only by analyzing the R^2 value. The other issue is that an over fitted model may produce misleadingly high R^2 value which result in inaccurate predictions. Overfitting may occur by increasing model predictors (i.e. higher order polynomials), which result in modeling the random noise in the data. Adjusted R-squared is computed by:

$$R^{*2} = R^2 - \frac{(1-R^2)k}{n-(k+1)} \tag{3-15}$$

Adjusted R-squared (R^{*2}) is always smaller than R^2. R^{*2} is adjusted for the number of variables included in the regression equation. If the value of R^{*2} is much lower than R^2 value, it is an indication that the regression method may be over-fitted [65], [66]. R^{*2} is a great measure to compare models with different numbers of predictors. By adding useless (useful) variable which has low correlation with output to the model, R^{*2} will be decreased (increased).

3.4. **Study System**

This chapter studies the SCE Valley system which appears to be a susceptible network to FIDVR events.

The valley network consists of a transmission system which contains of two 115 kV busses (section A&B and section C&D) as shown in Fig. 3-6. Each of the 115 kV substation busses feeds a meshed sub-transmission system.

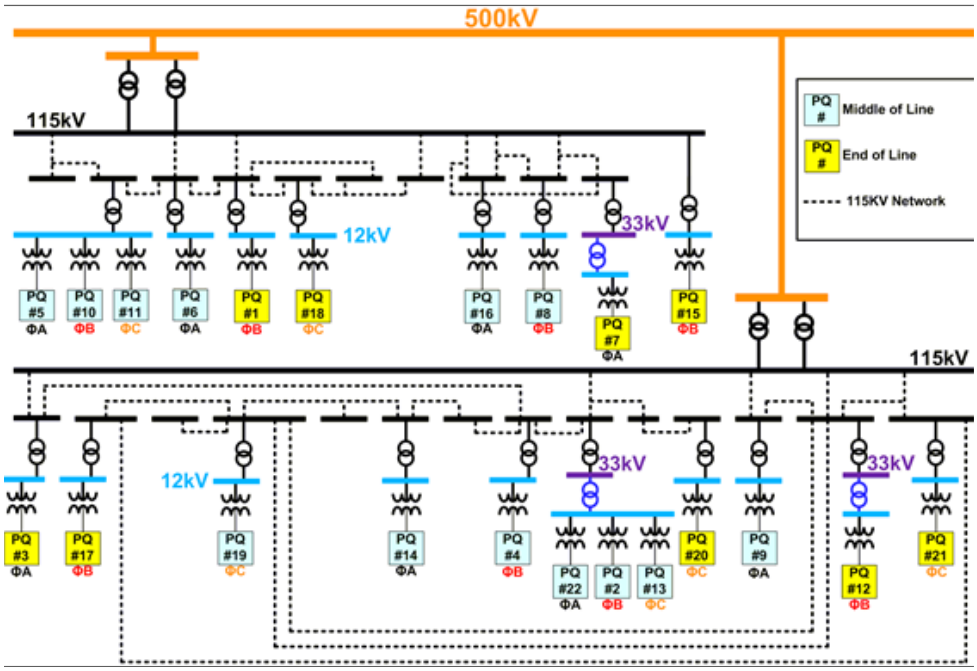

Fig. 3-6. Valley distribution system and PQubes Locations Diagram [38]

There are 24 meshed sub-transmission 115 kV substations in the network. Sub-transmission 115 kV substations contain two types of distribution circuits 33 kV and 12 kV, most of which are 12 kV. 12 kV distribution circuits used for both commercial and residential circuits with pad-mount and pole-mount transformers to serve customers. Longer distribution circuits are mainly rural which use 33 kV circuits instead of the 12 kV

distribution circuits. The PQube devices were installed in the pad-mount transformer's secondary (240 V) side that fed to customers. For each distribution circuit, the device installations were located either at the middle or the end of the line. These data recording devices were also placed on different phases of the circuits to acquire a diverse collection of event data.

3.4.1. Power Quality Recording Device specifications

The power quality recording devices (PQubes) installed in valley's network can record up to five voltages and five currents during steady-state conditions as well as during system events. These devices were programmed to record both root mean square (RMS) and sinusoidal waveforms when an event is triggered. RMS event data captured at 1 sample/cycle and sinusoidal waveform event data captured at 32 samples/cycle. Every device was equipped with an uninterruptible power supply (UPS) for up to nine minutes so that will record during events of low voltage without compromising the data. Each SCE PQube contains components such as PQube module, current module, power supply, circuit breaker, current transformers, din rail and enclosure.

3.5. Numerical Results

3.5.1. A. Database

Based on a comprehensive literature review, several single-variable regression and Multi-linear regression and polynomial regression models built using real FIDVR data

collected from SCE valley network [58]-[59]-[38]. The mentioned features in section III have been used for training the ML algorithms. Linear regression and polynomial regression techniques are robust and easy to implement ML models which can guarantee fast prediction of FIDVR events. However, these models may not have acceptable result based on single feature training. By using multiple features proposed in this study, the method accuracy can be improved, while the model simplicity can assure the fast prediction.

The goal is to train linear regression and polynomial regression on four features selected in section III. The features are weather temperature, fault voltage, post-fault voltage and pre-fault reactive power. The relation of these four features with FIDVR duration is a 5-dimension problem which cannot be shown in a 2-D plane.

3.5.2. Single Variable Regression

To demonstrate the behavior of each feature on FIDVR duration, four single-variable linear regression, polynomial with order of 2 and polynomial with order of 3 are analyzed respectively.

Fig. 3-7 demonstrates the mentioned regression methods when the input is historical weather temperature. As it can be seen, weather temperature has direct relation with the FIDVR duration. This result is similar to the one obtained by heatmap in Fig. 3-2 where the weather temperature and FIDVR duration correlation is +0.72. This relation is due to the fact that when weather temperature increases, more A/C motor loads are connected to the grid. The more the number of constant torque induction motors, the higher the risk of delayed voltage recovery. Hence, FIDVR duration increases by increase of induction loads. According to Fig. 3-7, polynomial regression with order of two and three has a

better fit to the data compared to linear regression. Since order two and three are approximately similar in fitting the data, it is better to use order two to reduce the model complexity.

Dependency of FIDVR duration on fault voltage is shown in Fig. 3-8. As it can be seen, lower fault voltage results in higher FIDVR duration. This result is confirming the correlation value (-0.43) obtained from heatmap on Fig. 3-2. The effect of this feature can be explained by the fact that low fault resistance results in high fault current which lead to voltage drop. The lower the voltage drop, the higher the possibility of stalling a greater number of induction motors. The higher the number of motors stall, the more time is needed to recover the dropped voltage. According to Fig. 3-8, second order and third order polynomial regression result in quite similar data fitting while the third order is more complex.

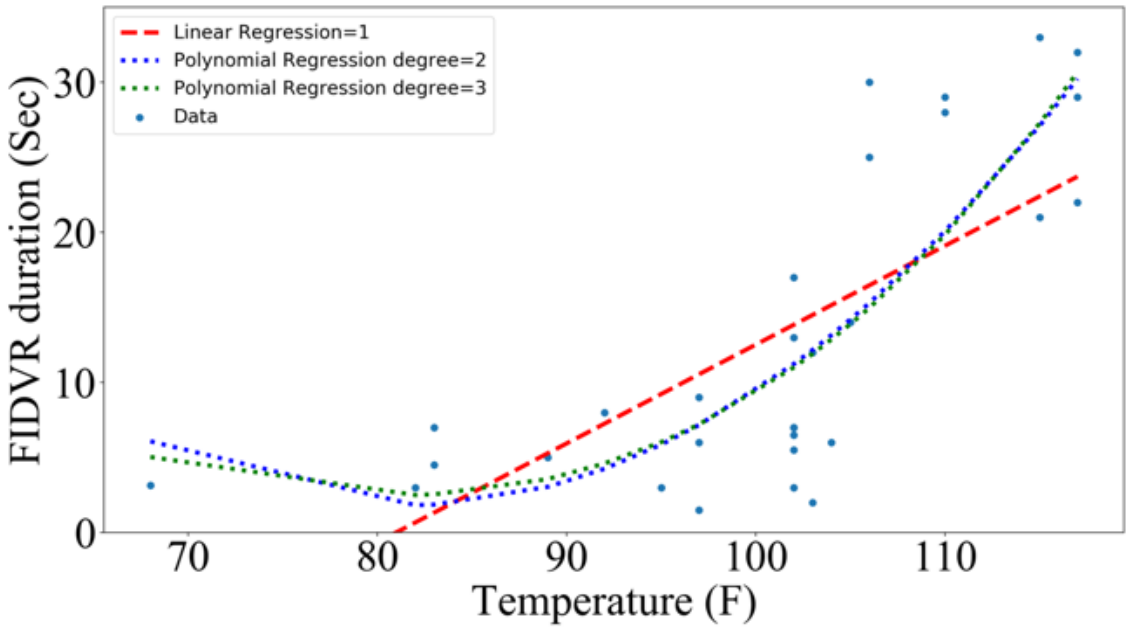

Fig. 3-7. The relation of the FIDVR duration and local weather

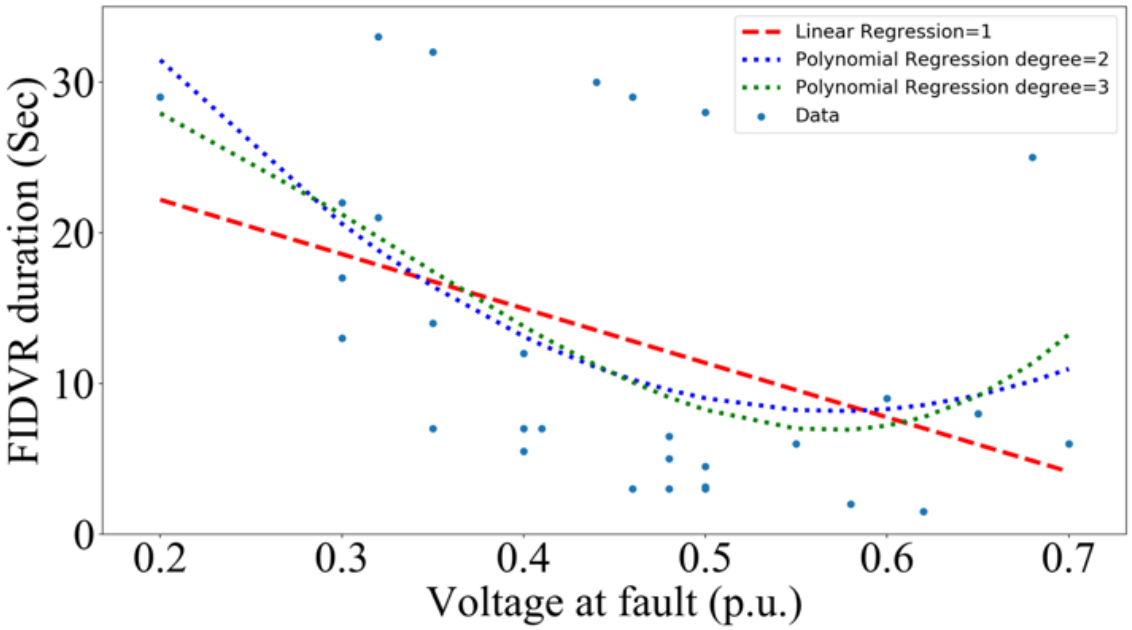

Fig. 3-8. The relation of the FIDVR duration and voltage magnitude at fault

Fig. 3-9 demonstrates the post-fault voltage magnitude's impact on FIDVR duration. As mentioned before, post-fault voltage magnitude refers to the voltage when the system goes to quasi-static state after the fault is cleared. Based on the fitted curves in Fig. 3-9, linear regression and second and third orders of polynomial regression have similar behavior. It can be seen that the slope of the curves fitting the post-fault voltage are higher than other features which means that the post-fault voltage has a better fit to the data compared to other features; meaning that it has stronger relation with the output which verifies the previous conclusion in the heatmap on Fig. 3-2 (correlation= -0.92). The reason why post-fault voltage has higher impact on FIDVR duration compared to fault voltage can be explained by the fact that the fault voltage is a transient value which may change when

the fault is cleared. However, post-fault is a slow varying quantity. In addition, it can be seen that the effect of post-fault voltage on the voltage delay duration is higher than weather temperature. This can be explained by the fact that the internal heat generated by stalling current is more crucial than external temperature.

Fig. 3-10 shows the pre-fault active power relationship with FIDVR duration. As it can be seen it has direct relation with the output. However, the fitting curve slopes are not high which demonstrates its low correlation with the duration. This is also shown in the heatmap on Fig. 3-2 where the pre-fault active power correlation with FIDVR duration is 0.35 which is the lowest correlation among the features.

The relation of pre-fault reactive power with FIDVR duration is plotted in Fig. 3-11. The data fitting using this feature has best performance after post-fault voltage and weather temperature. The correlation factor for this feature is 0.64 as shown in Fig. 3-2. The slope of the curves, fitting the pre-fault reactive power data, are higher than those of pre-fault active power data. This, shows the importance of pre-fault reactive power data in FIDVR analysis.

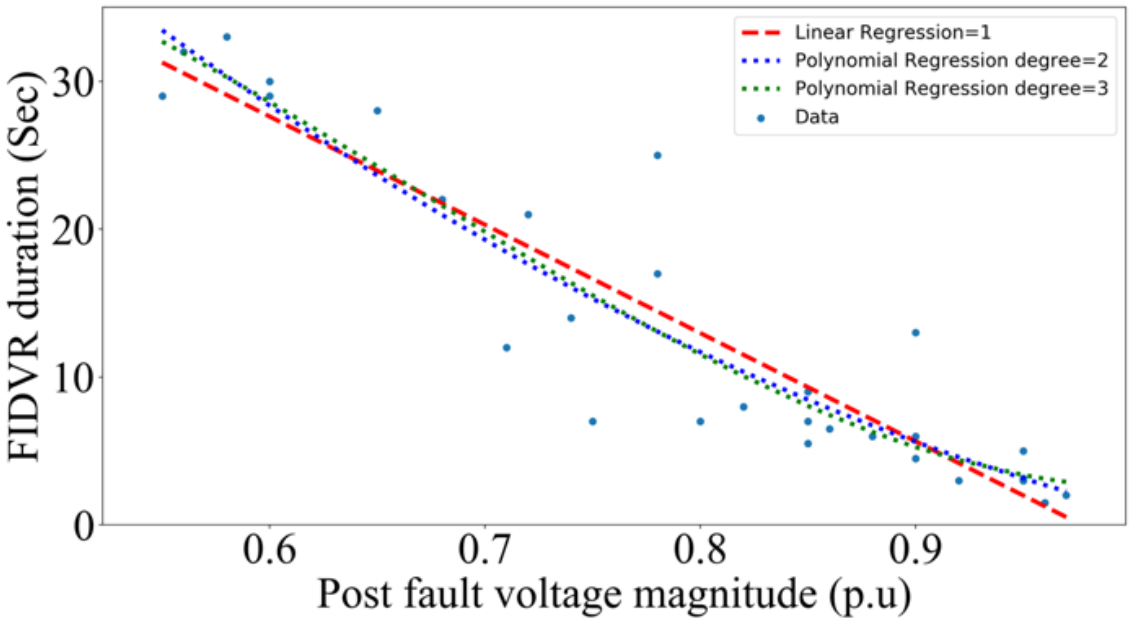

Fig. 3-9. The relation of the FIDVR duration and post-fault voltage magnitude

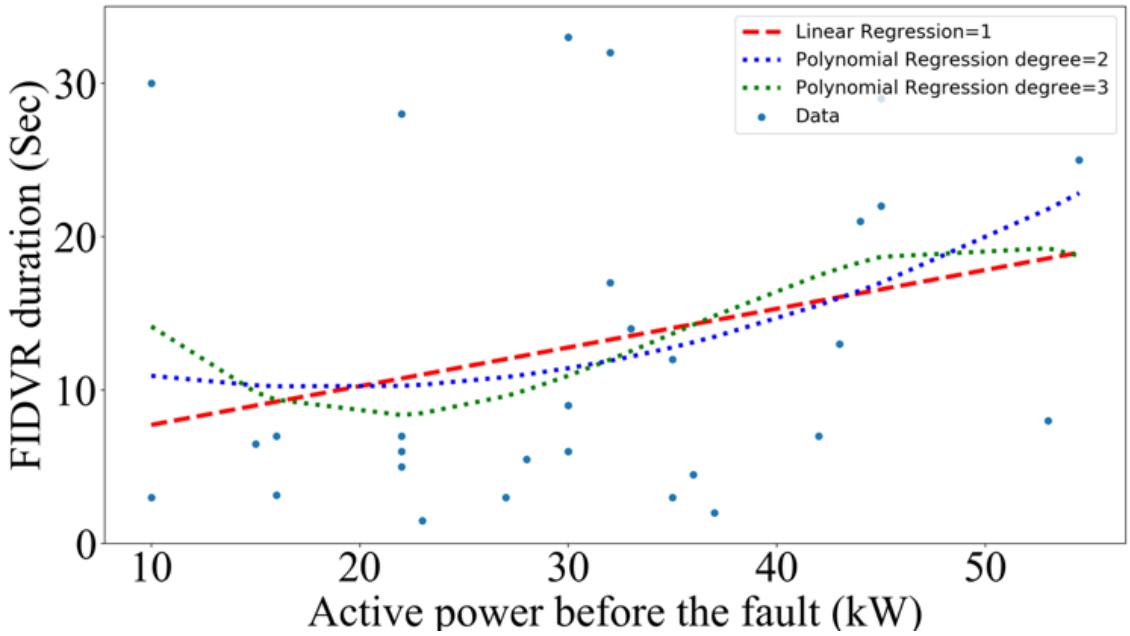

Fig. 3-10. The relation of the FIDVR duration and pre-fault active power magnitude

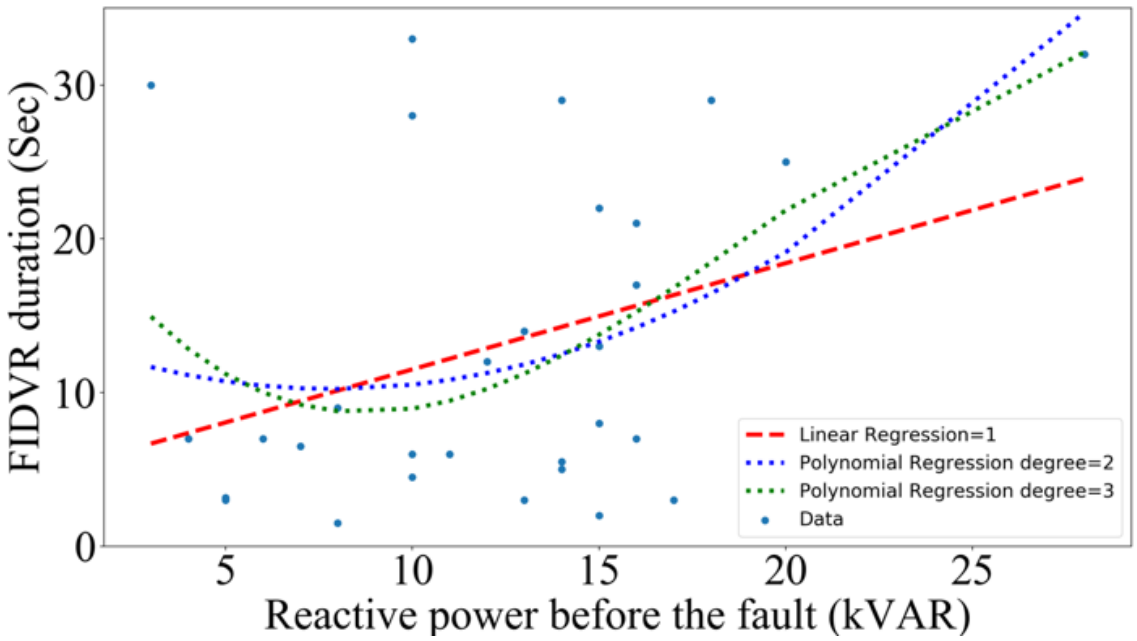

Fig. 3-11. The relation of the FIDVR duration and pre-fault reactive power magnitude

The pre-fault active power is eliminated as the features in multi-variable regression analysis due to its high variance and its low correlation with FIDVR duration. In addition, this feature is dependent on pre-fault reactive power, while the selected features for ML training should be independent.

Figs. 3-7 - 3-11 show the regression analysis verifying the correlation values shown in the heatmap in Fig. 3-2. Table 3-1 shows all conditions of single-variable linear, second order, third order and forth order polynomial regressions. F1-F5 features are weather temperature, fault voltage, post-fault voltage, pre-fault active power and pre-fault reactive power magnitudes, respectively. As shown, linear regression (LR) has only one coefficient, while second, third and fourth order polynomial has 2, 3 and 4 coefficient or parameters. It can be seen that the best model is the fourth order polynomial regression which uses post-fault voltage magnitude as the feature. MSE, RMSE and MAPE in this

case are 13.65, 3.6949 and 27.2079 respectively which are the lowest values and R^2 and R^{*2} are the highest values (0.873, 0.795) compared to other cases. Note that closer R^2 value to 1 indicated that a greater proportion of variance is accounted for by the model. R^2 equal to 0.837 means that the fit introduces 83.7% of the total variation in the data.

Table 3-1
SINGLE-VARIABLE REGRESSION MODELS VALIDATION

Feature	ML Model	Intercept	Coefficients	MSE	RMSE	MAPE	R^2	R^{*2}
F1	**LR**	13.0716	7.429	52.33448	7.2342	113.76	0.513	0.215
	Poly=2	10.174	9.838, 2.897	35.109	5.925	78.473	0.673	0.474
	Poly=3	10.0	9.309, 3.25, 0.22	34.9568	5.912	74.526	0.6749	0.476
	Poly=4	8.6647	17.3595, 7.418, -4.5108, -1.806	27.782	5.2708	68.5024	0.7416	0.5837
F2	**LR**	13.07166	-4.41705	88.0185	9.38182	113.3746	0.18144	-0.318
	Poly=2	10.53871	-4.9089, 2.5329	79.194	8.899	97.5764	0.2635	-0.1865
	Poly=3	10.4779	-6.8105, 2.439, 0.7956	78.045	8.8343	92.7	0.274	-0.169
	Poly=4	10.8889	-7.0353, 1.245, 0.87619, 0.318	77.8267	8.8219	92.0285	0.2762	-0.166
F3	**LR**	13.07	-9.574	15.8638	3.9829	33.696	0.852	0.762
	Poly=2	11.784	-8.953, 1.287	14.6026	3.82	27.2719	0.864	0.78
	Poly=3	11.633	-9.895, 1.716, 0.576	14.449	3.8012	28.631	0.865	0.783
	Poly=4	10.75	-8.295, 5.075, -1.067, -1.64	13.65	3.6949	27.2079	0.873	0.795
F4	**LR**	13.071	3.596	94.5955	9.726	121.568	0.1202	-0.4173
	Poly=2	13.201	3.6175, -0.12997	94.57467	9.72495	121.5749	0.12	-0.417
	Poly=3	13.253	2.7239, -0.247 0.4035	94.3235	9.712	121.3176	0.1228	-0.4132
	Poly=4	11.73	2.189, 3.981, 0.819, -1.226	91.71336	9.5767	123.519	0.147	-0.374
F5	**LR**	13.071	6.6215	63.684	7.98024	99.5429	0.4077	0.0458
	Poly=2	12.8515	6.4717, 0.22	63.555	7.9721	99.0865	0.4089	0.0477
	Poly=3	12.0716	8.4218, 1.46, -0.6767	61.9498	7.8708	94.9849	0.4238	0.0718
	Poly=4	12.001	8.0843, 1.775, -0.4515, -0.096	61.916	7.8686	95.042	0.424	0.072

Table 3-1 also shows that R^2 values of regression methods on F2 and F4 increases by model complexity. However, the R^{*2} values are negative. The negative R^{*2} may appear when SSE approaches to the SST. This condition occurs when the explanation towards response is very low or negligible. Hence, negative R^{*2} is the result of insignificance explanatory variables. R^{*2} value can be improved by increasing the variables or sample size or eliminating correlated independent variables. According to Table 3-1, except some mentioned conditions, by model complexity, MSE, RMSE and MAPE of all conditions decreased while R^2 and R^{*2} of all conditions increased. This means that with higher model complexity; the model fits the data better. On the other hand, the processing time increases with model complexity.

In the next section, the result of these single-variable regression methods will be compared to multi-variable methods.

3.5.3. Multi-variable Regression

In the previous section, five single variable regression methods have been used in order to demonstrate the effect of each feature on FIDVR duration. To enhance the prediction accuracy, features F1, F2, F3 and F5 are used for multi-variable ML algorithm training and the feature F4 is eliminated.

Table 3-2 demonstrates the value of validation indices for multi-variable linear regression, second order and third order polynomial using the four mentioned features as input data. MSE and RMSE of all multi-variable ML approaches are lower than single-variable regressions. In addition, R^2 and R^{*2} values are improved compare to Table 3-1. However, the increase of R^2 is not a reliable measure for comparing the results, since R^2 would not decrease when more variables are added to the model. Therefore, R^{*2} should be

used as a measure. Note that both R^2 and the R^{*2} provide the number of data points fall within the regression equation. However, R^2 assumes that every single variable presents the variation in the dependent variable. R^{*2}, on the other hand, provides the percentage of variation explained by only the independent variables that actually affect the dependent variable. In the process of comparing regression models that use the same dependent variable and the same estimation period, the RMSE decreases as adjusted R-squared increases. Hence, the model with the highest R^2 will have the lowest RMSE, and R^2 can be used as a measure. However, when comparing regression models in which the dependent variables are not the same, or different sets of observations are used, R^2 is not a reliable guide to model quality. As shown in Table 3-2, both R^2 and R^{*2} have higher values compared to Table 3-1. R^{*2} of second order polynomial regression shows that 95% of variation explained by the important selected features which is an acceptable value for FIDVR prediction. The results of third order polynomial regression demonstrates overfitting of this model since R^2 and R^{*2} are equal to 1 and MSE, RMSE and MAPE are approximately zero. This result is too accurate and shows that the model is fit too precise for dataset but may lose credibility for a new dataset.

Figs. 3-12, 3-13 and 3-14 show the prediction vs the real test output value using multi-variable linear regression, second order polynomial and third order polynomial regression, respectively. Note that by using four features, the total number of parameters for linear regression, second order polynomial and third order polynomial are 5, 9 and 13, respectively. X-axis is the real FIDVR duration and Y-axis is the predicted duration. The closer the points to the diagonal line, the more accurate the prediction is. It can be seen

from Fig. 3-12 that the prediction error using second order polynomial regression is lower than linear regression.

Table 3-2
MULTI-VARIABLE REGRESSION MODELS VALIDATION

Validation	*Linear Regression*	*Polynomial Regression*	
		Polynomial degree=2	*Polynomial degree=3*
MSE	13.065369	1.925118	3.07157e-21
RMSE	3.614605	1.387486	5.54217e-11
MAPE	35.60462	15.462729	6.50849e-10
R Squared	0.878494	0.952096	1
Adjusted-R Squared	0.804240	0.941155	1

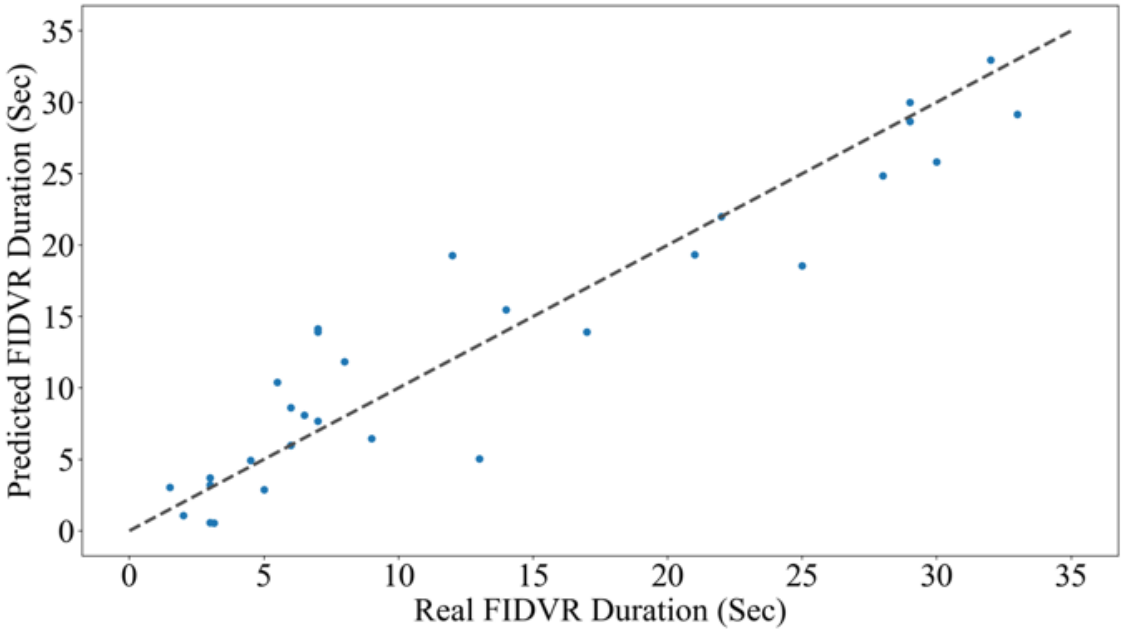

Fig. 3-12. Real value (Real FIDVR duration) vs predicted value (predicted FIDVR duration) using linear regression

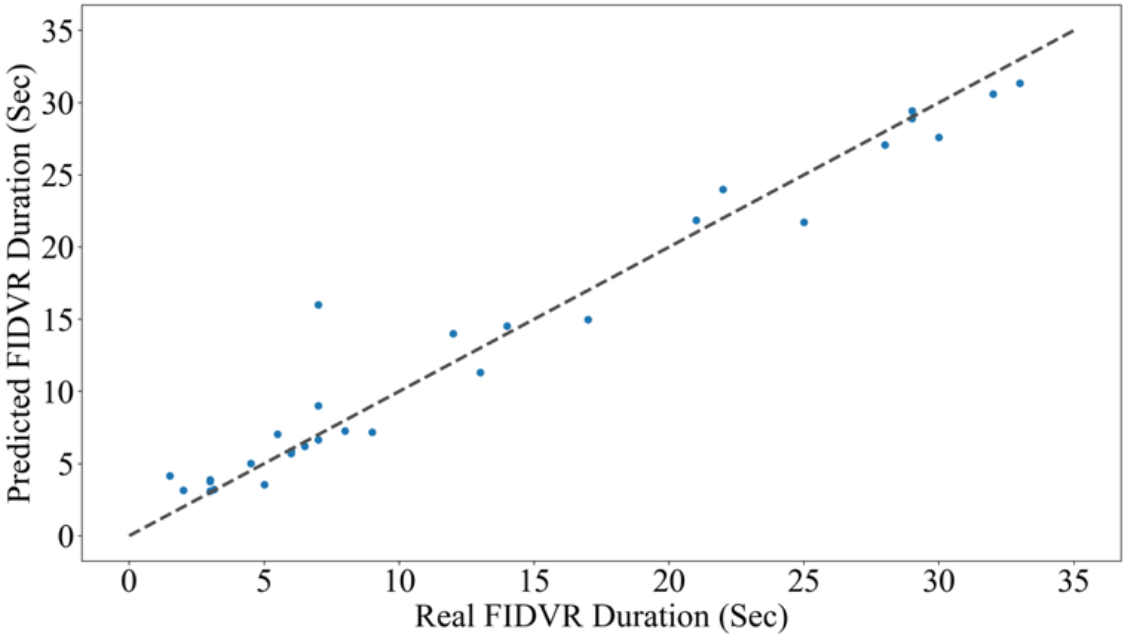

Fig. 3-13. Real value (Real FIDVR duration) vs predicted value (predicted FIDVR duration) using polynomial regression with degree of 2

Fig. 3-14 shows that the prediction is too accurate which is overfitting modeling error. When a function is too closely fit to the set of data points, overfitting occurs. The reason of overfitting is that the model becomes too complex to explain idiosyncrasies in the data. The measured data always has some degree of random error or noise within it which is ineluctable. Thus, attempting to design a model which conform too closely to the measured data can reduce the prediction power and infect the model with errors. Over time, when more data is added to the data set, model error on training data decreases and so does the error in the test data. If the model trains too much, the model error on train data still decreases while the model error on test dataset increases at one point. The reason behind continuous decrease in train data error is that the model in overfitting and learning irrelevant details and noise. The error of the test data set, on the other hand, rises due to the fact that the model loses the ability to generalize. The most appropriate model is when

the model error for both train and test datasets are in acceptable range. In this case, the second order polynomial model has good performance in predicting unseen test datasets. According to Table 3-2, the performance of second order polynomial regression on both training and test sets are acceptable, while the accuracy of linear regression is lower than the second order polynomial model and the third order polynomial model is overfit. This means that the model is trained only on the training data too much in such a way that will have noticeable error on new dataset.

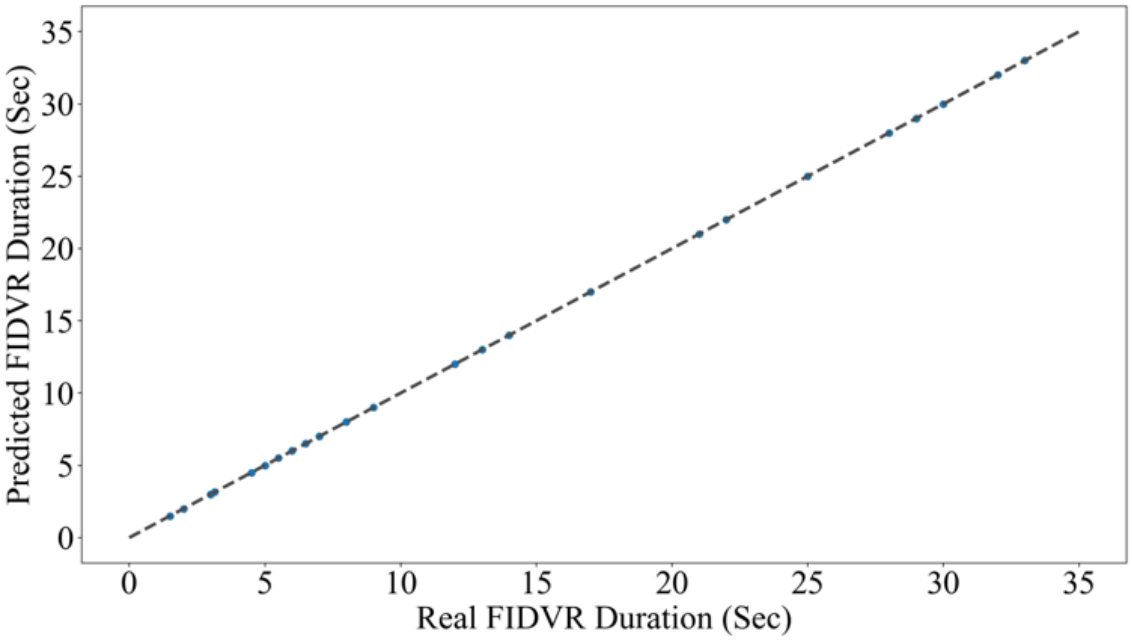

Fig. 3-14. Real value (Real FIDVR duration) vs predicted value (predicted FIDVR duration) using polynomial regression with degree of 3

The main purpose of FIDVR prediction is to employ appropriate LS approach to enhance the system voltage stability. Hence the prediction speed plays an important role in having proper voltage recovery. Late FIDVR duration prediction results in late LS

action which in severe cases may increase the risk of voltage collapse. The tests are performed on a computer Core i5 with 2.6 GHz CPU and 4 GB RAM. The process of ML training For second order polynomial regression, the average online training time of four probabilistic features is 2.1 *s*, while the average testing time is only 6.29 *ms*. Note that although real-time testing time is more important than the online training time, in case of multiple FIDVR events happening in a few minutes, it is important to have a fast online training in order to update the dataset based on the new data. According to a FIDVR event reported by [38], on July 29, 2014 a FIDVR event happens at 15:15:11 PDT and after 16 seconds another fault result in the next FIDVR. Another FIDVR event on August 12, 2012 at 15:51 PDT was observed and recorded by the PMU and after 8 minutes (15:59 PDT) another event was recorded. These data can be used for developing a more robust ML model and can be used for following real-time FIDVR prediction, if the training process becomes fast enough. To compare the method testing speed note that the FIDVR assessment time in the literature [46], [17], [44], [45] are 0.14, 1.5 *s*, 0.35 *s*, and 0.5 *s*, respectively. Although more complicated ML methods can be used for enhancing the FIDVR prediction accuracy, the simplicity and fast processing of polynomial and linear regression approaches can be beneficial in the view of online training and testing time. In addition, the accuracy of the model can be increased by adding more features to the model.

Fig. 3-15 shows the relation of FIDVR duration and fault voltage and post-fault voltage indices in a 3D demonstration. As it can be seen, the FIDVR duration increases by decrease of selected features. Fig. 3-16 shows the relation of FIDVR duration and weather temperature and pre-fault reactive power indices in a 3D demonstration. As it can be seen, the FIDVR duration increases by the increase of selected features.

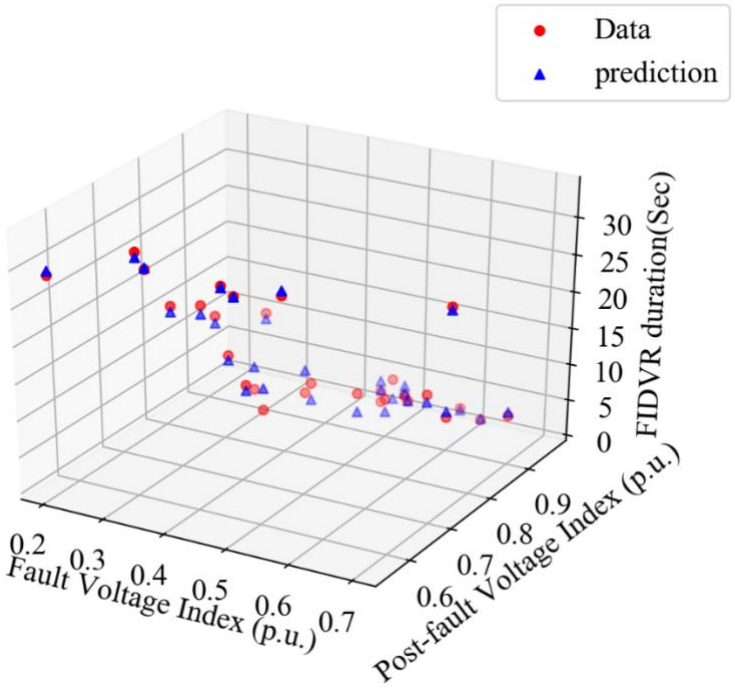

Fig. 3-15. 3D demonstration of the real value vs predicted value using multi-variable second order polynomial regression

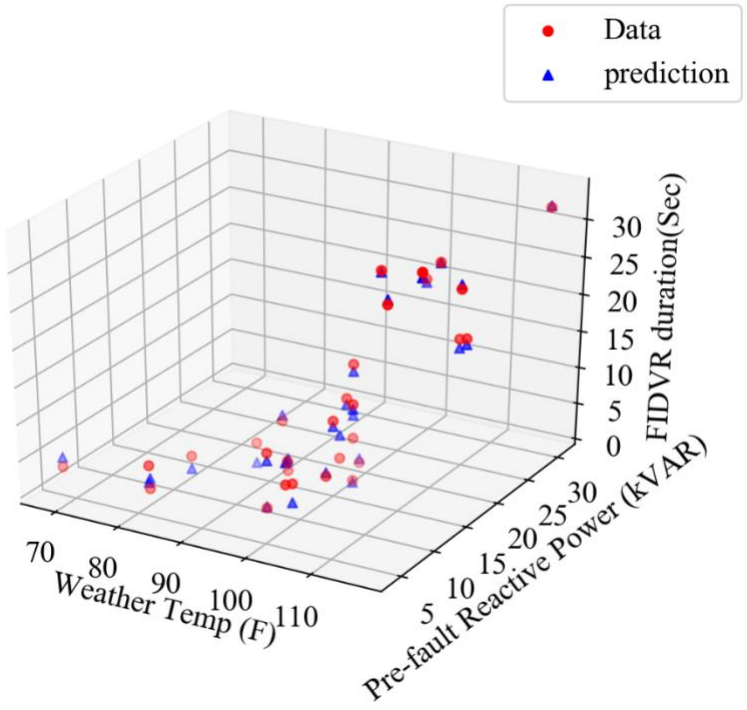

Fig. 3-16. 3D demonstration of the real value vs predicted value using multi-variable second order polynomial regression

3.6. **Conclusion**

This chapter proposed a data-driven multi-variable machine learning-based Decision-making model for prediction of FIDVR duration. Relation of several power system data with FIDVR duration has been investigated. Selected features have been utilized for online machine learning model development. The stored data of power quality recording devices (PQubes) are used for online machine learning model developing. Several simple regression models such as single-variable and multi-variable regression models i.e. linear regression and polynomial regression models have been developed and compared. The real-time PQubes data can be used for rapid real-time prediction of the FIDVR duration following a system disturbance. By using the developed model, FIDVR duration can be obtained using real-time data without using complicated load modeling. Simple regression models result in simple programming and fast prediction and on the other hand, using multiple features for model development enhance the model accuracy. As a result, the FIDVR prediction can be accomplished very fast, yet with acceptable accuracy.

Chapter 4 : CLASSIFICATION AND REGRESSION DECISION-MAKING FOR REAL-TIME FAULT-INDUCED DELAYED VOLTAGE RECOVERY ASSESSMENT BASED ON A PROBABILISTIC TIME-SERIES DATA-DRIVEN MULTI-VARIABLE APPROACH

4.1. Introduction

As it has been explained in section 3.1.1 and 3.1.2, FIDVR events occur frequently since the density of induction motors is increasing with continuing market penetration of low-inertia air conditioning (A/C) loads which are not equipped with compressor under voltage protection. A comprehensive literature review has been done. Among all the methods available in the literature, employing a fast FIDVR detection and implementing appropriate emergency control i.e. load shedding seems to be the best approach for enhancing the power system stability margins. Wide area measurement systems and the subsequent data-driven control methods provide real-time post-fault system characteristics measurement [47] which can be used for FIDVR detection. By taking benefit from machine learning methods for FIDVR analysis, a fast real-time decision making along with less data requirement and less computation time can be achieved while analytical data-driven methods did not guarantee fast computation for FIDVR analysis. Chapter 3 focused on FIDVR duration analysis using simple machine learning methods. In chapter 3, several features have been introduced and a comprehensive sensitivity

analysis has been implemented to demonstrate the relation of each feature with FIDVR duration. A data-driven single variable and multi-variable linear and polynomial regression method has been proposed to predict the duration of FIDVR. In this chapter, a time-series data-driven multi-variable simultaneous classification and regression decision making method is proposed which can operate in the first milliseconds after the system fault to simultaneously predicts whether the event can be categorized as FIDVR or not; if yes, assesses the FIDVR duration.

4.1.1. Problem Identification

In real world, applying inaccurate prediction may result in false alarm which lead to unnecessary emergency control i.e. load shedding or missing as alarm which in severe cases may lead to voltage collapse. In order to enhance the machine learning assessment accuracy, reference [68] evaluates the credibility of the decisions made by the learning model and reference [69] designs a time-series classifier for MG islanding detection however, this model only focuses on binary classification problems, which is not applicable to FIDVR duration assessment. Therefore, a more advanced method needs to be specially designed to improve the real-time FIDVR assessment speed. To overcome the aforementioned issues, this study proposes a robust validation technique to enhance the real-time prediction accuracy and presents a time-series probabilistic decision making approach to simultaneously predict the possibility of the FIDVR event occurrence and predict the duration of the event in the case of classifying it as FIDVR.

4.1.2. Contribution

In this chapter a simultaneous classification and regression assessment is proposed using a time-series measurement-based machine learning technique for real-time FIDVR prediction. The proposed method employs the same features similar to the explained features in section 3.3. In addition to the mentioned features, new indices will be proposed and their correlation with FIDVR duration will be investigated. These indices are active power deviation index (PDI), reactive power deviation indices (QDI), active power increment slope index (PISI) and reactive power increment slope index (QISI). The novel contributions of the proposed method are summarized as follows:

1) An advanced feature analysis is performed to investigate the effectiveness of power system data on FIDVR phenomenon.

2) A probabilistic k-fold validation method is employed for enhancing the proposed FIDVR prediction accuracy.

3) A real-time probability assessment of FIDVR phenomenon occurrence is accomplished as a time-series classification machine learning decision making algorithm.

4) The FIDVR duration is assessed in a continuous manner. The real-time FIDVR prediction is executed as a time-series regression method.

5) An ensemble machine learning approach has been utilized in each step of both time-series classification and regression decision making models which guarantees the test score improvement in real power system event.

The proposed method has been tested on SCE system and demonstrated higher accuracy and speed compared to other state-of-the-art methods.

4.2. Advanced feature analysis

To further investigate the relation of power system data and FIDVR duration, several indices are proposed and compared to each other in this section. Note that the accuracy of machine learning methods can be improved by increasing the features which have high correlation with the output variable. Therefore, feature analysis and feature selection can play an essential role in FIDVR assessment.

4.2.1. Proposed indices

Dynamic behavior of induction motor loads is the major cause for FIDVR events. These motors decelerate and stall following a large disturbance, resulting in low voltages in a significant portion of the power system. The stalling results the subsequent increase of reactive power demand of the induction motors and the subsequent reactive power increase prevents quick voltage recovery. According to the recorded FIDVR events occurred in SCE power network [58], [59], [38], the reactive power deviations of the system in the area of depressed voltage is an important feature which can be used in model development [42]. Some sample FIDVR events are listed below to illustrate the importance of investigating the relation of active power and reactive power deviation and FIDVR event.

The RMS data of a FIDVR event occurred at 14:14 PDT on August 10, 2012 is shown in Fig.4-1. This FIDVR event occurred as a result of a lightning strike in the SCE network and as it can be seen, the event lasted for approximately 12 seconds. It can be seen that a noticeable active power and reactive power jump occurs in the very beginning of the event. In another event happened at 17:02 PDT on August 30, 2013, active power and reactive power deviation in one second after the fault is 2.7 and 4.3 times their pre-fault values. According to the recorded data shown in Fig. 4-2, the FIDVR event lasted approximately 3 seconds before reaching pre-event voltage. The reason behind this behavior could be explained by the fact that huge reactive power deviation in the beginning of the event is due to the high stalling current drawn by the motors. High current flowing in the line leads to higher possibility of thermal protection trip. As a result, motors disconnect from the network faster and the FIDVR duration decreases.

On the other hand, a fault at 15:51 PDT on August 12, 2012 caused voltage to dip to 35% of nominal, resulting in a severe FIDVR event. As demonstrated in Fig. 4-3, the slope of FIDVR increment is very low at the beginning. The FIDVR event duration was approximately 32 seconds until the voltage recovered to the pre-fault value (after the over-voltage clearance). The relation behind power deviation slope and FIDVR can be explained similar to the aforementioned events. Low active power and reactive power increment slope in the first milliseconds of the fault is due to the fact that the motors stalling current is not high or very few motors are stalled. These element-based reasons cannot be measured for FIDVR analysis since there are lots of uncertainties such as the number of A/C units, the manufacturing models of the machines, the settings of A/C units, initial conditions when stalling, accessibility of A/C units for doing the analysis, etc.

Therefore, the power system data (active power and reactive power values) can be used for FIDVR assessment.

Analyzing the FIDVR events occurred in SCE shows that slow power increment slope results in late thermal protection trip which is due to the inverse time-overcurrent characteristic of thermal protections explained in section 3.1.1. The lower the stalling current, the higher the tripping time.

To take advantage of this important feature analysis, four indices, active power deviation index (PDI), reactive power deviation indices (QDI), active power increment slope index (PISI) and reactive power increment slope index (QISI) are proposed and their correlation with FIDVR is investigated.

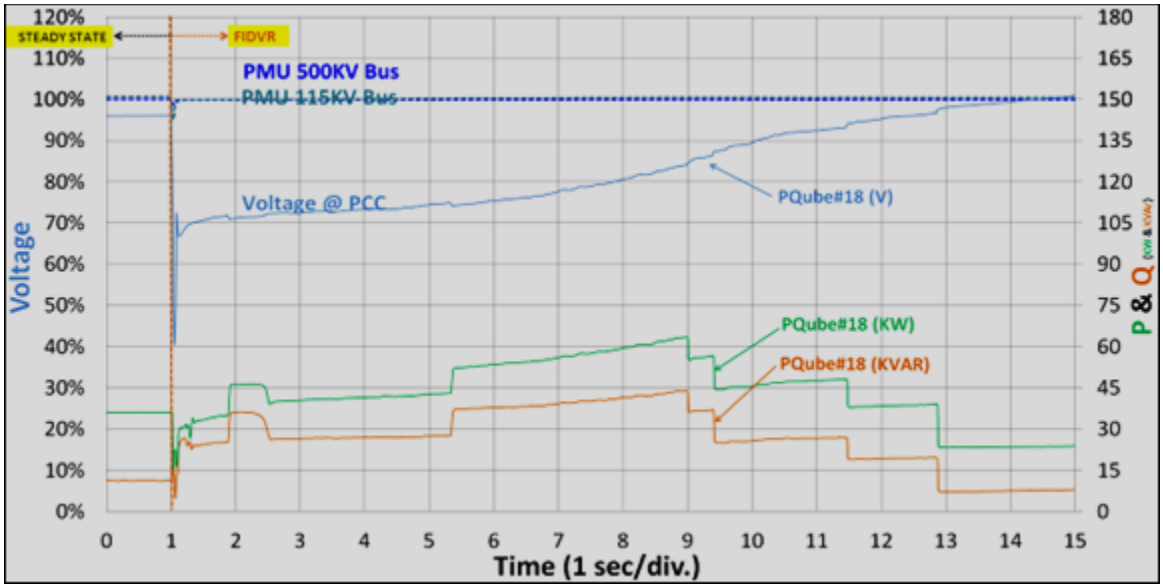

Fig. 4-1 RMS data of voltage, active power and reactive power of an FIDVR event occurred at 14:14 PDT on August 10, 2012 in the SCE network [58]

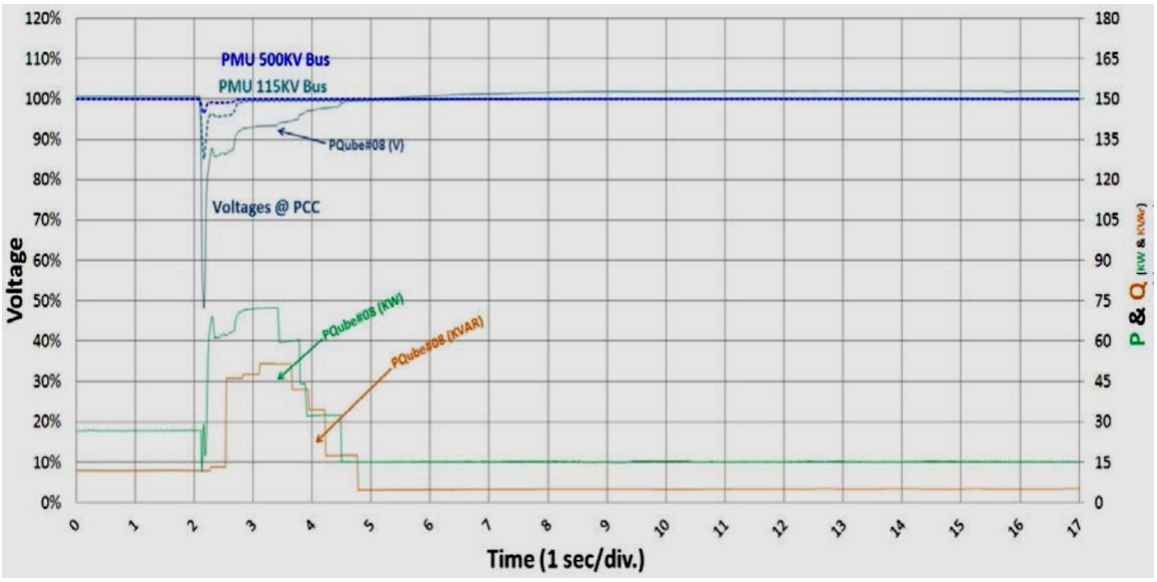

Fig. 4-2 RMS data of voltage, active power and reactive power of an FIDVR event occurred at 17:02 PDT on August 30, 2013 in the SCE network [59]

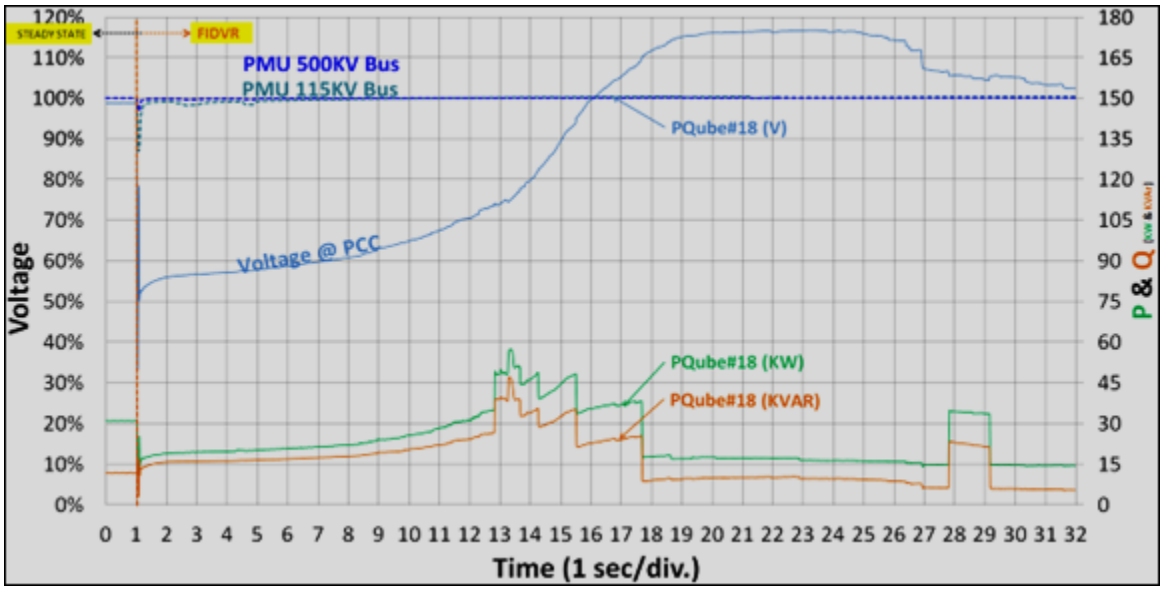

Fig. 4-3 RMS data of voltage, active power and reactive power of an FIDVR event occurred at 15:51 PDT on August 12, 2012 in the SCE network [58]

4.2.1.1. Active power deviation Index (PDI)

According to the analysis in section 4.2.1, active power deviates by increase in stalling current of the motors. Active power deviation index (PDI) is defined as a numerical index to quantify the active power deviation due to the FIDVR event. PDI is defined as a continuous index and can be obtained as follows:

$$\begin{cases} PDI_t = \dfrac{P_t - P_{t_f}}{P_{t_f}}, if \ \left|\dfrac{P_t - P_{t_{pre-f}}}{P_{t_{pre-f}}}\right| > \delta \\ 0, \qquad otherwise \end{cases} \quad \forall\, t \in [t_f, T] \tag{4-1}$$

where P_t is the active power at time *t* and P_{t_f} denotes the pre-fault active power magnitude and T is the considered voltage recovery time frame. $\delta = 20\%$ is adopted in this study. The value of δ can be adjusted depending on the practical problem requirement, but a different δ value will not affect the effectiveness of the proposed FIDVR assessment method.

To illustrate the relation of PDI_t and FIDVR, the average of PDI (APDI) is defined as follows:

$$APDI = \frac{\sum_{t=t_f}^{T} PDI_t}{N} \quad \forall\, t \in [t_f, T] \tag{4-2}$$

where *N* is the number of time steps from t_f to *T*. Linear regression, second and third order polynomial regression are used to demonstrate the relation of APDI and FIDVR with the result shown in Fig. 4-4. As it can be seen, the correlation is poor.

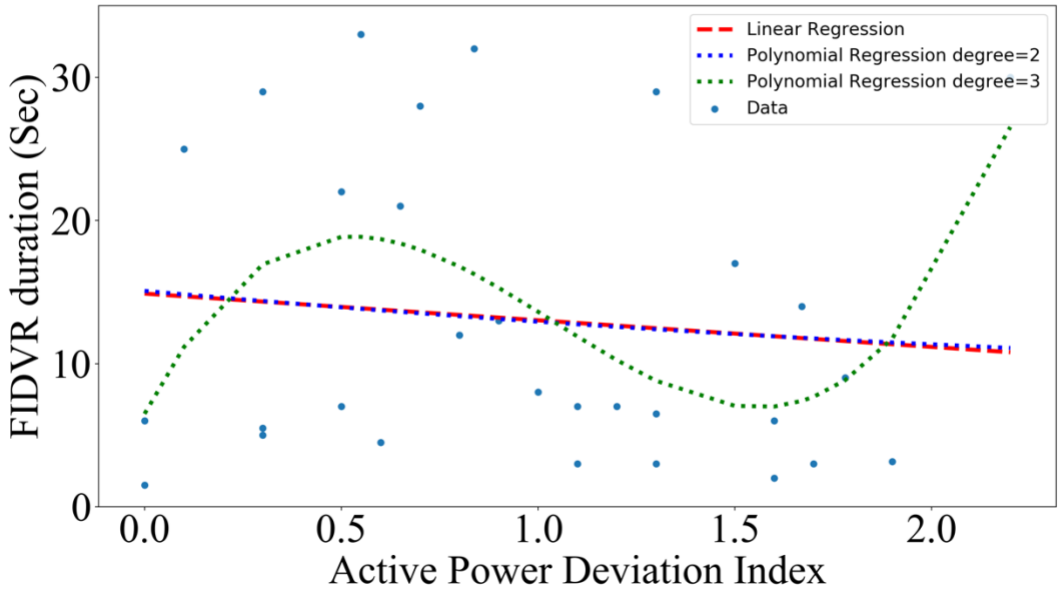

Fig. 4-4. The relation of the FIDVR duration and average active power deviation index using polynomial regression

4.2.1.2. Reactive power deviation Index (QDI)

Stalling induction motors absorb huge amount of reactive power from the grid. Reactive power deviation index (QDI) is defined as a numerical index to quantify the reactive power deviation due to the FIDVR event. QDI is defined as a continuous index and can be obtained as follows:

$$\begin{cases} QDI_t = \dfrac{Q_t - Q_{t_f}}{Q_{t_f}}, if \ \left|\dfrac{Q_t - Q_{t_{pre-f}}}{Q_{t_{pre-f}}}\right| > \delta \\ 0, \qquad otherwise \end{cases} \quad \forall\ t\ \in [t_f, T] \tag{4-3}$$

where Q_t is the active power at time t and Q_{t_f} denotes the pre-fault active power magnitude and T is the considered voltage recovery time. Similarly, $\delta = 20\%$ is adopted for QDI.

The average of QDI_t (AQDI) is defined as follows:

$$AQDI = \frac{\sum_{t=t_f}^{T} PQI_t}{N} \qquad \forall\, t \in [t_f, T] \tag{4-4}$$

To illustrate the relation of AQDI and FIDVR, linear regression, second and third order polynomial regression are used. Corresponding is the result shown in Fig. 4-5. As it can be seen, the correlation of QDI and FIDVR is better than FDI and FIDVR. However, the variance of the data shows that this feature could not be an appropriate feature for FIDVR analysis.

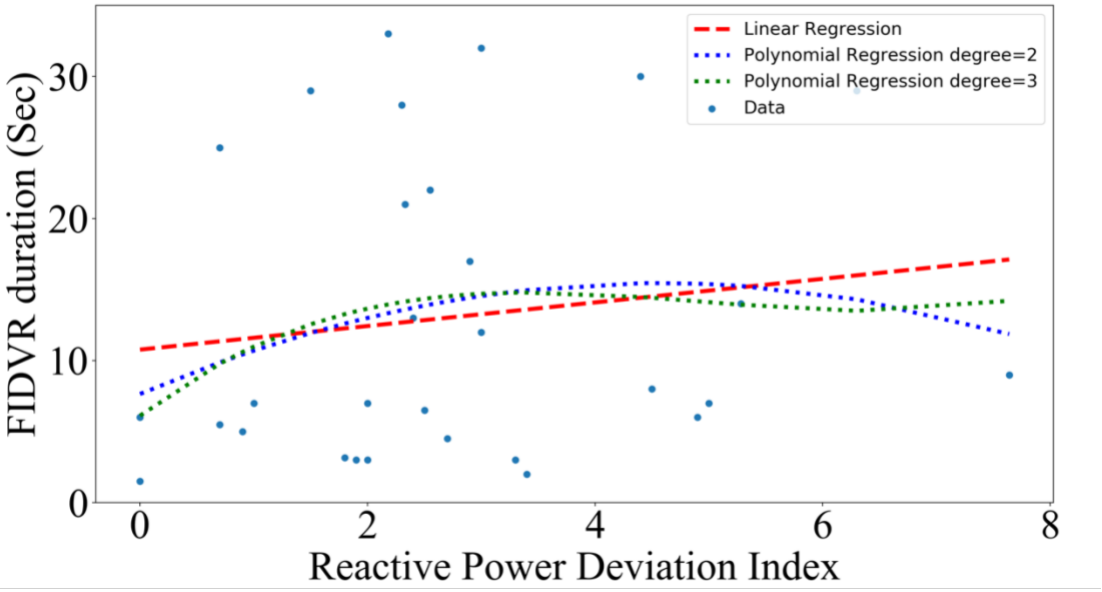

Fig. 4-5. The relation of the FIDVR duration and reactive power deviation index using polynomial regression

4.2.1.3. Active Power Increment Slope Index (PISI)

As explained in 4.2.1.1, PDI was not successful in representing a proper correlation with FIDVR. Hence, a new numerical index, called active power increment slope index (PISI), is defined to help the algorithm in better assessing FIDVR.

$$PISI_t = \frac{PDI}{V_t \times (t - t_f)} \qquad \forall\, t \in [t_f, T] \tag{4-5}$$

The average of $PISI_t$ (APISI) can be calculated as:

$$APISI = \frac{\sum_{t=t_f}^{T} PISI_t}{N} \quad \forall\, t \in [t_f, T] \tag{4-6}$$

To better demonstrate the relation of APSIS and FIDVR, the result of linear regression, second and third order of polynomial regression are shown in Fig. 4-6. According to Fig. 4-6, PISI and FIDVR has inverse relation with each other.

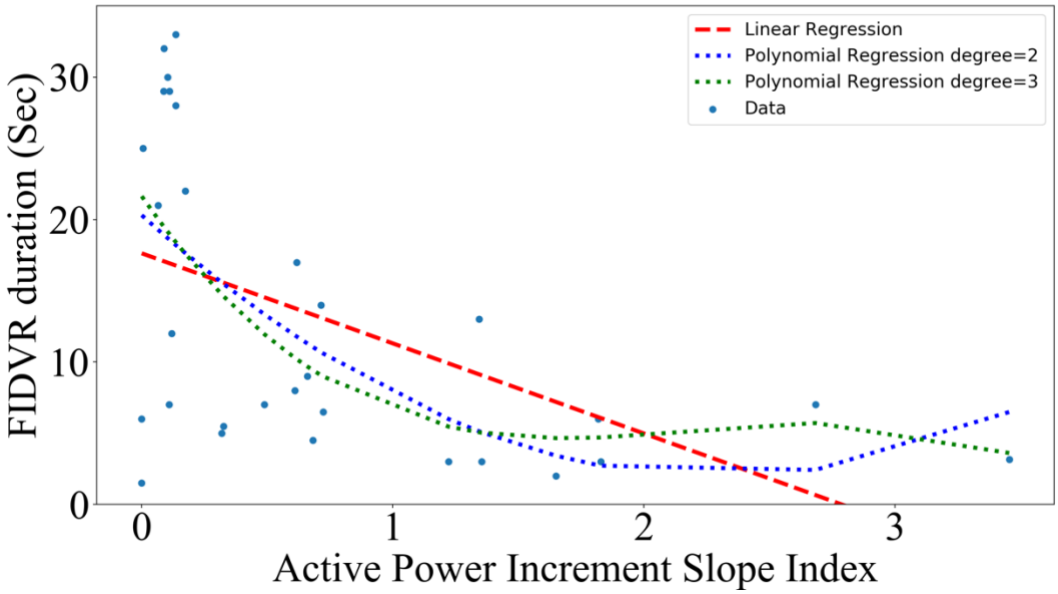

Fig. 4-6. The relation of the FIDVR duration and active power increment slope index using polynomial regression

4.2.1.4. Reactive Power Increment Slope Index (QISI)

As explained in 4.2.1.2, QDI was not successful in having a proper correlation with FIDVR. Hence, a new numerical index, called reactive power increment slope index (QISI), is defined to help the algorithm in better assessing FIDVR.

$$QISI_t = \frac{QDI}{V_t \times (t - t_f)} \qquad \forall\, t \in [t_f, T] \tag{4-7}$$

The average of $QISI_t$ (AQISI) can be calculated as:

$$AQISI = \frac{\sum_{t=t_f}^{T} QISI_t}{N} \qquad \forall\, t \in [t_f, T] \tag{4-8}$$

To better demonstrate the relation of AQISI and FIDVR, the result of linear regression, second and third order of polynomial regression are shown in Fig. 4-7. According to Fig. 4-7, QISI and FIDVR has inverse relation with each other.

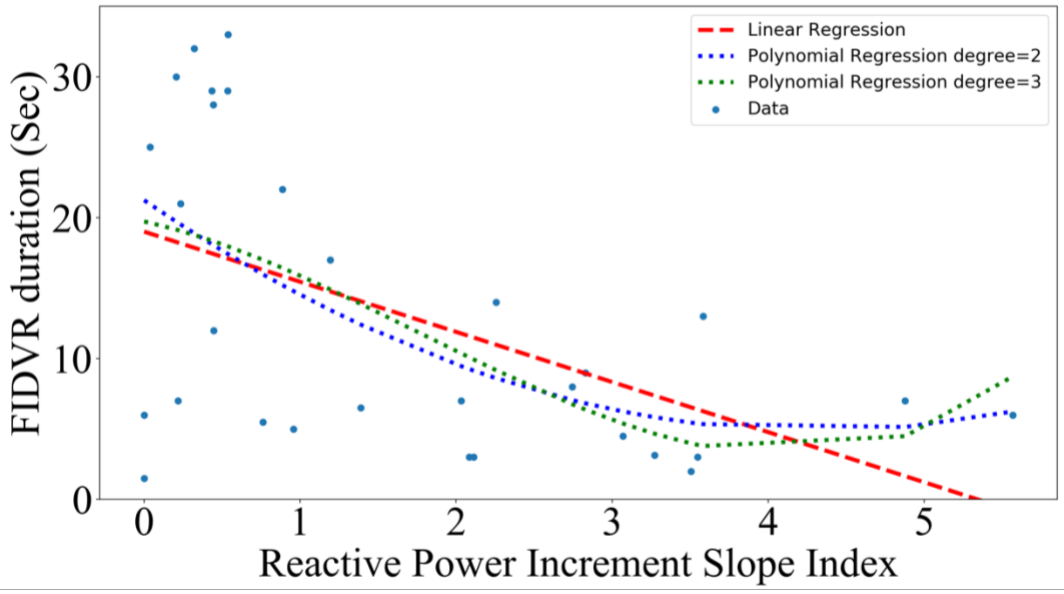

Fig. 4-7. The relation of the FIDVR duration and reactive power increment slope index using polynomial regression

4.2.2. Proposed Feature Sensitivity Analysis

As it has been mentioned in section 3.3.2, a straightforward data visualization can highly help in having valuable correlation analysis between the selected features and FIDVR. The FIDVR data in this section is extracted from a comprehensive literature review on the recorded data in SCE PQubes.

4.2.2.1. HeatMap

Heatmaps are used to show the data dependency on two or more variables as a color-coded image plot. As it can be seen in Fig. 4-8, the correlation of PDI and QDI with FIDVRD is -0.11 and 0.14 which is very low. On the other hand, APISI and AQISI have high correlation with FIDVRD which shows that the proposed feature can be effective in assessing the FIDVR.

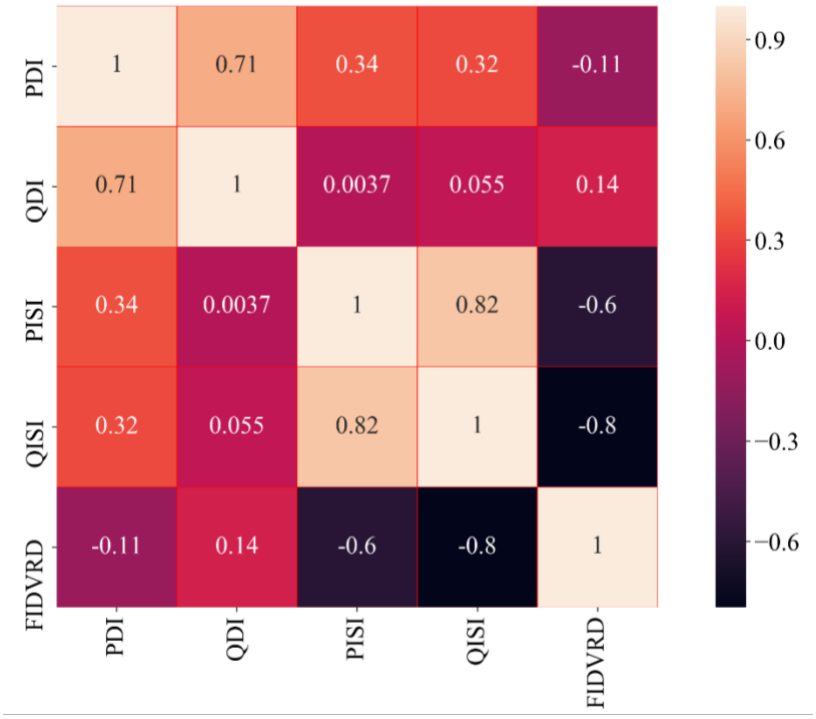

Fig. 4-8. Heat Map of the dataset which demonstrates the correlation between selected features and FIDVRD

4.2.2.2. PairGrids

As explained in section 3.2.2.2, pair plots use different pair of variables for each subplot and forms a matrix of sub-plots. Pair grid is used to understand the best set of features to explain a relationship between more than three variables in the dataset.

The distributions and relationships in a dataset can be seen in Fig.4-9 where each variable of the dataset is mapped onto a column and row in a grid of multiple axes. The diagonal plots showcase the histograms which demonstrate the distribution of a single variable.

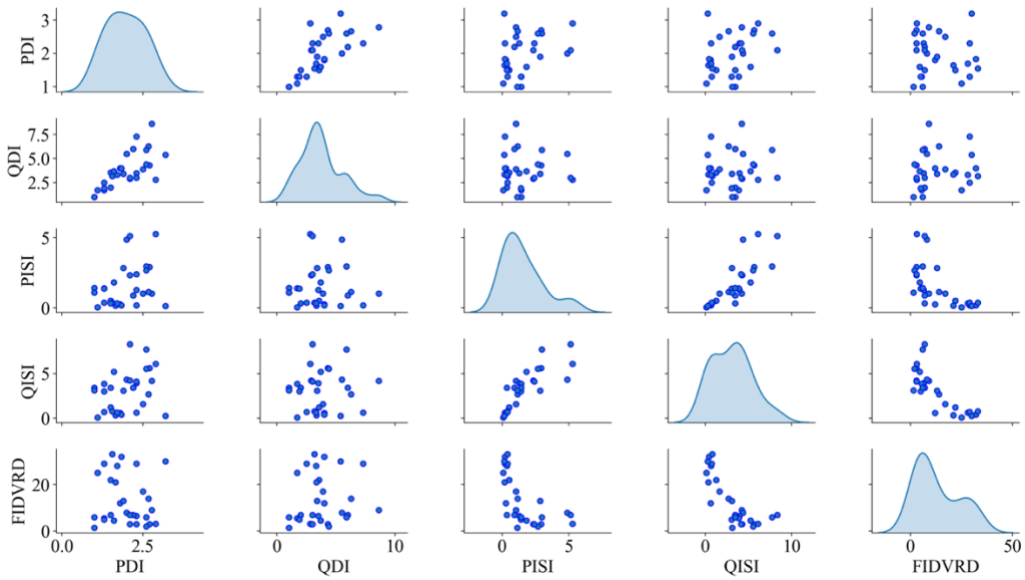

Fig. 4-9. Pair Grid plot of the selected features

More detailed pairplots are shown in Fig. 4-10 where the FIDVR cases which their duration are lower than 10 s, between 10 and 20 s and more than 20 s are colored blue, orange and green respectively. Suppose pair grid matrix is called PG. As it can be seen from element PG_{51} and PG_{52}, APDI and AQDI distribution do not provide valuable information for FIDVR analysis since approximately all three FIDVRD margins occur for APDI and AQDI values of 1-2.5 and 0-3, respectively. On the other hand, by analyzing PG_{53} and PG_{54} one can conclude that with lower APISI and AQISI, FIDVRD will be higher. On the top triangle, bivariate kernel density estimation of each two feature of the dataset has been demonstrated to visualize the distribution of the dataset. Blue, red and green kernel density estimations demonstrate FIDVR events of less than 10s, between 10 s and 20 s and more than 20 s, respectively. Elements of the 6^{th} column of PG demonstrate the kernel density estimation of each of the features and the FIDVR duration.

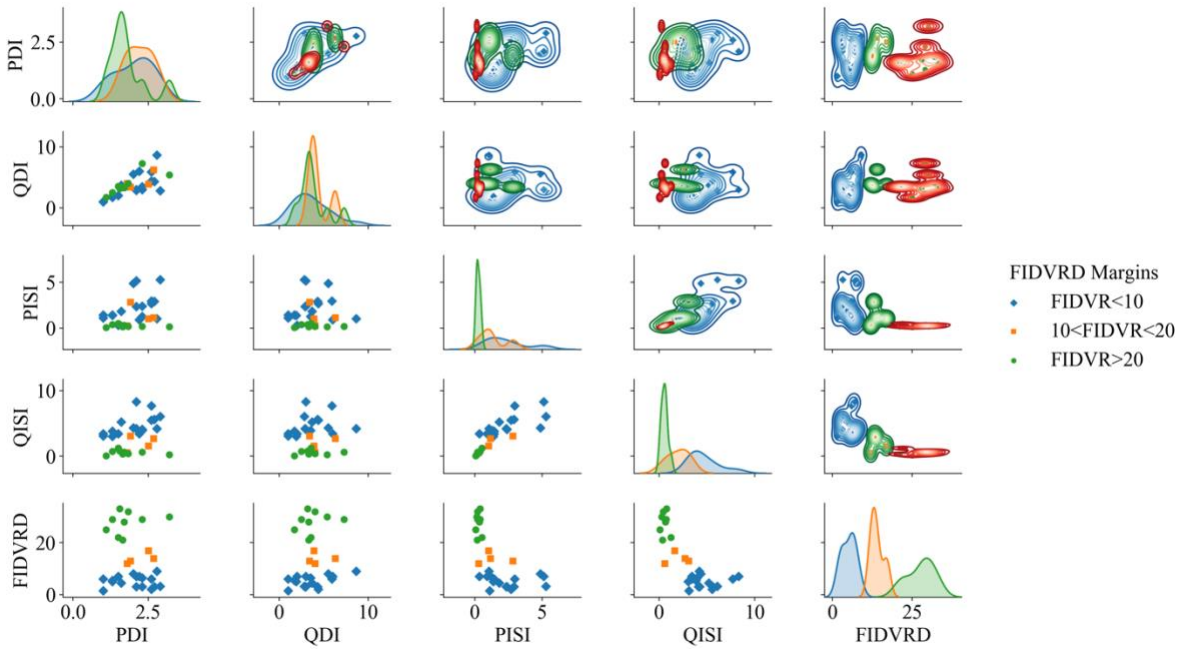

Fig. 4-10. Pair Grid plot of the features categorized according to the FIDVR duration. The higher triangle shows the bivariate kernel density estimation

4.2.3. Partitioning data

An issue when fitting a ML model is its performance behavior while applying it to a new data. To address this issue, the data set has been split into three partitions of test, train and validation in section 3.3.3. In this chapter, a powerful and popular resampling technique is used which is called *k*-fold validation technique.

4.2.3.1. *K*-Fold Cross Validation

K-fold validation is a useful method in ML which allows training and testing the model for *k* times on different subsets of data and generate the skill and performance estimation of the ML model on unseen data. The method has a parameter *k* that refers to the number of subsets that the data is split into. *K*-Folds Cross Validation splits the data into *k* different

subsets or folds while *k-1* subsets are selected for training the model and the last fold as test data. The model is trained on *k*-1 folds and after finalizing the model, the model is tested on the test set. This step is repeated for *k* times until all the *k* fold experience being a test set. Finally, the accuracy of the model can be calculated by taking average of all folds [63], [70], [62], [71].

The process of the algorithm is demonstrated in Fig. 4-11. The procedure is as follows:

1) The dataset is shuffled randomly
2) The dataset is split to *k* subsets
3) For each unique group take one subset as test set
4) For each group take the remaining subsets as training set
5) For each group fit the model on training set (for either random forest classifier (RFC) or random forest regressor (RFR))
6) For each group evaluate model on the test set (for either RFC or RFR)
7) For each group store the evaluation score and discard the model
8) Aggregate the evaluation scores of all the groups
9) Calculate the model's final validation score by taking average of validation score of all groups

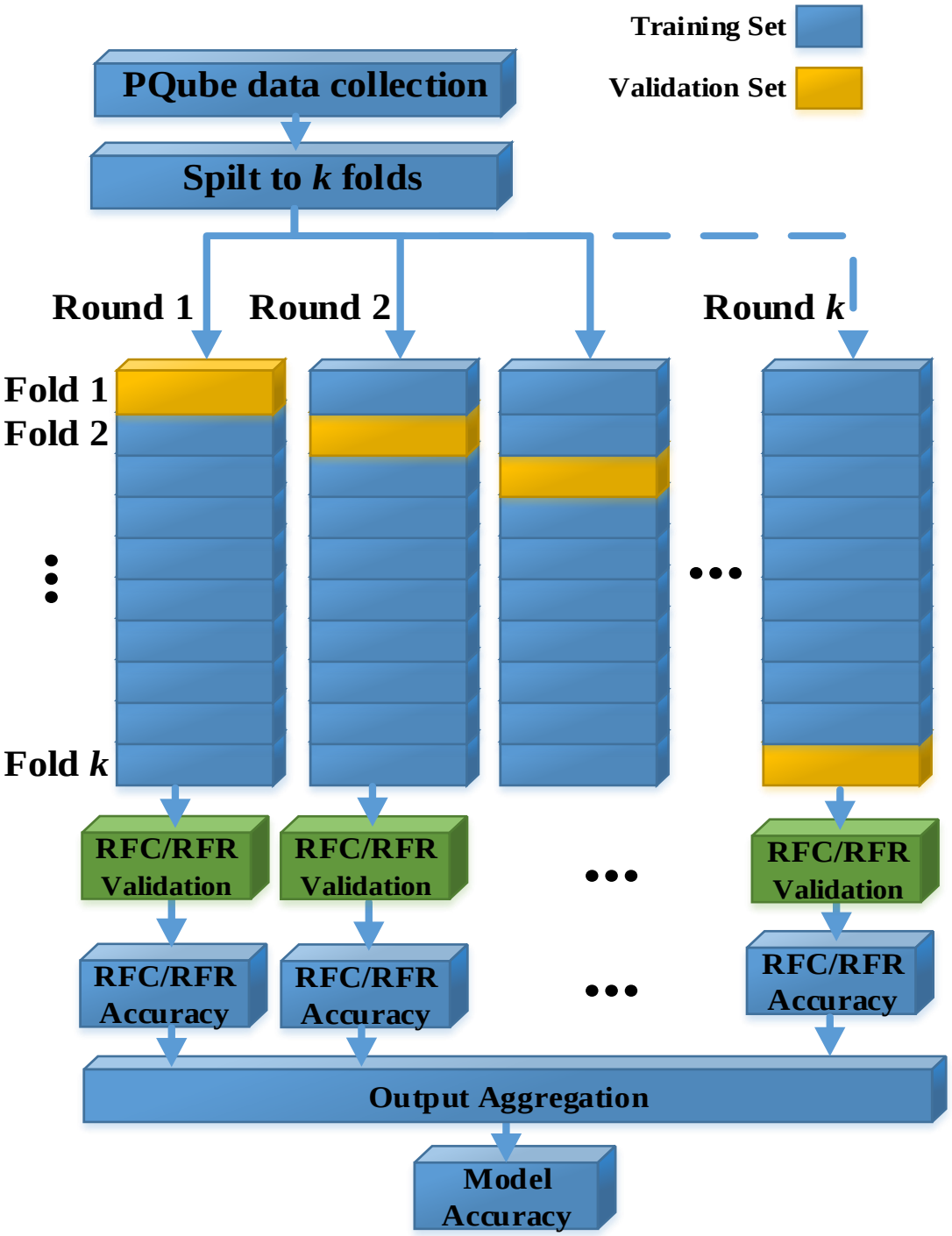

Fig. 4-11 Offline training and model creation using K-fold cross validation procedure

One of the most important tasks in *k*-fold cross validation is that each observation in the data sample should be assigned to an individual group and stays in that group for the duration of the procedure. This process gives the opportunity to each sample to be used in the test set 1 time and used to train the model *k*-1 times.

Some of the advantages of cross validation are:

1) It is a great out-of-sample estimator since it provides the opportunity for each sample to participate in test and train subsets.
2) The variance of the resulting estimate is reduced as *k* increases.
3) The algorithm generally results in a less biased (less optimistic) estimate of the model skill compared to train/test split method.

The only challenge of cross validation is that the computation time of training algorithm increases as the number of k folds increases since the algorithm has to be rerun from scratch k times.

A poorly chosen value for *k* may result in a mis-representative idea of the skill of the model, such as a score with a high variance which may change a lot based on the data used to fit the model, or a high bias such as an overestimate of the skill of the model. In this study, the value for k is fixed to 10, a value that has been found through experimentation to generally result in a model skill estimate with low bias a modest variance.

4.3. **Ensemble Learning**

Ensemble machine learning methods use multiple learning models so that final error from the aggregated output can be reduced and finally gain better predictive results. In the literature, ensemble learning has been applied to power system stability analysis. In [72] an artificial neural network ensemble is proposed for transient stability assessment and decision tree-based ensembles are developed in [73], [74], [75] for dynamic security assessment.

The proposed FIDVR assessment is explained in section 4.4. This method is a time-series data-driven decision making for real-time FIDVR prediction. The proposed technique needs separate classifier/regressors to be trained for each time step, hence its overall training burden can become higher than a single classifier/regressors. In addition, each of the classifier/regressors in the ensemble learning consists of a number of features and single learners. Therefore, the overall model training time can be increased which is not suitable for real world applications since it diminishes the reliability of the model for online or real-time applications. In this case, since the fast learning capability of random forest (RF) can significantly reduce the overall training burden, RF can be an ideal algorithm as an ensemble learning method for FIDVR analysis. In this study, each RF in an ensemble in each-time step is trained by a random portion of the training samples according to the proposed *k*-fold cross validation technique [76], [63], [70], [62].

4.3.1. Decision Tree-Based Ensemble Learning

4.3.1.1. Background

RFs, also known as random decision forests, are a popular ensemble method that can be used to build predictive models for both classification and regression problems. RF model creates an entire forest of random uncorrelated decision trees (DTs) to arrive at the best possible answer [76], [63].

DTs are simple models that employ top-down approach in which the root node splits the data until a certain criterion is met. The continuous splitting of nodes results in predicting values in final nodes based on the values of interior nodes. Although DTs are intuitive methods, they have some imitations that prevent them from being useful in ML applications. Some of these issues are their tendency to have high variance when they utilize different training and test sets of the same data which result in overfitting on training data. This leads to poor performance on unseen data which limits the usage of DTs in predictive modeling. A single DT typically has less predictive capability than an NNET or SVM [77], [78]. Another disturbing issue of the DT model is that it is not stable to small changes in the learning data since the selection of the splitting variable depends on the particular distribution of observations in the learning sample [73]. Thus, the entire tree structure could be altered if the first splitting variable is chosen differently due to a small change in the learning data and results in a high variability of DT predictions.

However, by using RF learning model of Breiman [79] as a DT-based ensemble method, models can be created that utilize underlying DTs as a foundation for producing powerful results. RFs choose a subsample of the feature space at each split which aims to make the

trees de-correlated and prune the trees by setting a stopping criterion for node splits. Random forests are not the only way to ensemble DT learning, but they have the great advantage of fast tuning with almost no user input, except for the number of trees in the ensemble [76], [63], [70], [62].

Fig. 4-12 demonstrates the RF algorithm.

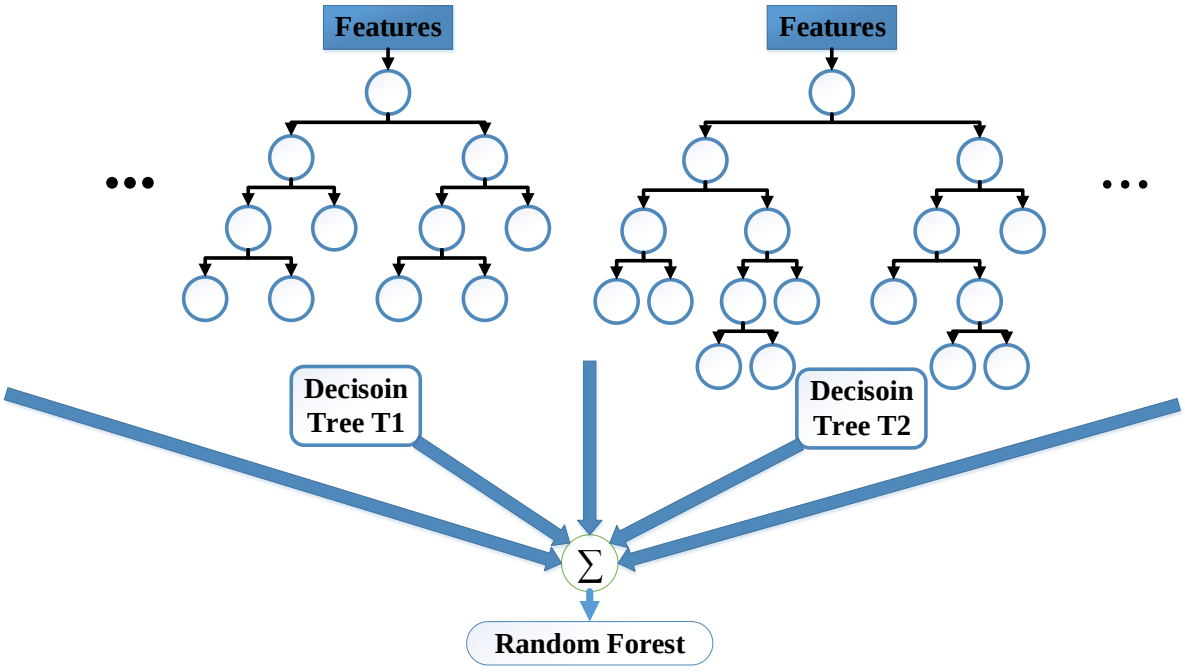

Fig. 4-12. Random Forest topology

4.3.1.2. Random Forest

Suppose that we are given a training sample $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ of independent and identically distributed (i.i.d.) $[0, 1]^d \times R$-valued random variables ($d \geq 2$) with the same distribution as an independent generic pair (X,Y) satisfying $EY^2 < \infty$. The space $[0, 1]^d$ is equipped with the standard Euclidean metric. For fixed $x \in [0,1]^d$, the goal is to estimate the regression function $r(x) = \mathbb{E}[Y|X = x]$ using the data D_n. In this respect, we

say that a regression function estimate r_n is consistent if $E[r_n(X) - r(X)]^2 \to 0$ as $n \to \infty$ [79], [80].

A random forest is a predictor consisting of a collection of randomized base regression trees $\{rn(x, \Theta m, Dn), m \geq 1\}$, where $\Theta 1, \Theta 2, \ldots$ are i.i.d. outputs of a randomizing variable Θ. These random trees are combined to form the aggregated regression estimate:

$$\bar{r}_n(X, D_n) = \mathbb{E}_\theta[r_n(X, \theta, D_n)] \tag{4-9}$$

where E_Θ denotes expectation with respect to the random parameter, conditionally on X and the data set D_n. Note that, in practice, the above expectation is evaluated by Monte Carlo, i.e., by generating M (usually large) random trees, and taking the average of the individual outcomes. The randomizing variable Θ is used to determine how the successive cuts are performed when building the individual trees, such as selection of the coordinate to split and position of the split.

In the model, the variable Θ is assumed to be independent of X and the training sample D_n which excludes in particular any bootstrapping or resampling step in the training set and rules out any data-dependent strategy to build the trees, such as searching for optimal splits by optimizing some criterion on the actual observations. However, Θ is allowed to be based on a second sample, independent of, but distributed as, D_n.

Each individual random tree is constructed as follows [79], [80]:

All nodes of the tree are associated with rectangular cells such that at each step of the construction of the tree, the collection of cells associated with the leaves of the tree forms a partition of $[0,1]^d$ and the root of the tree is $[0,1]^d$ itself. The following procedure is then repeated $\lfloor \log_2 k_n \rfloor$ times, where $k_n \geq 2$ is a deterministic parameter, fixed beforehand by the user, and possibly depending on n. At each node, a coordinate of $X = (X^{(1)},\ldots,X^{(d)})$ is

selected, with the *j*-th feature having a probability $p_{nj} \in (0, 1)$ of being selected. At each node, once the coordinate is selected, the split is at the mid-point of the chosen side.

Each randomized tree $r_n(X, \Theta)$ outputs the average over all Y_i for which the corresponding vectors X_i fall in the same cell of the random partition as X. In other words, letting $A_n(X,\Theta)$ be the rectangular cell of the random partition containing X [79], [80]:

$$r_n(X,\theta) = \frac{\sum_{i=1}^{n} Y_i 1_{[X_i \epsilon A_n(X,\theta)]}}{\sum_{i=1}^{n} 1_{[X_i \epsilon A_n(X,\theta)]}} 1\varepsilon_n(X,\theta) \tag{4-10}$$

where the event $\varepsilon_n(X,\theta)$ is defined by

$$\varepsilon_n(X,\theta) = \left[\sum_{i=1}^{n} 1_{[X_i \epsilon A_n(X,\theta)]} \neq 0\right] \tag{4-11}$$

Hence, by convention, the estimate is set to 0 on empty cells. Taking finally expectation with respect to the parameter Θ, the random forests regression estimate takes the form [79], [80]:

$$\bar{r_n}(X, D_n) = \mathbb{E}[r_n(X,\theta)] = \mathbb{E}\left[\frac{\sum_{i=1}^{n} Y_i 1_{[X_i \epsilon A_n(X,\theta)]}}{\sum_{i=1}^{n} 1_{[X_i \epsilon A_n(X,\theta)]}} 1\varepsilon_n(X,\theta)\right] \tag{4-12}$$

Note that, by construction, each individual tree has exactly $2^{[log_2 k_n]}$ ($\approx k_n$) terminal nodes, and each leaf has Lebesgue measure $2^{-[log_2 k_n]}$ ($\approx 1/k_n$). Thus, if $\boldsymbol{X}$ has uniform distribution on $[0,1]^d$, there will be on average about n/k_n observations per terminal node. In particular, the choice $k_n = n$ induces a very small number of cases in the final leaves, in accordance with the idea that the single trees should not be pruned. During the

construction of the tree, at each node, each candidate coordinates $X^{(j)}$ may be chosen with probability $p_{nj} \in (0, 1)$. This implies in particular $\sum_{j=1}^{d} p_{nj} = 1$ [79], [80].

To have better understanding about random forest classifier suppose our training data set is represented by D_n and suppose data set has M features (or attributes or variables). $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ and Xi is input vector $\{X_{i1}, X_{i2}, \ldots X_{iM}\}$ and y_i is the label or output. Suppose the number of trees in our forest is S then S datasets should be created from random resampling of data in D_n with-replacement (n times for each dataset). This will result in $\{T_1, T_2, \ldots T_S\}$ datasets. RF creates S trees and uses m $(=sqrt(M)$ or $=floor(ln_{M+1}))$ random sub-features out of M possible features to create any tree (random subspace method. So, for each T_i bootstrap dataset a tree K_i is created. To classify some input data $D = \{x_1, x_2, \ldots, x_M\}$, it pass through each tree and produce S outputs (one for each tree) which can be denoted by $Y = \{y_1, y_2, \ldots, y_s\}$. Final prediction is a majority vote on this set [79], [80].

4.3.1.3. Random Forest Tuning

RF has nearly the same hyper-parameters as a decision tree or a bagging classifier. RF adds additional randomness to the model, while growing the trees. RF searches for the best feature among a random subset of features, instead of searching for the most important feature while splitting a node. Adding randomness to the model lead to a wide diversity that generally results in a better model. Hence, in RF, only a random subset of the features is taken into consideration by the algorithm for splitting a node. Deep decision trees might suffer from overfitting, while RF prevents overfitting most of the time, by creating random subsets of the features and building smaller trees using these subsets. Afterwards, it combines the subtrees [63], [70], [62], [79] , [80].

The Hyperparameters in random forest are either used to increase the predictive strength of the model or to make the model faster. Some of the important hyperparameters of *sklearns* built-in RFs are as follows [79] , [80]:

- **Number of estimators**

This hyperparameters is the number of trees the algorithm builds before taking the maximum voting or taking averages of predictions. Higher number of trees increases the performance of the model which makes the predictions more stable, however it also slows down the computation.

- **Max depth**

Max depth represents the depth of each tree in the forest. The deeper the tree, the more splits it has which captures more information about the data.

- **Max features**

This hyperparameters is the maximum number of features RF considers to split a node. The number of features to consider when looking for the best split are as follows:

- If int, then consider max_features features at each split.
- If float, then max_features is a fraction and int(max_features × n_features) features are considered at each split.
- If "sqrt", then max_features=sqrt(n_features) (same as "auto").
- If "log2", then max_features=log2(n_features).

- **Min Sample Leaf**

It determines the minimum required number of leaves to split an internal node which is similar to min samples splits, except this describe the minimum number of samples at the leaves.

- **N jobs**

This parameter informs the engine the number of processors it is allowed to use. The value of 1 indicates that it can only use one processor, while a value of -1 means that there is no limit which increases the Model's Speed.

- **Random state**

This hyperparameters makes the model's output replicable.

- **Out Of Bag Score**

It is a RF cross validation method. In this sampling, about one-third of the data is not used to train the model and can be used to evaluate its performance which are called the out of bag (OOB) samples. OOB samples or OOB error is very similar to the leave-one-out cross-validation method, however almost no additional computational burden goes along with it.

- **Min samples split**

Min samples split represents the minimum number of samples required to split an internal node which can vary between considering at least one sample at each node to considering all of the samples at each node. When we increase this parameter, each tree in the forest becomes more constrained as it has to consider more samples at each node.

One of the big issues in ML is overfitting. However, most of the times, RFs are immune to overfitting since there are enough trees in the forest to protect the classifier/regressor from overfitting.

4.4. **Proposed RFC/RFR based decision making model**

In the literature, the ML methods use a pre-selected input vectors to classify/predict the system's stability status [50], [52], [53], [51], [54], [55]. To decide whether an event can be categorized as FIDVR or not or to predict the duration of FIDVR, a basic ML method can implement the assessment on real-time system data of a specific time after the fault and the final decision can be made by simply using this data. However, since the system has transient behavior after the fault and during the FIDVR, if the event is considered as FIDVR, it is unacceptable to only use the data of one glance. If the transient behavior of the system after the prediction analysis causes the model to categorize an event as an FIDVR fallaciously and/or predict a higher/lower FIDVR duration, subsequent emergency control methods may not be triggered in the appropriate time and the system operator fails to ensure the stability of the system. However, the error existing in ML prediction process cannot be properly captured and controlled, therefore by having higher assessment accuracy, system stability can be guaranteed.

To improve the decision-making accuracy and speed, this study proposes a probabilistic time-series FIDVR assessment method which progressively and simultaneously decides the possibility of occurring of FIDVR event and predicts the FIDVR duration, if the event categorized as FIDVR. Before explaining the real-time prediction algorithm, note that to have a real-time assessment, the proposed random forest classification (RFC) and random forest regression (RFR) models are trained for each time window according to the proposed *k*-fold cross validation technique proposed in section 4.2.3. The training is

accomplished based on the input feature values of each time step captured by event recorder. The features are weather temperature, voltage trajectory, pre-fault reactive power and QISI, where pre-fault reactive power is a fixed value during the analysis and other features change during the real-time FIDVR assessment. Once the RFR and RFC models are built, they can be used in real-time prediction.

The proposed FIDVR assessment method is illustrated in Fig. 4-13 where the curves at the top represent the post-fault voltage values collected in real-time from the installed meters and t_f is the fault time. $t_f,\ t_f + \Delta t, t_f + 2\Delta t, \dots, t_f + N\Delta t$ are the time points when the FIDVR assessment action is progressively executed, and $t_f + N\Delta t$ is the latest decision time after which the load shedding actions must be taken to be effective. $T=N\Delta t$ is the considered voltage recovery time explained in section 4.2.1. Based on the N time points, N sliding time-windows are defined as the time period between each two consecutive time points. The process is implemented as follows:

At each time step, snapshots of selected features are collected within the time-step (Δt) between current and the previous time points, and the voltage magnitudes are transformed into voltage deviation and PISI is calculated. The principle is to assess the FIDVR event and predict the FIDVR duration simultaneously in each time step. To do so, the features are given to RFC to probabilistically decide whether the event can be categorized as FIDVR or not. The prediction result comes in the form of the binary value which categorizes the event either as FIDVR or non-FIDVR. In the proposed mode, the probabilistic prediction not only predicts the possibility of FIDVR happening, but also predicts the FIDVR duration. To do so, if the event categorized as FIDVR, the data will be given to RFR to predict the FIDVR duration based on the features in the snapshot. If

the event is not considered as FIDVR by RFC, the decision will be saved, and the probabilistic prediction will be progressed at the next time point with the updated feature snapshots. In practice, a maximum decision time $t_f + N\Delta t$ is needed, at which time the FIDVR assessment decision must be delivered. Since N time windows are defined, N probabilistic predictors (RFC and RFR) should be prepared respectively for different time windows. The final FIDVR assessment decision is made once sufficient decision probability criterion is achieved. Using the proposed method, the assessment decision can be reliably delivered as early as possible without impairing the accuracy. Thus, overall decision-making speed will be significantly improved.

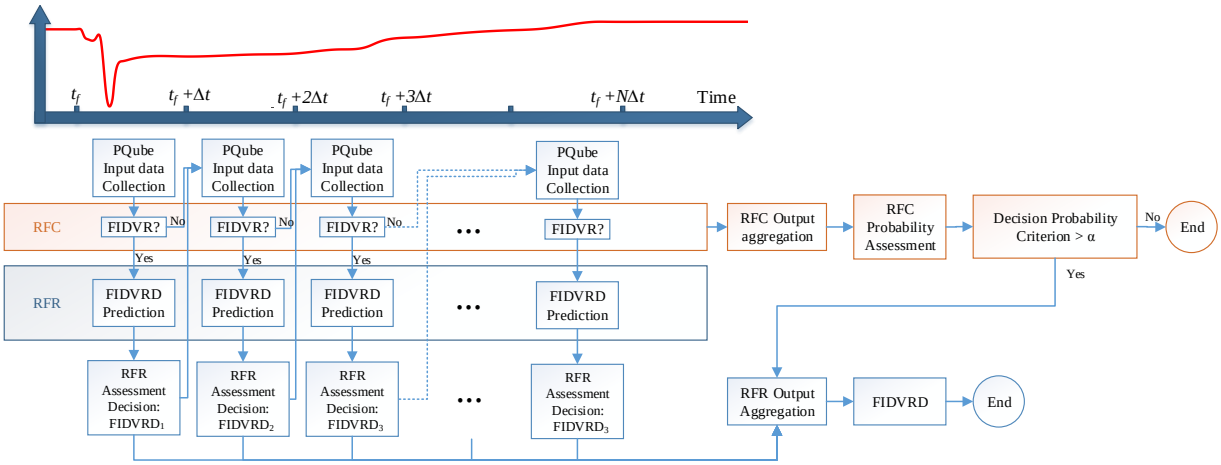

Fig. 4-13. Real-time application of the proposed RFC/RFR predictor

4.5. **Numerical Studies**

The proposed model is built and tested in SCE valley network [58] which appears to be a susceptible network to FIDVR events. In the first part of numerical studies real FIDVR data is used collected from a comprehensive literature review on SCE valley network [38].

To investigate the detailed characteristics of FIDVR events in distribution circuits, SCE installed 22 power quality meters (PQubes) on 17 of its Valley Substation's 24 sub-transmission circuits that serve the utility's residential and commercial customers. By recording these data, lots of information can be obtained such as the understanding of how FIDVR events evolve and impact local residential and commercial customers [58]-[59]-[38].

In the second part of numerical study, several FIDVR events are created using WECC composite load model, air conditioner load model and thermal protection model as well as random load disconnection model. By modeling the system in EMTP-RV, several FIDVR events created and used for investigating FIDVR duration prediction.

4.5.1. Numerical Studies Using SCE Network

The proposed model built using real FIDVR data collected from SCE valley network [58]-[59]-[38]. The mentioned features in section 4.4. have been used for training the ML algorithm. The goal of real-time investigating FIDVR is to employ appropriate emergency control approach when it is needed. Hence, having a precise FIDVR assessment is necessary for consecutive emergency control.

4.5.1.2. Test System

As shown in Fig. 4-14, the valley network consists of a transmission system which contains of two 115 kV busses (section A&B and section C&D). Each of the 115 kV substation busses feeds a meshed sub-transmission system.

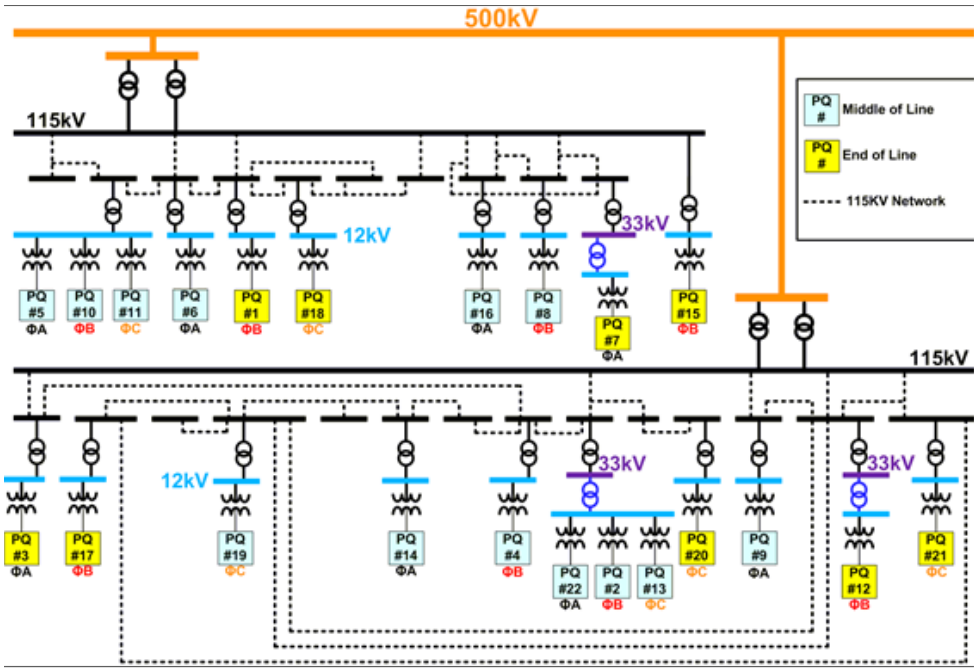

Fig. 4-14. Valley distribution system and PQubes Locations Diagram [38]

There are 24 meshed sub-transmission 115 kV substations in the network. Sub-transmission 115 kV substations contain two types of distribution circuits 33 kV and 12 kV, most of which are 12 kV. 12 kV distribution circuits supply both commercial and residential circuits with pad-mount and pole-mount transformers to serve customers. Longer distribution circuits are mainly rural which use 33 kV circuits instead of the 12 kV distribution circuits. The PQube devices were installed in the pad-mount transformer's secondary (240 V) side supplying customers. For each distribution circuit, the device installations were located either at the middle or the end of the line. These data recording devices were also placed on different phases of the circuits to acquire a diverse collection of event data.

4.5.1.3. Offline and Real-time train/test procedure

The power quality recording devices (PQubes) installed in valley's network can record up to five voltages and five currents during steady-state conditions as well as during system events. These devices were programmed to record both root mean square (RMS) and sinusoidal waveforms when an event is triggered. RMS event data captured at 1 sample/cycle and sinusoidal waveform event data captured at 32 samples/cycle. Every device was equipped with an uninterruptible power supply (UPS) for up to nine minutes so that will record during events of low voltage without compromising the data. Each SCE PQube contains components such as PQube module, current module, power supply, circuit breaker, current transformers, din rail and enclosure [38].

Based on the 1 sample/cycle (0.01667 s step size) of PQubes data sampling, the time window width Δt is set to 0.1 s, so each time window consists of 6 points of each feature. As mentioned, the features are voltage deviation and QISI (pre-fault reactive power is a fixed number feature). Hence, in each time window, the number of training features is 13 for each busbar (2 features×6 trajectory points+1 pre-fault reactive power). The latest decision time $N\Delta t$ is set to 1 s. Thus, N = 10 RFR and RFC probabilistic predictors are needed.

The computational time efficiency of the 10 probabilistic predictors is listed in Table 4-1 where all the tests are performed on a Laptop Core i5 with 2.6 GHz CPU and 4 GB RAM. Benefiting from the fast learning speed of RFR and RFC, the total offline training time of 10 probabilistic predictors is only 563 s, although the computation process in each time frame includes RFR and RFC of 620 trees using 10-fold validation for each RFR/RFC. For online testing, the computational performance requirement for each

RFR/RFC is that the computation time consumed by each probabilistic predictor must be shorter than the time window width Δt to make sure there is no overlap between two successive predictions. In Table 4-1, the average and the longest computation time of the probabilistic prediction on a trajectory snapshot are both listed. It can be seen that the longest testing time 12.31 *ms* is much shorter than Δt (100 *ms*), meaning the proposed probabilistic prediction method is fully compatible with the proposed time-series application.

Table 4-1
FIDVR Assessment Time Efficiency

	Offline Time		**Real-Time**	
RFC/RFR Offline Training Time (s)	Average Aggregation Performance (10-fold) Validation Time (s)	Total Average offline Computation Time (s)	Average Testing Time on a Trajectory Snapshots (ms)	Longest Testing Time on a Trajectory Snapshot (ms)
449.39	63.644	563.034	11.22	12.31

Following the training process in Fig. 4-13, RFR and RFC ensembles are simultaneously trained as a classifier to select the event as FIDVR and a FIDVR duration predictor using the features in each time window, respectively. The number of trees in each ensemble RFC/RFR is set to 620. For the probabilistic prediction performed at each time point, the number of estimators (trees) of the RFRs is separately tuned, and the tuning result is shown in Fig. 4-15 where the different colors refer to the tuning results for different time points. The number of estimators' range with the lowest testing MSE is empirically selected and as it can been, the RFR trained for a later time point result in a lower MSE. Such phenomenon coincides with the initial thought, i.e., higher prediction accuracy is at the cost of slower decision speed.

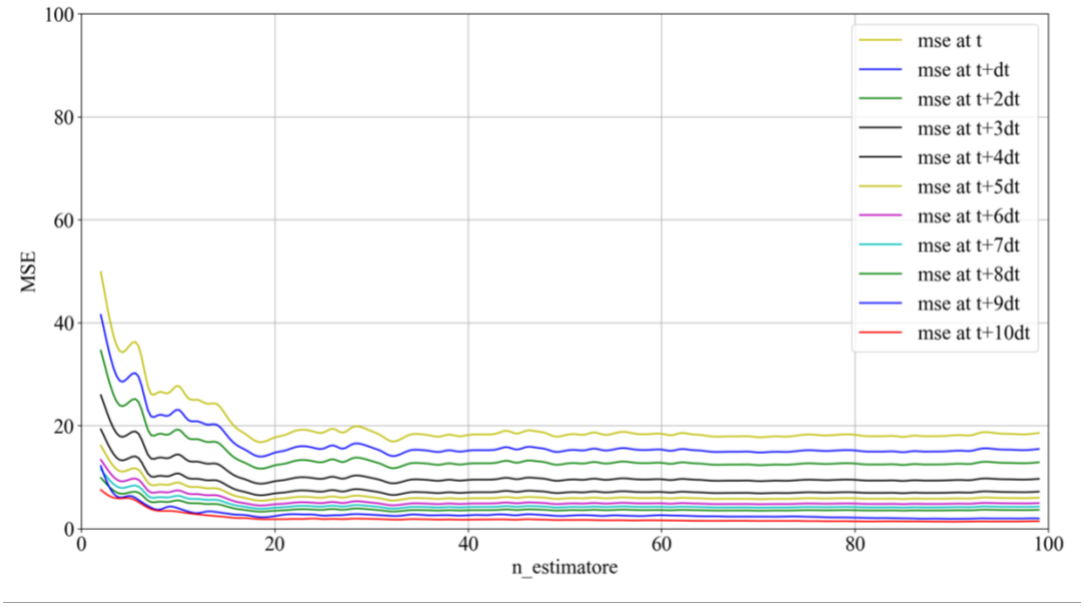

Fig. 4-15. Number of estimators tuning results for ten time-steps

The proposed simultaneous RFR/RFC based FIDVR assessment method is applied on the testing samples to verify its performance. For the probabilistic prediction performed, the accuracy of the *k*-fold RFC/RFR for each time point and also the average accuracy is calculated and shown in Table 4-2. Score provides the coefficient of determination for the trained model on the given data. By analyzing the data used for training, the training score can be obtained, while by analyzing unseen test-data, the validation accuracy can be calculated. Note that the results of Table 4-2 is based on 620 number of estimator. As it can be seen, in the first time-step, the test result of RFR/RFC are below 99%. The results of 5^{th} time step -10^{th} time step demonstrates high efficiency of the model. For most time points, the misclassification rates achieve below 1%. The overall misclassification rate is as low as 0.987%, demonstrating high accuracy in FIDVR detection. Such short assessment time in both RFR and RFC verifies the exceptional early assessment capability of the proposed method. Another aspect of the simultaneous RFR/RFC assessment

performance to be considered is how early the unstable samples can be successfully detected even in the first five time-steps.

Table 4-2
RFR/RFC Testing and Testing Results

		Accuracy										Average Accuracy
	Time points	1	2	3	4	5	6	7	8	9	10	
RFR	Training accuracy	0.991	0.921	0.993	0.992	0.9929	0.9956	0.9967	0.9955	0.996	0.9972	0.9942
	Validation accuracy	0.989	0.988	0.99	0.989	0.9854	0.991	0.9926	0.9919	0.9921	0.992	0.9901
RFC	Training accuracy	0.989	0.9899	0.993	0.942	0.9974	0.9959	0.9964	0.9959	0.9971	0.9962	0.9945
	Validation accuracy	0.977	0.979	0.984	0.989	0.996	0.994	0.9956	0.9949	0.9962	0.9954	0.99013

To obtain the relation of number of estimators and average of training and validation accuracy of RFR, the number of estimators (trees) of the RFRs are separately tuned, and the tuning result is shown in Fig. 4-16 where the accuracy score is average score of all time points. The blue curve and green curve are respectively the average training and validation score (R squared) of the proposed prediction method for different value of number of trees. It can be seen that the best result is obtained when the number of estimators is equal 620.

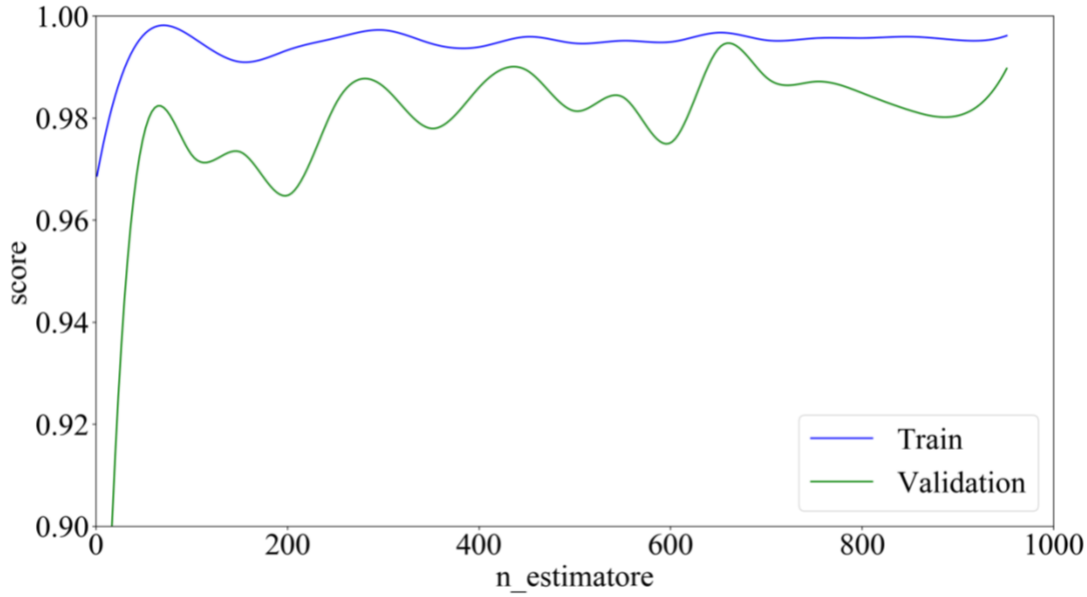

Fig. 4-16. Predictor's training and validation average score based on tuning the number of estimators

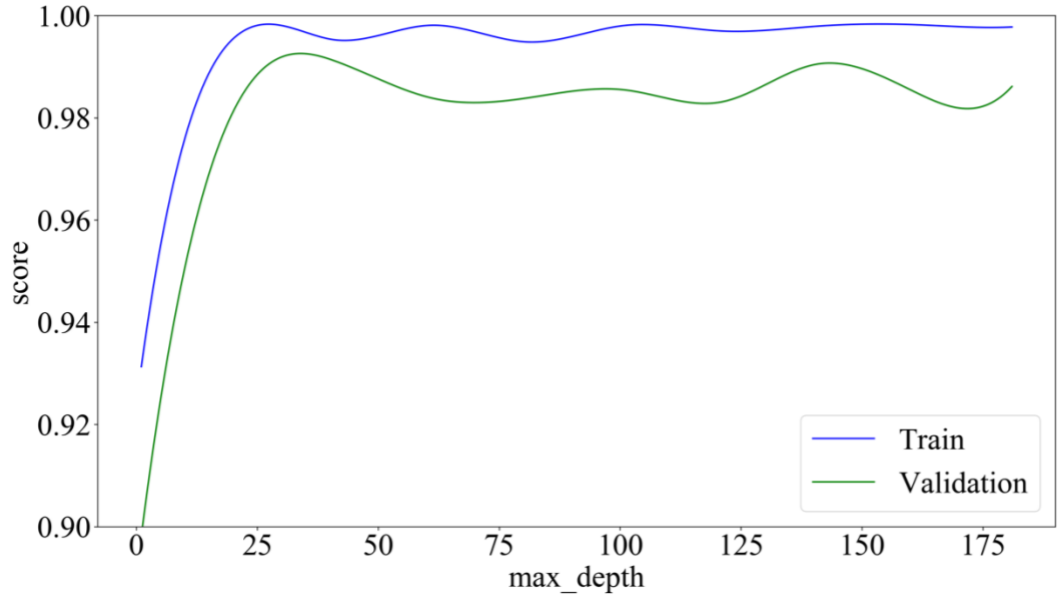

Fig. 4-17. Predictor's training and validation average score based on tuning the max depth

The same analysis is accomplished by changing the max depth of RFRs while number of estimators are set to 620. It can be seen from Fig. 4-17 that the best value for max depth is 27.

Figs. 4-18 and 4-19 demonstrate the average OOB error at the addition of each new tree during training of RFRs and RFCs, respectively. The resulting plot allows a practitioner to approximate a suitable value of number of estimators at which the error stabilizes. The RFR/RFC is trained using bootstrap aggregation, where each new tree is fit from a bootstrap sample of the training observations. The OOB error is the average error for each calculated using predictions from the trees that do not contain in their respective bootstrap sample. This allows the RFC to be fit and validated whilst being trained [81]. OOB error provides the coefficient of determination using OOB method.

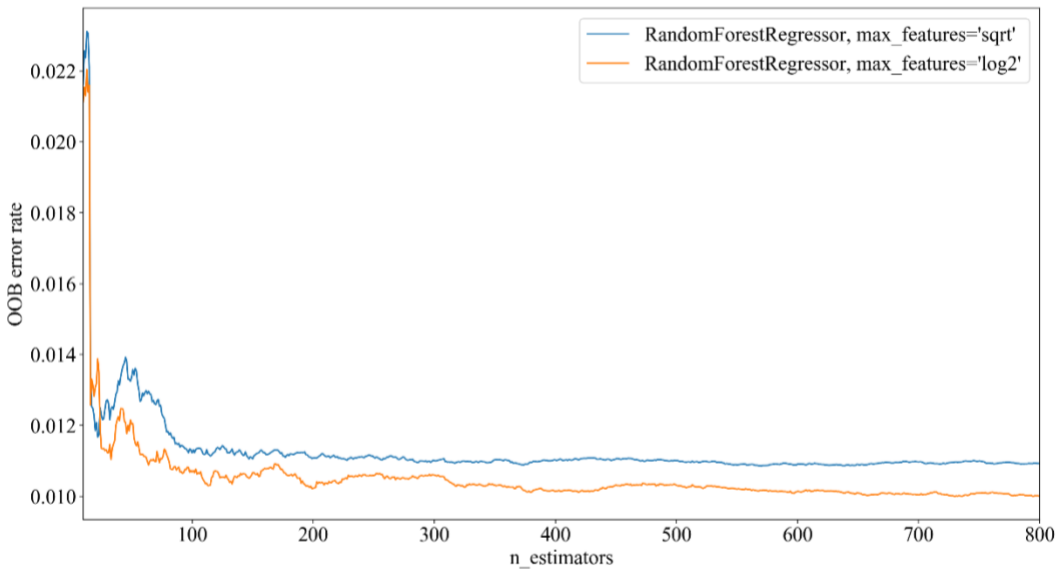

Fig. 4-18**.** RFR Predictor's training and validation average OOB rate based on tuning the number of estimators

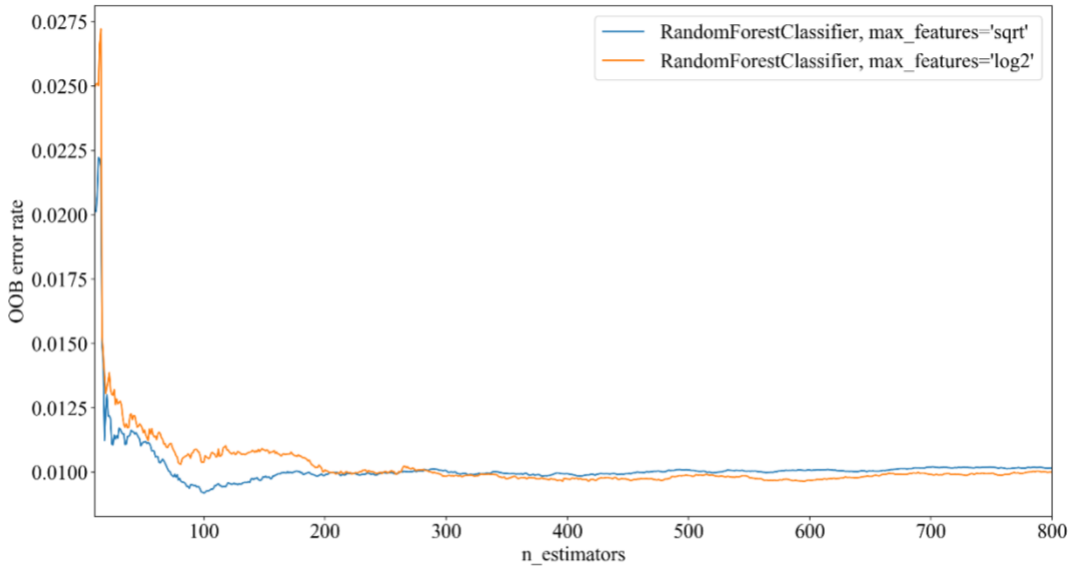

Fig. 4-19. RFC estimator's training and validation average OOB rate based on tuning the number of estimators

4.5.2. Numerical Studies Using Test System

To demonstrate the effectiveness of the proposed time-series simultaneous classification regression FIDVR assessment on a bigger dataset, 90 events are created in a test system using EMTP-RV and the features are collected from each event to build the proposed time-series RFC/RFR prediction model.

WECC load modeling approach has been utilized to create FIDVR event. In section 4.5.2.1, load modeling is explained and in the next chapters the result of RFR/RFC are demonstrated.

4.5.2.1. Load Modeling

4.5.2.1.1. Various load models

The dynamic load model structures for measurement-based load modeling in the literature are i) static constant impedance-current-power (ZIP) component and a dynamic induction motor (IM) component [82], [83], ii) complex load model (CLOD) developed by PSS/E [84] and iii) Western Electricity Coordinating Council (WECC) composite load model. ZIP+IM model has been widely used due to its simple structure and CLOD model has several more components including two IMs of different torque-speed and current-speed curves. However, these models were not successful in modeling FIDVR phenomenon [24], [25]. To overcome the drawbacks of mentioned modeling techniques, the WECC composite load model has been developed [85], which includes a single-phase IM component among other enhancements.

The WECC has been investigated FIDVR events on 27 air-conditioner (A/C) units during voltage and frequency deviations. According to these studies, it has been noticed that A/C units typically stall within 3 cycles. Their stalling voltage with the outdoor temperature of 80°F, 100°F and 115°F are 60%, 65% and 70% respectively. Operation time of thermal overload protection switches has inverse relation with the stall current. Depending on the stall current, thermal protection switches open to disconnect the A/C units within 2 to 24 seconds depending on the stalling current. As shown in Table 4-3, the WECC load model has an extensive list of 121 parameters used to describe its static and dynamic behaviors under disturbances [85], while a simple load model such as ZIP+IM model only has 13 parameters [86]. See [85] for more details. These parameters can be

categorized to represent the WECC composite load model as a model consist of substation and feeder, load model components and the fraction for each load component. Fig. 4-20 illustrates the structure of the WECC composite load model which consists of a substation transformer model, a feeder equivalent model, and six load model components [85]. The load components include three three-phase motors (A, B and C), one single-phase motor (air conditioner D), one static ZIP load, and an electronic load, all connected in parallel.

Among all aforementioned load components, the high penetration of single-phase induction motors in power systems results in FIDVR phenomenon and single-phase A/C units represent a significant fraction of the summer on-peak load in WECC and are considered to be contributing to dynamic performance issues in WECC [85].

Table 4-3
List of WECC Composite Load Model Parameters and Values

Motor A		**Ftr2A**	0.47	**TrclB**	0.6	**TtrlC**	0.02	**VrstD**	0.9	**Static Load**		**Feeder**	
FMA	0.167	**Vrc2A**	0.639	**Vtr2B**	0.7	**FtrlC**	0.2	**TrstD**	0.4	**Pfs**	-0.99	**Bss**	0
MtypA	3	**Trc2A**	0.73	**Ttr2B**	0.02	**VrclC**	0.65	**FuvrD**	0.17	**P1e**	2	**Rfdr**	0.04
LFmA	0.7	**Motor B**		**Ftr2B**	0.3	**TrclC**	0.6	**Vtr1D**	0.65	**P1c**	0.54546	**Xfdr**	0.05
RsA	0.04	**FMB**	0.167	**Vrc2B**	0.85	**Vtr2C**	0.7	**Ttr1D**	0.02	**P2e**	1	**Fc**	0.75
LsA	1.8	**MtypB**	3	**Trc2B**	inf	**Ttr2C**	0.02	**Vtr2D**	0.9	**P2c**	0.45454	**Xxf**	0.08
LpA	0.1	**LFmB**	0.8	**Motor C**		**Ftr2C**	0.3	**Ttr2D**	5	**Pfrq**	-1	**Tfixhs**	1
LppA	0.083	**RsB**	0.34	**FMC**	0.167	**Vrc2C**	0.85	**Vc1offD**	0.4	**Q1e**	2	**Tfixls**	1
TpoA	0.092	**LsB**	1.8	**MtypC**	3	**Trc2C**	inf	**Vc2offD**	0.4	**Q1c**	-0.5	**LTC**	1
TppoA	0.002	**LpB**	0.16	**LFmC**	0.8	**Motor D**		**Vc1onD**	0.45	**Q2e**	1	**Tmin**	0.9
HA	0.05	**LppB**	0.12	**RsC**	0.34	**FmD**	0.167	**Vc2onD**	0.45	**Q2c**	1.5	**Tmax**	1.1
EtrqA	0	**TpoB**	0.1	**LsC**	1.8	**MtypD**	1	**VthD**	30	**Qfrq**	-1	**Step**	0.00625
Vtr1A	0.75	**TppoB**	0.0026	**LpC**	0.16	**LFmD**	1	**Th1tD**	0.3	**MBase**	0	**Vmin**	1
TtrlA	Inf	**HB**	1	**LppC**	0.12	**CompPFD**	0.97	**Th2tD**	2.05	**Electronic Load**		**Vmax**	1.02
FtrlA	0.2	**EtrqB**	2	**TpoC**	0.1	**VstallD**	0.6	**TvD**	0.025	**Fel**	0.167	**Tdel**	30
VrclA	0.9	**Vtr1B**	0.5	**TppoC**	0.0026	**RstallD**	0.1			**Pfel**	1	**Tdelstep**	5

TrclA	Inf	**TtrlB**	0.02	**HC**	1	**CstallD**	0.1			**Vd1**	0.75	**Rcmp**	0
Vtr2A	0.5	**FtrlB**	0.2	**EtrqC**	2	**TstallD**	0.02			**Vd2**	0.65	**Xcmp**	0
Ttr2A	0.02	**VrclB**	0.65	**Vtr1C**	0.5	**FrstD**	0			**Frcel**	0.25		

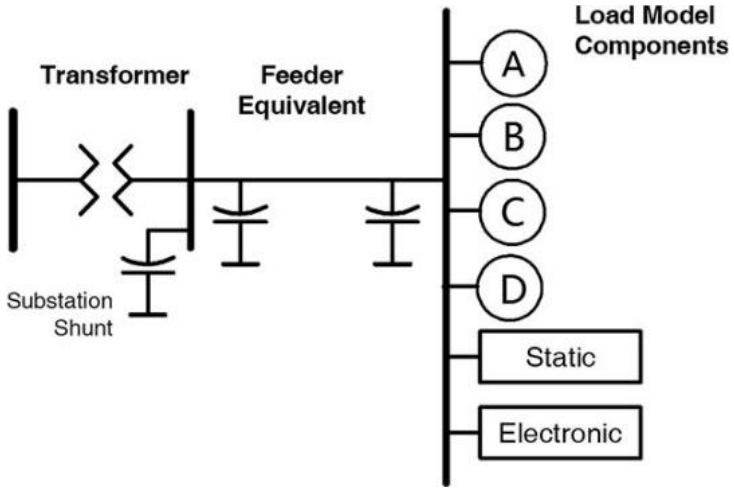

Fig. 4-20. Schematic of the WECC composite load model [87].

4.5.2.1.2. Single-phase air-conditioner modeling

Most single-phase compressor motors are capacitor-run capacitor-start motors. As shown in Fig. 4-21, the motor actually has two windings: main or run winding (R), and auxiliary or start winding (S) which are T connected to a common terminal. The single-phase induction motor model in EMTP-RV is shown in Fig. 4-22. The supply voltage and the capacitor in series with the winding creates the auxiliary winding voltage. There is an internal thermal relay located in the common wire. A large number of single-phase residential A/C units were tested by Bonneville Power Administration,

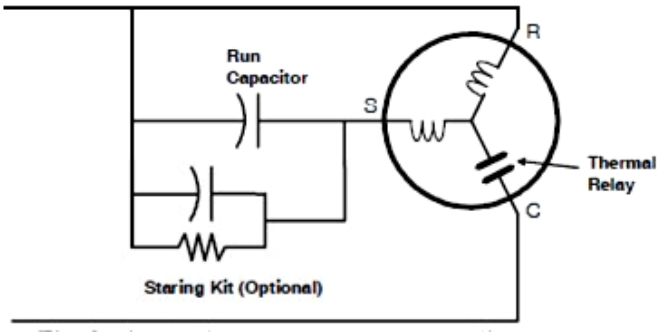

Fig. 4-21 Schematic of a single-phase compressor motor diagram [85]

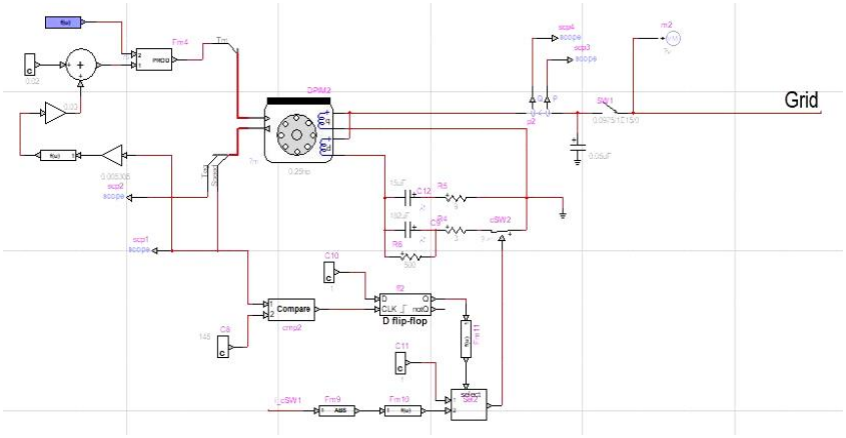

Fig. 4-22. Schematic of a single-phase compressor motor in EMTP-RV

Southern California Edison and EPRI investigated the dynamic behavior of these loads. The Voltage and frequency oscillations, frequency ramps and steps, voltage ramps, voltage sags and faults of various magnitude, duration and recovery have been analyzed in these studies [85].

To demonstrate the real and reactive power sensitivity with respect to voltage, slow voltage ramp down tests are performed. Fig. 4-23 shows trajectories of compressor real and reactive power versus supply voltage at temperatures of 80 °F in the lower line, middle lines at 100 °F and upper lines at 115 °F. As shown, the motor stalls at approximately 60-65% voltage. The higher the voltage, the higher temperature. Once the motor is stalled,

even if the voltage recovers, it remains stalled since the compressor is unable to restart against the full head of pressure. It takes one to five minutes to restart since the pressure has to be equalized.

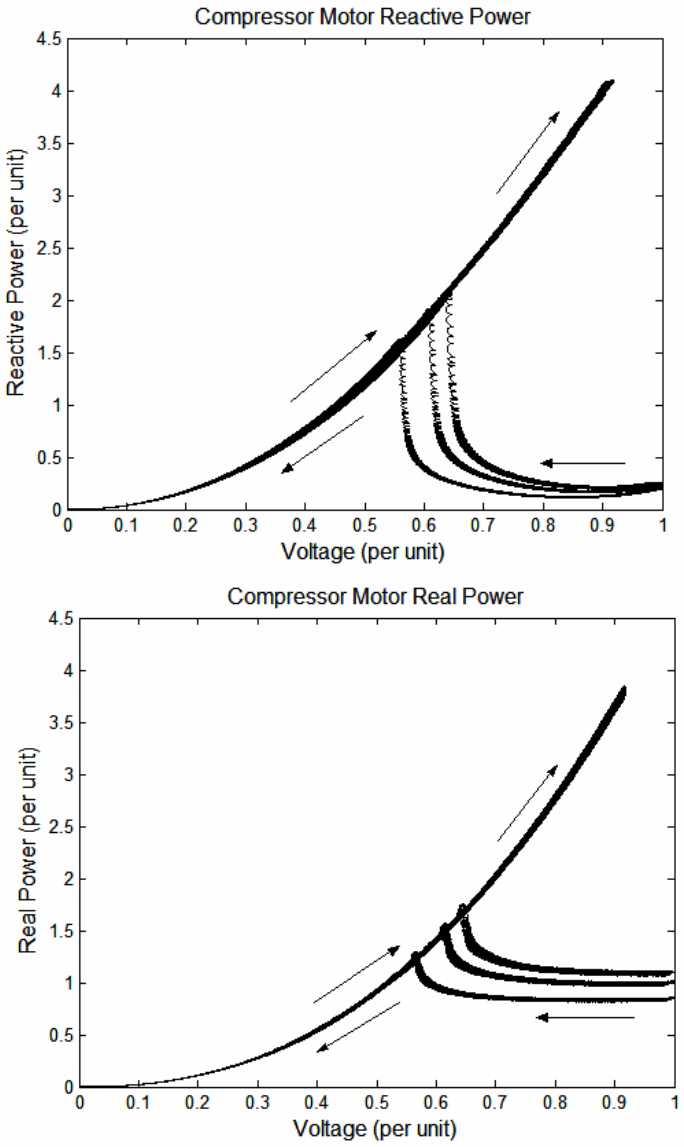

Fig. 4-23. Compressor motor real and reactive power versus supply voltage (voltage ramp down test).

A. Thermal Protection modeling

The WECC model contains four explicit motor models representing high and low inertia motors. In addition, two constant torque and torque varying with speed squared models

are also considered. The low inertia, constant torque model which are mainly air conditioner units has thermal protection element.

Thermal protection determines the amount of stalled air conditioners trip over time. The delay in thermal protection provides the delay in delayed voltage recovery. Since this characteristic is critical for FIDVR analysis, it is reasonable to implement appropriate motor protection model to better simulate FIDVR phenomenon [60].

Fig. 4-24 shows the experimental data for thermal protection operation as a function of stalled voltage for each of the tested air conditioner units. As it can be seen, stalled voltage has inverse relation with thermal protection tripping time since the stall current increase as the stall voltage increase. With the increase of stall current, the thermal protection would operate faster. As it can be seen in Fig. 4-24, when exposed to a fault which cause delayed voltage recovery, most of the tested A/C units stalled from 1.0 to 20 seconds before the thermal protection switch tripped to protect the compressor. The right end of each plot is where the unit started stalling (stall threshold) and the left end is where the contactor opens (contactor dropout voltage) [60].

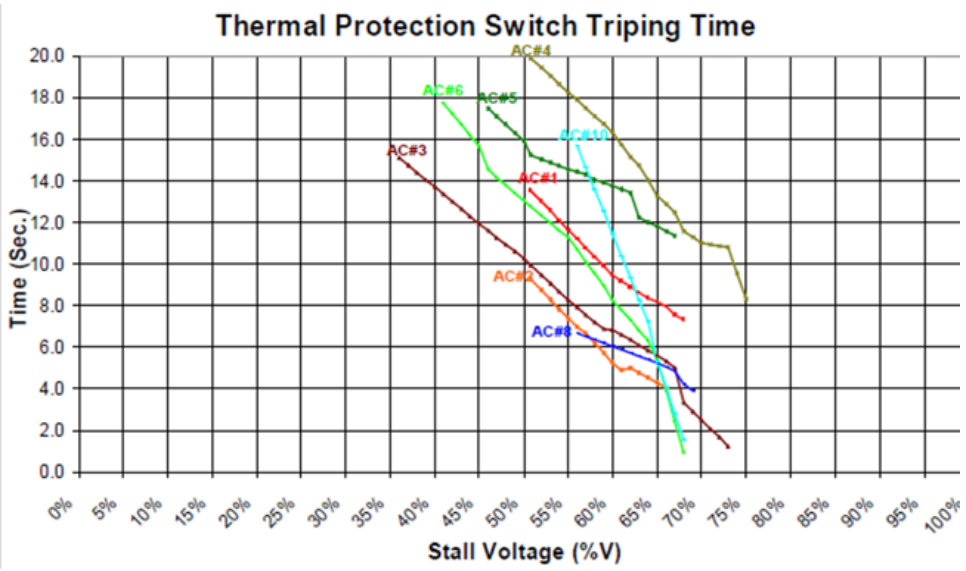

Fig. 4-24. Relation between thermal protection switch tripping time and voltage [60]

Thermal relay plays a critical role in the simulating voltage stability events which is protecting the motor from overheating due to extended stall currents. Without the thermal protection switch, the motor might overheat and fail as a result of absorbing high currents for long periods. The inverse relation between time delay in operation of thermal protection and stall current is typical for these protective devices. The higher the current the faster the operation. A block diagram model for a thermal relay is presented in Fig. 4-25. The transfer function in delays the response in the operation of the thermal relay. The delay decreases with increased stall current. The input of the model ($R(I_C)^2$) is the heat produced by the compressor motor current. The current at which the temperature stays at the equilibrium is shown by I_E. An integrator has been used and its input is the heat to estimate the compressor winding temperature raise (θ). A feedback K is used to model heat dissipation. Once the temperature exceeds the threshold (θ_{TRIP}), the compressor motor will trip [85].

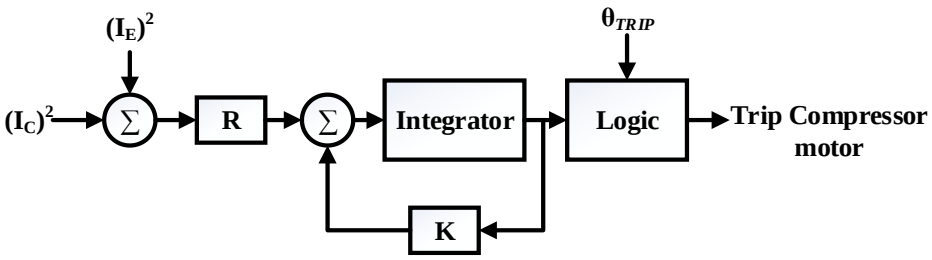

Fig. 4-25. Thermal protection relay model [60]

Finally, by utilizing aforementioned load modeling, a 9-bus system is modeled in EMTP-RV to create FIDVR and non-FIDVR events. The system topology is shown in Fig. 4-26.

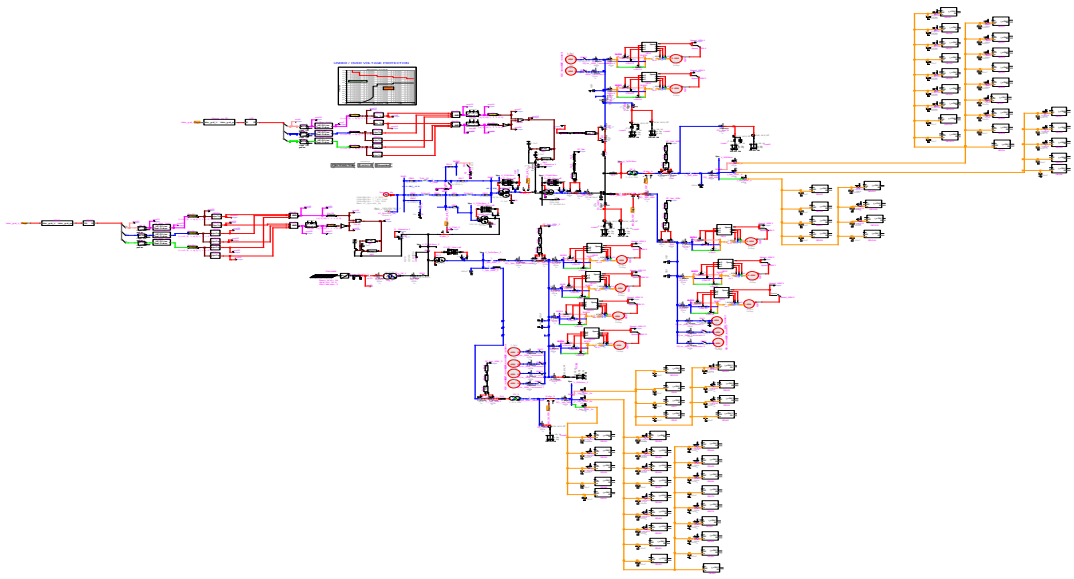

Fig. 4-26. The 9-Bus test system

4.5.2.1.3. Random Load Disconnection

Small motors with thermal protection may take from 1 to 20 s to trip after stall [60], [34], [5], while larger motors with under voltage protection trip typically from 0.4 to 1.2 seconds [47].

These motors trip before they stall while the system voltage drops below the protection threshold (between 50% and 90% typically). To implement a random thermal protection relay trip, a random disconnection is modeled as a normal distribution with a mean value and standard deviation are considered for both thermal protection and under voltage protection elements. The assumption for the under-voltage protection trip is between 0 and 2 s and that the thermal protection trip is between 1 and 30 s. Note that the FIDVR duration is considered to be the duration between the fault and the time over voltage finishes. For each FIDVR case, the portion of motor loads are also randomly selected between 0% and 80%.

4.5.2.2. Feature Analysis

As it has been explained, feature analysis basically explains which features are more important in training of model. The selected features for training the proposed time series RFC/RFR algorithms are voltage deviations, pre-fault reactive power and QISI, respectively. In order to demonstrate the effect of selected features, average of voltage deviations and average of QISI (AQISI) are used and shown in Fig. 4-27 –Fig. 4-29. As it can be seen the most important feature is voltage deviation which is similar to the result obtained from SCE network (see Fig. 4-14). The next important feature is AQISI and finally Pre-fault reactive power feature correlation with FIDVR is 0.64. The results are quite similar to those of SCE dataset. Pair grid plot of the selected features are shown in Fig. 4-28 and the categorized pair grid plot of the features can be seen in Fig. 4-29. According to the expectations, average voltage deviation and AQISI have inverse

relationship with FIDVR while pre-fault reactive power has direct relation with the delayed voltage.

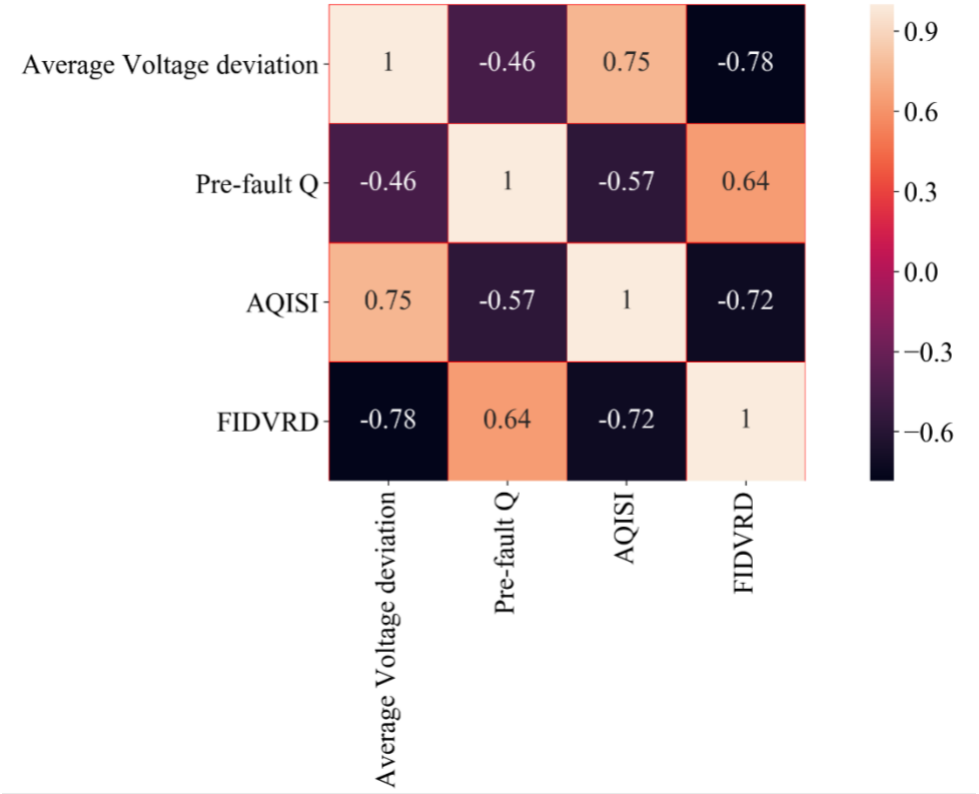

Fig. 4-27. Heat Map of the dataset which demonstrates the correlation between selected features and FIDVRD

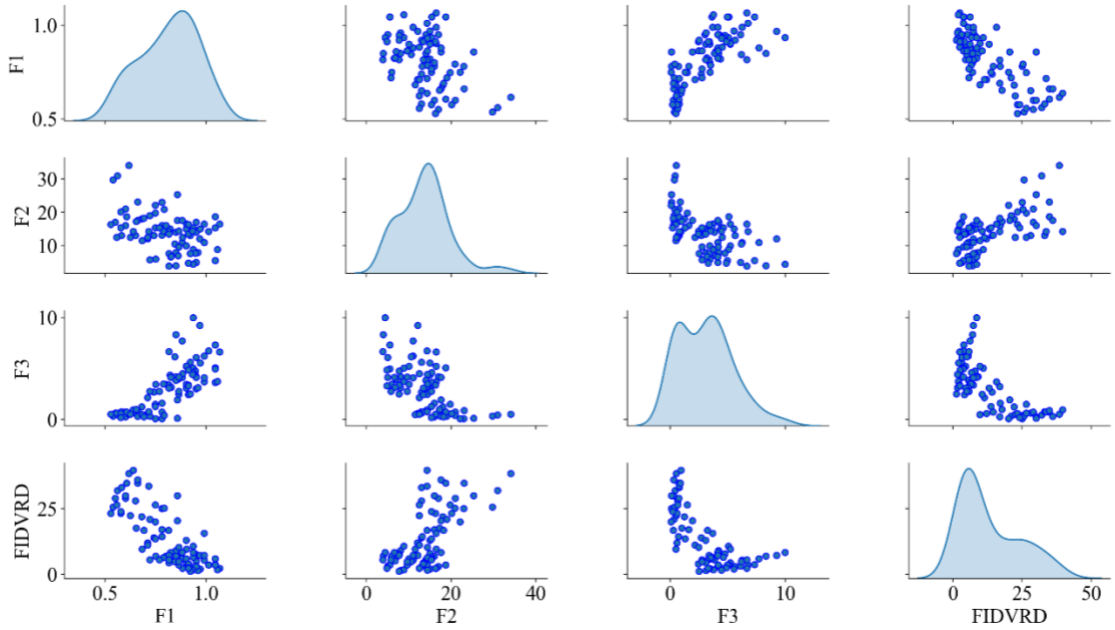

Fig. 4-28. Pair Grid plot of the selected features

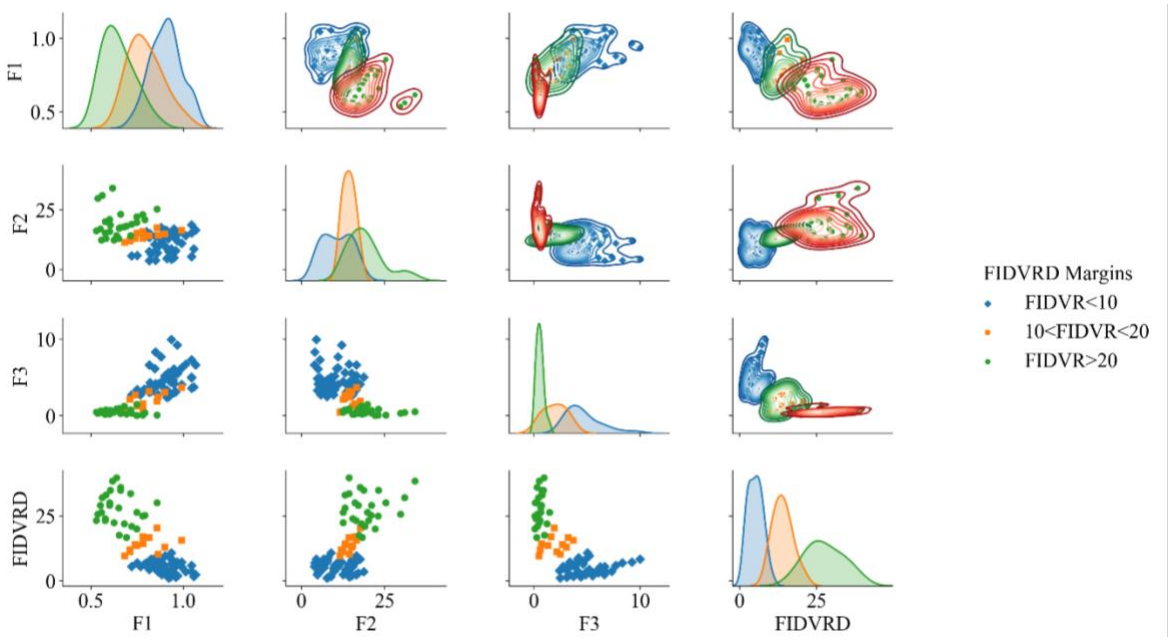

Fig. 4-29. Pair Grid plot of the features categorized according to the FIDVR duration. The higher triangle shows the bivariate kernel density estimation

4.5.2.3. RFC/RFR Result Analysis

After modeling the loads and implementing it in the system, FIDVR events have been created by triggering faults in the transmission lines. The time window width Δt is set to 0.1 s and each 0.01 s a snapshot of the system features is stored, so each snapshot of the selected features consists of 10 points. Considering all the buses in the system, the number of training features is 189 (9 buses ×2 features× 10 trajectory points+9 pre-fault reactive power for each bus). The latest decision time $N\Delta t$ is set to 1 s. Thus N = 10 probabilistic predictors are needed.

Following the training process, RFR and RFC ensembles are simultaneously trained as a classifier to select the event as FIDVR and a FIDVR duration predictor using the features in each time window, respectively. The number of trees in each ensemble RFC/RFR is set to 620. For the probabilistic prediction performed at each time point, the number of estimators (trees) of the RFRs is separately tuned, and the tuning result is shown in Fig. 4-30 where the different colors refer to the tuning results for different time points. The number of estimator's range with the lowest testing MSE is empirically selected and as it can been, the RFR trained for a later time point result in a lower MSE. Such phenomenon coincides with the initial thought, i.e., higher prediction accuracy is at the cost of slower decision speed.

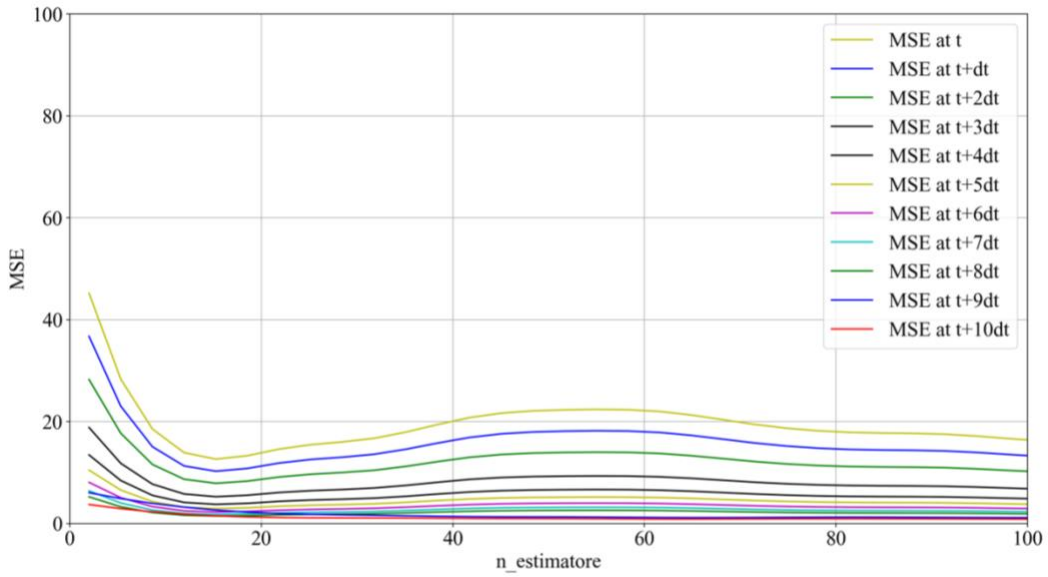

Fig. 4-30. Number of estimators tuning results for ten time-steps

The result of RFC classification is compared to some widely used classification approaches to demonstrate the effectiveness of the proposed methodology. However, since the numbers of features are more than two, the results cannot be shown in a plane. Therefore, a mathematical procedure named principal component analysis (PCA) methodology is used to transforms a number of features into small number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The number of principal components are selected to be two in order to plot the results in the plane.

According to the North American Electric Reliability Corporation (NERC) Transmission Issues Subcommittee, an event is FIDVR if a voltage condition initiated by a fault and characterized by stalling of induction motors where initial voltage recovery after the clearing of a fault is less than 90% of pre-contingency voltage and slow voltage recovery occur more than 2 seconds to expected post-contingency steady-state voltage

level [34]. According to NERC definition, events can be classified as FIDVR or non-FIDVR.

The classification methods are Logistic Regression, Linear Kernel, Sigmoid Kernel, KNN, Gaussian Naive Bayes and Decision Tree and their classification results are shown in Figs. 4-31 to 4-36, respectively. As it can be seen, all aforementioned method have noticeable errors in classifying FIDVR events and non-FIDVR events.

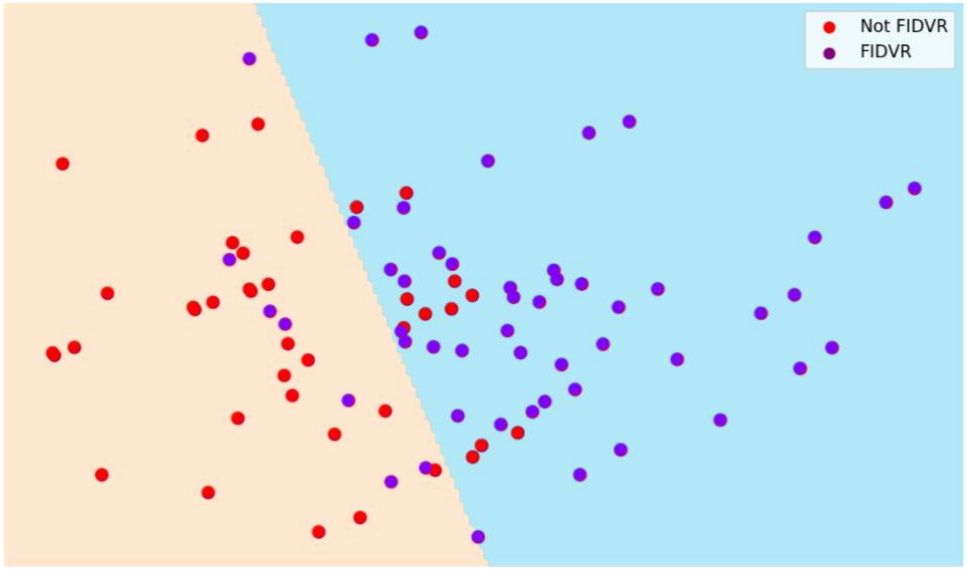

Fig. 4-31. Classification result using Logistic Regression classifier

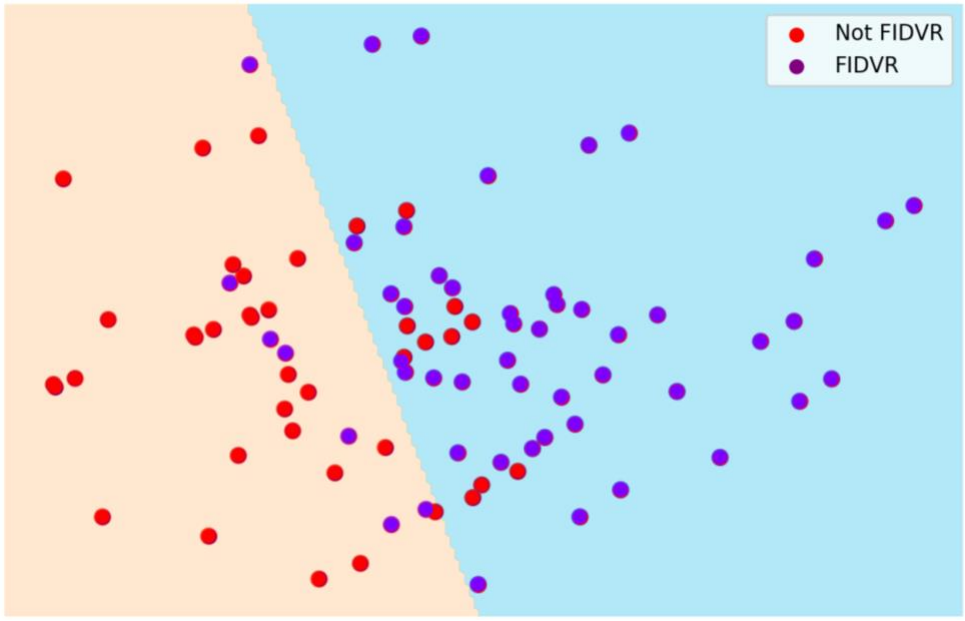

Fig. 4-32. Classification result using Linear Kernel classifier

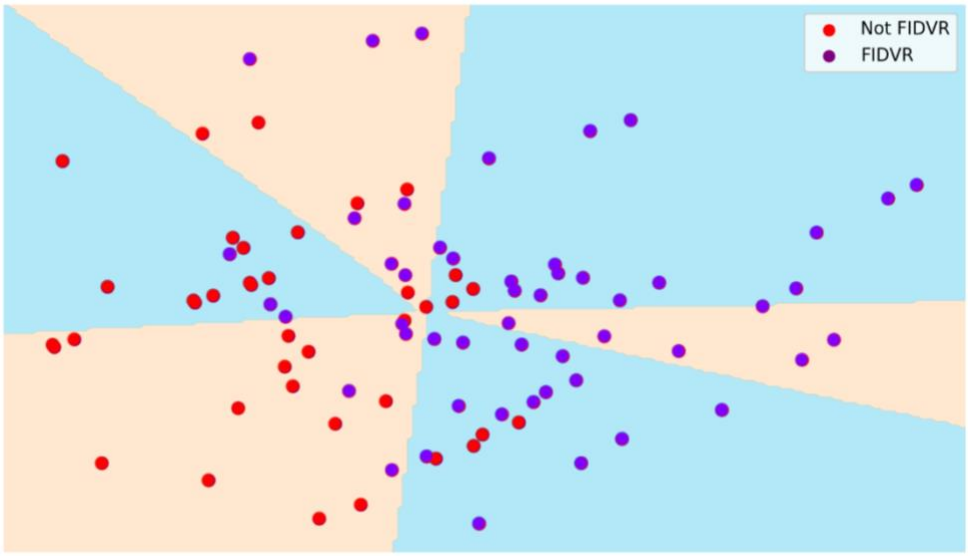

Fig. 4-33. Classification result using Sigmoid Kernel classifier

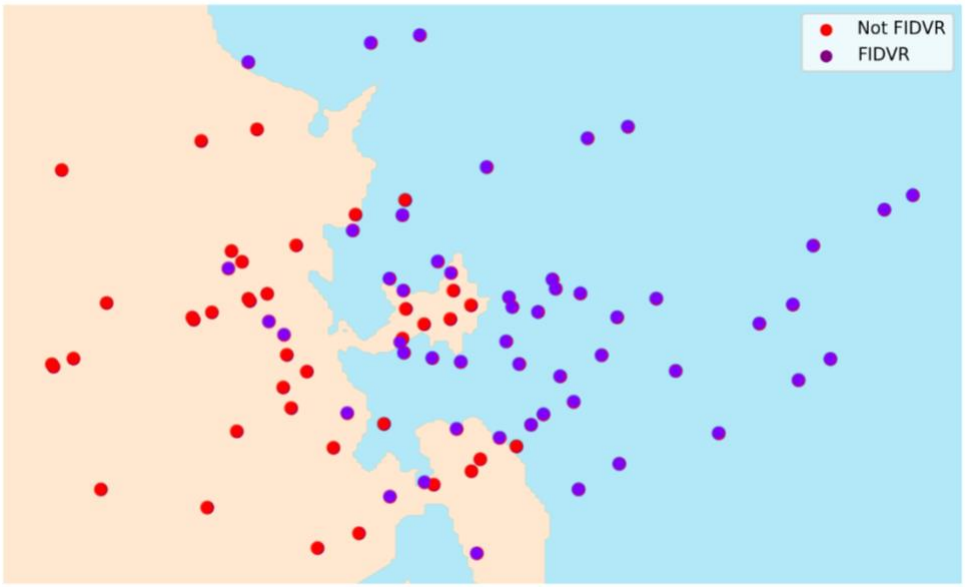

Fig. 4-34. Classification result using KNN classifier

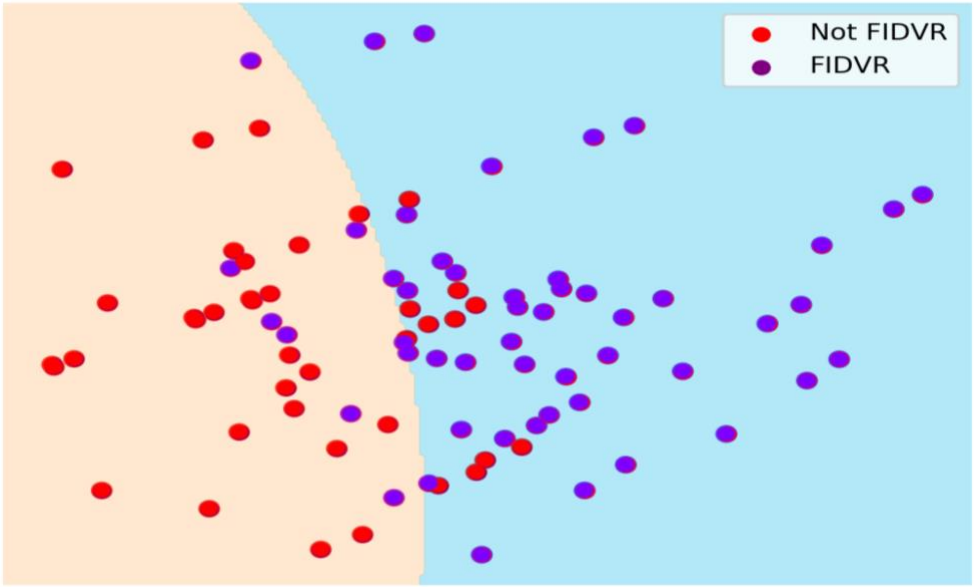

Fig. 4-35. Classification result using Gaussian Naive Bayes classifier

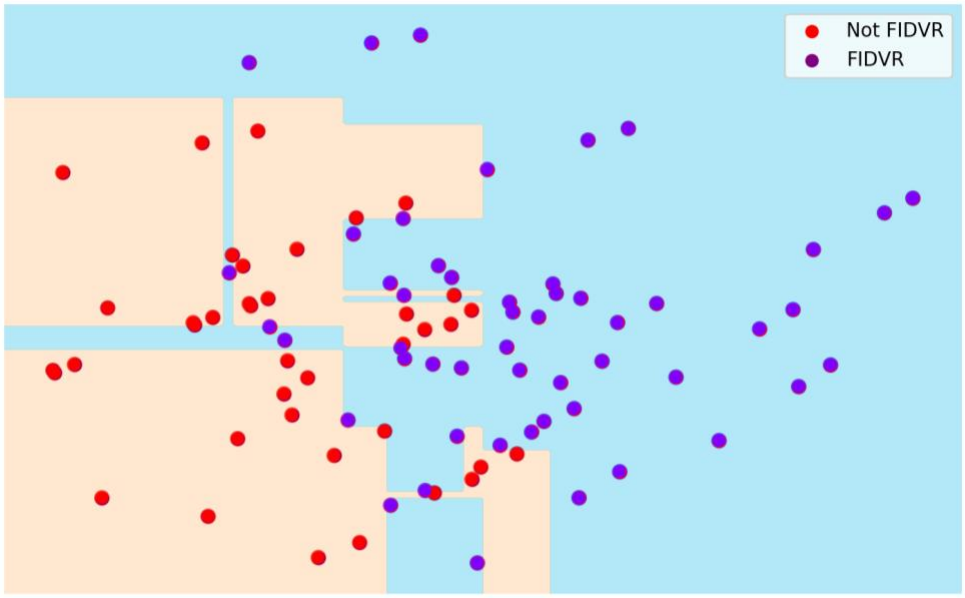

Fig. 4-36. Classification result using Decision Tree classifier

In RFC method, multiple estimators are combined to reduce the effect of this overfitting and increase the classification accuracy. In the proposed time-series RFC method, the performance of the RFCs in each time step is calculated and the final decision is based on the average of all time-steps. By using the bagging (RFC) method which makes use of an ensemble of parallel estimators, and using time-step method, the estimator can find a better classification. In this study, optimized ensemble of randomized decision trees is implemented in the RFC estimator of *Scikit-Learn* package. The results of the classification using proposed time-series RFC based method is shown in Fig. 4-37 where the classification accuracy is much better than the Fig. 4-31 –Fig. 4-36. The red and purple dot are the data points showing the Non-FIDVR and FIDVR events, respectively. The yellow and blue areas are the classified by the classifier. Note that X-axis and Y-axis are the X and Y values created by using PCA method and does not represent the value of selected features.

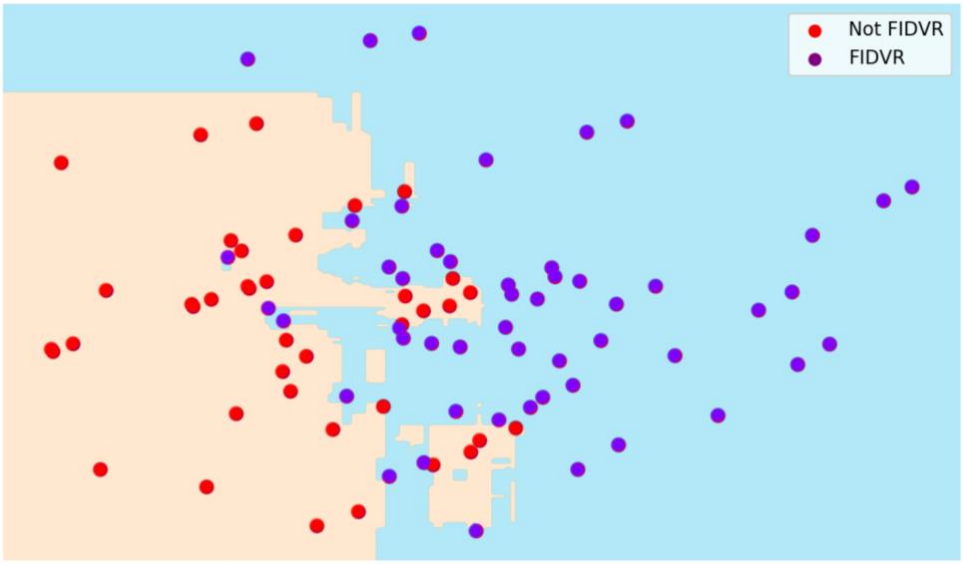

Fig. 4-37. Classification result using proposed time-series RFC classifier

4.6. **Conclusion**

This chapter proposed a probabilistic time-series simultaneous RFR/RFC-based FIDVR assessment method based on pre-fault reactive power, post-fault real-time voltage trajectories and post-fault reactive power increment slope index (QISI). The assessment decision can be made as early as possible without impairing its accuracy. To achieve this, voltage magnitude and QISI, as time-series features for FIDVR analysis, are measured at each sliding time window using PQubes, and the assessment decision is made in such a way that the event is analyzed using RFCs to categorize event either as FIDVR or non-FIDVR and if selected as FIDVR, the delayed voltage duration is predicted using RFR. In doing so, the assessment speed is significantly high along with high decision accuracy which can be reliably obtained at each time window. The final decision is based on the majority vote of all the time steps. The excellent performance of the proposed method has

been demonstrated on SCE network real data and on a test power system. By utilizing the proposed fast and highly accurate FIDVR detection and prediction, the subsequent load shedding approaches can be employed not only faster, but also with the most optimum amount of load to be shed.

Chapter 5 : THESIS SUMMARY AND FUTURE WOKS

5.1. Thesis summary

Two LS techniques were developed in the second chapter, i) a simplified LS technique and ii) a complete LS approach. The simplified LS technique derives PF-based P-V curve according to the type, size, and the power factor of the system loads. The obtained PF-based P-V curve as well as the frequency droop characteristics are used for calculating UVA, UFA and OFA and the final LSA is calculated. The complete LS method is based on the sensitivity factors in pre-disturbance conditions to address the dynamic changes of the operating conditions subsequent to the disturbance. In this method, several PF-based P-V curves are generated and two PF-based P-V curves, i.e., upper and lower PF values, closest to the loads are selected and the LSA is calculated. The proposed LS approach is validated based on the detailed time-domain simulation in the PSCAD program.

A comparative analysis in an islanded IBMG based on the time-domain simulation and taking into account various load models is carried out to investigate the effectiveness of the proposed LS techniques, compared with a previously published method. Three possible disturbances are considered in the study system. The study reveals that the proposed complete LS technique results in the lowest LSA and effectively restores both voltage and frequency of the islanded IBMG to the desired values. Moreover, the system stability during disturbance and after the LS has been investigated which demonstrates the

effectiveness of UVFLS in enhancing the system stability.

The third chapter proposed a data-driven multi-variable machine learning-based Decision-making model for prediction of FIDVR duration. Relation of several power system data with FIDVR duration has been investigated. Selected features have been utilized for online machine learning model development. The stored data of power quality recording devices (PQubes) are used for online machine learning model developing. Several simple regression models such as single-variable and multi-variable regression models i.e. linear regression and polynomial regression models have been developed and compared. The real-time PQubes data can be used for rapid real-time prediction of the FIDVR duration following a system disturbance. By using the developed model, FIDVR duration can be obtained using real-time data without using complicated load modeling. Simple regression models result in simple programming and fast prediction and on the other hand, using multiple features for model development enhance the model accuracy. As a result, the FIDVR prediction can be accomplished very fast, yet with acceptable accuracy.

The forth chapter proposed a probabilistic time-series simultaneous RFR/RFC-based FIDVR assessment method based on pre-fault reactive power, post-fault real-time voltage trajectories and post-fault reactive power increment slope index (QISI). The assessment decision can be made as early as possible without impairing its accuracy. To achieve this, voltage magnitude and QISI, as time-series features for FIDVR analysis, are measured at each sliding time window using PQubes, and the assessment decision is made in such a way that the event is analyzed using RFCs to categorize event either as FIDVR or non-FIDVR and if selected as FIDVR, the delayed voltage duration is predicted using RFR. In

doing so, the assessment speed is significantly high along with high decision accuracy which can be reliably obtained at each time window. The final decision is based on the majority vote of all the time steps. The excellent performance of the proposed method has been demonstrated on SCE network real data and on a test power system. By utilizing the proposed fast and highly accurate FIDVR detection and prediction, the subsequent load shedding approaches can be employed not only faster, but also with the most optimum amount of load to be shed.

5.2. **Contributions**

The second chapter of this thesis proposed an enhanced method for deriving P-V curves according to load PF and two simple and complete load shedding techniques to recover bus voltages and frequency of the system. In the simple load shedding method, last pre-disturbance PF-based P-V curves are utilized for calculating load shedding amount. However, in complete load shedding method, an algorithm has been proposed which calculates several PF-based P-V curves for each load bus and selects two of these curves according to loads PF in post-disturbance. Finally, one of these two PF-based P-V curve which results in the optimum load shedding amount and maintains both bus voltage and frequency into desired values is used for load shedding. By implementing this procedure, the closest PF-based P-V curve to the loads PF in post-disturbance will be chosen.

There is not any literature which studies loads PF and load type on P-V curves in microgrids to analyze the load shedding amount according to load model. In addition, the available load shedding strategies are typically based on pre-fault sensitivity factors and,

hence, do not address the dynamic changes due to a contingency. For instance, if loads power factor changes during the disturbance, the pre-disturbance values are used for obtaining load shedding amount. In the complete load shedding method, the load shedding is accomplished according to the load PF during the disturbance. In addition, microgrid is modeled based on state-space equation. All the loads, lines and inverters are modeled in state-space equation and any change in the system characteristics can be done by changing the corresponding matrixes. Hence, if any changes occur in the system configuration such as change in load/line/DG sizing/type/locations, the state-space equations will be updated accordingly and, as a result, the responses of the system will be updated correspondingly. For instance, if the resistance of a line increases, the corresponding matrix related to the respective line will be changed and voltage of respective bus connected to the line will drop and the slope of respective PF-based P-V curve of that bus will increase. Whatever the topology of the system is, the value of bus voltages and therefore PF-based P-V curves can be obtained by using state-space equation and time-step simulation. Based on the obtained PF-based P-V curves and the frequency of the system, load shedding amount can be obtained according to the technique proposed in the second chapter. The proposed method is independent of the topology of the system since it is based on simple mathematical computations such as algebraic operations, maximum and minimum operators, and rule-based decision makings.

To encapsulate, the contributions of this study are as follows:

1. Comprehensive investigation on loads type and power factor of P-V curves in microgrid.

2. Proposing the algorithm for obtaining PF-based P-V curve which considers load type and power factor.
3. Proposing under voltage-frequency load shedding strategy based on the fault sensitivity factors so as to address the dynamic changes due to a contingency.
4. Implementing under voltage-frequency load shedding method based on complete state-space model of microgrid
5. The LSA using simple LS and complete LS methods are respectively 32.6% and 37.28% lower than the LSA using the existing method in the literature.

The third chapter focused on investigating FIDVR duration dependency on the system characteristic as a comprehensive sensitivity analysis. The most important features are used for ML training in the next step. A comprehensive data-driven measurement-based ML-based technique for real-time FIDVR prediction is proposed which employs several power system data as an input for online ML design. The proposed method guarantees fast and accurate decision making for the consecutive LS methods. The novel contribution of the proposed method is summarized as follows:

1) A comprehensive feature selection to investigate the relation of each power system data with FIDVR prediction.

2) Instead of using complicated method which may reduce the real-time prediction speed and increase computational programming, simple regression methods are employed to enhance the processing speed along with several important power system features to increase the accuracy of the model.

3) An online ML training is accomplished to build a model for real-time FIDVR assessment. Linear and polynomial regression models are used and compared.

4) Validation methods have been proposed and used for analysing the proposed FIDVR prediction accuracy.

5) The FIDVR duration is assessed in a continuous manner. The real-time FIDVR prediction is executed as a regression method.

6) Since the model is simple, the real-time FIDVR prediction processing is fast enough which provides enough time for consecutive LS.

In the fourth chapter a simultaneous classification and regression assessment is proposed using a time-series measurement-based machine learning technique for real-time FIDVR prediction. The proposed method employs pre-fault reactive power, post-fault voltage magnitudes for the model training. In addition to the mentioned features, new indices were proposed and their correlation with FIDVR duration will be investigated. These indices are active power deviation index (PDI), reactive power deviation indices (QDI), active power increment slope index (PISI) and reactive power increment slope index (QISI). Finally, QISI was utilized for training the proposed time-series data-driven simultaneous RFC/RFR prediction method. The novel contribution of the proposed method is summarized as follows:

1) An advanced feature analysis to investigate the effectiveness of power system data on FIDVR phenomenon.

2) A probabilistic k-fold validation methods has been utilized for enhancing the proposed FIDVR prediction accuracy.

3) A real-time probability assessment of FIDVR phenomenon occurrence is accomplished as a time-series classification machine learning decision making algorithm.

4) The FIDVR duration is assessed in a continuous manner. The real-time FIDVR prediction is executed as a time-series regression method.

5) An ensemble machine learning approach has been utilized in each step of both time-series classification and regression decision making models which guarantees the test score improvement in real power system event.

5.3. **Future works**

As future work on under-voltage frequency load shedding, the effect of more complicated loads i.e. voltage and frequency dependent loads and active loads on PF-based P-V curves and load shedding of islanded microgrids can be analyzed.

With regard to assessing FIDVR events in power systems, more complicated methods i.e. neural networks can be utilized for building a machine learning model and predict FIDVR more accurately. However, there is a trade-off between the model accuracy and computation time. More features can also be analyzed and employed in building the model according to their relationship with FIDVR. The researches in this study focuses on assessing the FIDVR and predicting its duration. As future study, more complicated AI algorithms such as sequential neural networks can be employed in order to predict the voltage curve. When the delayed voltage curve is predicted, high information would be available for subsequent remedial control techniques.

PUBLICATIONS

1) S. Rahmani and A. Rezaei-Zare, " A Time-Series Data-Driven Multi-Variable Simultaneous Classification and Regression Decision Making for Real-Time Fault-Induced Delayed Voltage Recovery Prediction," in *IEEE Transactions on Sustainable Energy* (Submitted).

2) S. Rahmani, A. Rezaei-Zare and A. Hooshyar, " A Data-Driven Multi-Variable Polynomial Regression Analysis for Real-Time Fault-Induced Delayed Voltage Recovery Prediction," in *IEEE Transactions on Power Systems* (Submitted).

3) S. Rahmani, A. Rezaei-Zare, M. Rezaei-Zare and A. Hooshyar, "Voltage and Frequency Recovery in an Islanded Inverter-Based Microgrid Considering Load Type and Power Factor," in *IEEE Transactions on Smart Grid*. doi: 10.1109/TSG.2019.2899912

4) S. Rahmani and A. Rezaei-zare, " Prediction of System Voltage Recovery due to Single Phase Induction Motor Stall Using Machine Learning Techniques," *2019 IEEE Workshop on Electrical Machine Design, Control and Diagnostics*, Athens, Greece.

5) S. Rahmani and A. Rezaei-zare, "Under Voltage-Frequency Load Shedding in an Islanded Inverter-based Microgrid using Power Factor-based P-V curves," *2018 IEEE Electrical Power and Energy Conference (EPEC)*, Toronto, ON, 2018, pp. 1-6.
doi: 10.1109/EPEC.2018.8598403

Other Publications

6) M. Saeidi, S. Rahmani and A. Pirayesh, "Modeling microgrids with voltage and frequency dependent loads," *2017 IEEE 7th International Conference on Power and Energy Systems (ICPES)*, Toronto, ON, 2017, pp. 101-105.
doi: 10.1109/ICPESYS.2017.8215929

7) S. Rahmani, M. Saeidi and A. Pirayesh, "A combinational power sharing strategy based on master-slave and droop methods for power-electronics-interfaced distributed generation units operating in a DC micro-grid," *2017 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, Oshawa, ON, 2017, pp. 1-6.
doi: 10.1109/SEGE.2017.8052767

8) S. Rahmani, R. Chabanloo and M. Saeidi, "Voltage estimation in an islanded micro-grid using transient state estimation technique," *2017 Iranian Conference on Electrical Engineering (ICEE)*, Tehran, 2017, pp. 2117-2122.
doi: 10.1109/IranianCEE.2017.7985411

9) M. Saeidi, S. Rahmani and S. H. Fathi, "Analysis of wind speed effect on grid integration of PMSG-based WECS," *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, Shiraz, 2016, pp. 389-394.
doi: 10.1109/IranianCEE.2016.7585552

10) S. Rahmani and A. A. Razi-Kazemi, "Investigation of very fast transient over voltages in gas insulated substations," *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, Tehran, 2015, pp. 428-435.
doi: 10.1109/KBEI.2015.7436083

BIBLIOGRAPHY

[1] T. Van Cutsem, "Voltage instability: Phenomena, countermeasures, and analysis methods," *Proc. IEEE*, vol. 88, no. 2, pp. 208–227, 2000.

[2] M. G. L. Prabha Kundur, Neal J. Balu, "Power System Stability And Control." 1994.

[3] R. Hooshmand and M. Moazzami, "Optimal design of adaptive under frequency load shedding using artificial neural networks in isolated power system," *Int. J. Electr. Power Energy Syst.*, vol. 42, no. 1, pp. 220–228, 2012.

[4] J. Tang, J. Liu, F. Ponci, and A. Monti, "Adaptive load shedding based on combined frequency and voltage stability assessment using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 2035–2047, 2013.

[5] A. P. Ghaleh, M. Sanaye-Pasand, and A. Saffarian, "Power system stability enhancement using a new combinational load-shedding algorithm," *IET Gener. Transm. Distrib.*, vol. 5, no. 5, p. 551, 2011.

[6] A. Saffarian and M. Sanaye-Pasand, "Enhancement of power system stability using adaptive combinational load shedding methods," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1010–1020, 2011.

[7] V. N. Chuvychin and N. S. Gurov, "An adaptive approach to load shedding and spinning reserve control during underfrequency conditions," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 1805–1810, 1996.

[8] P. Mahat, Z. Chen, and B. Bak-jensen, "Under frequency load shedding for an islanded distribution system with distributed generators," *IEEE Trans. Power Deliv.*, vol. 25, no. 2, pp. 911–918, 2010.

[9] V. V. Terzija, "Adaptive underfrequency load shedding based on the magnitude of the

disturbance estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1260–1266, 2006.

[10] E. E. Aponte, S. Member, and J. K. Nelson, "Time optimal load shedding for distributed power systems," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 269–277, 2006.

[11] D. Y. Yang, G. W. Cai, Y. T. Jiang, and C. Liu, "Centralized adaptive under frequency load shedding schemes for smart grid using synchronous phase measurement unit," *J. Electr. Eng. Technol.*, vol. 8, no. 3, pp. 446–452, 2013.

[12] Q. Xu, B. Yang, C. Chen, F. Lin, and X. Guan, "Distributed Load Shedding for Microgrid with Compensation Support via Wireless Network," 2017.

[13] A. Mondal and M. S. Illindala, "Improved Frequency Regulation in an Islanded Mixed Source Microgrid Through Coordinated Operation of DERs and Smart Loads," *IEEE Trans. Ind. Appl.*, vol. 54, no. 1, pp. 112–120, 2018.

[14] E. Pashajavid and A. Ghosh, "Frequency support for remote microgrid systems with intermittent distributed energy resources - A two-level hierarchical strategy," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2760–2771, 2018.

[15] V. H. Bui, A. Hussain, and H. M. Kim, "A Strategy for Flexible Frequency Operation of Stand-Alone Multimicrogrids," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1636–1647, 2018.

[16] U. Rudez and R. Mihalic, "Monitoring the first frequency derivative to improve adaptive underfrequency load-shedding schemes," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 839–846, 2011.

[17] S. M. Halpin, K. A. Harley, R. A. Jones, and L. Y. Taylor, "Slope-permissive under-voltage load shed relay for delayed voltage recovery mitigation," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1211–1216, 2008.

[18] R. Tzoneva, A. C. Adewole, and A. Apostolov, “Adaptive under-voltage load shedding scheme for large interconnected smart grids based on wide area synchrophasor measurements,” *IET Gener. Transm. Distrib.*, vol. 10, no. 8, pp. 1957–1968, 2016.

[19] [H. Seyedi and M. Sanaye-Pasand, “New centralised adaptive loadshedding algorithms to mitigate power system blackouts,” *IET Gen., Transm., Distrib*, vol. 44, no. 6, pp. 202–206, 2015.

[20] M. Choobineh, D. Silva-Ortiz, and S. Mohagheghi, “An Automation Scheme for Emergency Operation of a Multi-Microgrid Industrial Park,” *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 6450–6459, 2018.

[21] K. Seethalekshmi, S. N. Singh, and S. C. Srivastava, “A synchrophasor assisted frequency and voltage stability based load shedding scheme for self-healing of power system,” *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 221–230, 2011.

[22] and D. S. D. Prasetijo, W. R. Lachs, “A new load shedding scheme for limiting underfrequency,” *IEEE Trans. Power Syst. Vol. 9, No. 3*, vol. 9, no. 3, pp. 1371–1378, 1994.

[23] K. Mollah and N. C. K. Nair, “Coordinated strategy for under-voltage and under-frequency load shedding,” *Power Eng. Conf. (AUPEC), 2010 20th Australas. Univ.*, pp. 1–6, 2010.

[24] B. Hoseinzadeh, F. M. Faria Da Silva, and C. L. Bak, “Adaptive Tuning of Frequency Thresholds Using Voltage Drop Data in Decentralized Load Shedding,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2055–2062, 2015.

[25] S. Nourollah, A. Pirayesh, and F. Aminifar, “Combinational scheme for voltage and frequency recovery in an islanded distribution system,” *IET Gener. Transm. Distrib.*, vol. 10, no. 12, pp. 2899–2906, 2016.

[26] A. Mahmoudi, S. H. Hosseinian, M. Kosari, and H. Zarabadipour, "A new linear model for active loads in islanded inverter-based microgrids," *Int. J. Electr. Power Energy Syst.*, vol. 81, pp. 104–113, 2016.

[27] N. Pogaku, S. Member, M. Prodanovic, T. C. Green, and S. Member, "Modeling, analysis and testing of autonomous operation of an inverter-based microgrid," *IEEE Trans.Power Electron*, vol. 22, no. 2, pp. 613–625, 2007.

[28] K. Lindén and I. Segerqvist, "Modelling of Load Devices and Studying Load / System Characteristics," *M.S. thesis, Dept. Electron. Electr. Power Eng. Chalmers Univ*, no. 131, p. 151, 1992.

[29] G. Wood, A., Wollenberg, B., Sheble, "Power generation, operation and control'," 2014.

[30] HYDRO ONE NETWORKS INC., "Distributed Generation Technical Interconnection Requirements Interconnections at Voltages 50kV and Below," vol. 3, p. 183, 2013.

[31] W. Ma, "Dynamical Systems Analysis in Electric Power Systems," Cornell University, Ithaca, NY, USA, 1992.

[32] D. Hill, "Definition and Classification of Power System Stability IEEE/CIGRE Joint Task Force on Stability Terms and Definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, 2004.

[33] C. Grande-moran and R. Wells, "A Case Study on Aggregate Load Modeling in Transient Stability Studies," pp. 1–24.

[34] "A Technical Reference Paper Fault - Induced Delayed Voltage Recovery," no. June, 2009.

[35] B. R. Williams, W. R. Schmus, and D. C. Dawson, "No. 3," vol. 7, no. 3, pp. 1173–1181, 1992.

[36] P. Irminger *et al.*, "Air conditioning stall phenomenon - Testing, model development, and

simulation," *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, pp. 1–8, 2012.

[37] M. Glavic *et al.*, "See it fast to keep calm: Real-time voltage control under stressed conditions," *IEEE Power Energy Mag.*, vol. 10, no. 4, pp. 43–55, 2012.

[38] J. Eto, "2014 FIDVR Events Analysis on Valley Distribution Circuits," 2014.

[39] L. Y. Taylor and S. Member, "Transmission Voltage Recovery Following a Fault Event," *IEEE Power Eng. Soc. Summer Meet.*, vol. 00, no. c, pp. 1–6, 2000.

[40] I. A. Hiskens, J. Alseddiqui, and S. Member, "Sensitivity , Approximation , and Uncertainty in Power System Dynamic Simulation," vol. 21, no. 4, pp. 1808–1820, 2006.

[41] M. Paramasivam, A. Salloum, V. Ajjarapu, V. Vittal, N. B. Bhatt, and S. Liu, "Dynamic optimization based reactive power planning to mitigate slow voltage recovery and short term voltage instability," *IEEE Trans. Power Syst.*, vol. 28, no. 4, p. 3865, 2013.

[42] A. P. S. Meliopoulos, G. Cokkinides, and G. Stefopoulos, "and Dynamic VAR Sources," pp. 124–131, 2006.

[43] Y. Xu, Z. Y. Dong, K. Meng, W. F. Yao, R. Zhang, and K. P. Wong, "Multi-objective dynamic VAR planning against short-term voltage instability using a decomposition-based evolutionary algorithm," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2813–2822, 2014.

[44] H. Bai and V. Ajjarapu, "A novel online load shedding strategy for mitigating fault-induced delayed voltage recovery," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 294–304, 2011.

[45] Y. Dong, X. Xie, K. Wang, B. Zhou, and Q. Jiang, "An Emergency-Demand-Response Based under Speed Load Shedding Scheme to Improve Short-Term Voltage Stability," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3726–3735, 2017.

[46] Y. Zhang, Y. Xu, Z. Y. Dong, and P. Zhang, "Real-Time Assessment of Fault-Induced Delayed Voltage Recovery: A Probabilistic Self-Adaptive Data-driven Method," *IEEE*

Trans. Smart Grid, vol. 3053, no. c, pp. 1–10, 2018.

[47] F. Li *et al.*, "Smart transmission grid: Vision and framework," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 168–177, 2010.

[48] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, "Real-time monitoring of short-term voltage stability using PMU data," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3702–3711, 2013.

[49] Z. Y. Dong, Y. Xu, P. Zhang, and K. P. Wong, "Using is to assess an electric power system's real-time stability," *IEEE Intell. Syst.*, vol. 28, no. 4, pp. 60–66, 2013.

[50] R. Diao *et al.*, "Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 832–839, 2009.

[51] Y. Xu *et al.*, "Assessing Short-Term Voltage Stability of Electric Power Systems by a Hierarchical Intelligent System," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 8, pp. 1686–1696, 2016.

[52] I. Kamwa, S. R. Samantaray, and G. Joós, "Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 258–270, 2009.

[53] I. Kamwa, S. R. Samantaray, and G. Joós, "On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 152–161, 2012.

[54] L. Zhu, C. Lu, and Y. Sun, "Time Series Shapelet Classification Based Online Short-Term Voltage Stability Assessment," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1430–1439, 2016.

[55] L. Zhu, C. Lu, Z. Y. Dong, and C. Hong, "Imbalance Learning Machine-Based Power

System Short-Term Voltage Stability Assessment," *IEEE Trans. Ind. Informatics*, vol. 13, no. 5, pp. 2533–2543, 2017.

[56] W. Wang, M. Diaz-Aguilo, K. Ben Mak, F. De Leon, D. Czarkowski, and R. E. Uosef, "Time Series Power Flow Framework for the Analysis of FIDVR Using Linear Regression," *IEEE Trans. Power Deliv.*, vol. 8977, no. c, pp. 1–10, 2018.

[57] Y. Zhang, Y. Xu, Z. Y. Dong, and R. Zhang, "A Hierarchical Self-Adaptive Data-Analytics Method for Power System Short-term Voltage Stability Assessment," *IEEE Trans. Ind. Informatics*, vol. 3203, no. c, 2018.

[58] J. Eto, "2012 FIDVR Events Analysis," 2012.

[59] J. Eto, "2013 FIDVR Events Analysis on Valley Distribution Circuits," 2013.

[60] L. B. N. Laboratory, "Load modeling transmission research," *Presentations*, 2004.

[61] M. Hubert, P. Rousseeuw, and P. Segaert, "Multivariate Functional Outlier Detection."

[62] P. Norvig, "Machine learning for programming," in *Conference on Systems, Programming, and Applications: Software for Humanity, {SPLASH} '14, Portland, OR, USA, October 20-24, 2014 - Companion Volume*, 2014, p. 3.

[63] P. Norvig, "Machine Learning for Learning at Scale," in *Proceedings of the Second {ACM} Conference on Learning @ Scale, L@S 2015, Vancouver, BC, Canada, March 14 - 18, 2015*, 2015, p. 215.

[64] B. S., "Brandt S. (1999) Linear and Polynomial Regression. In: Data Analysis. Springer, New York, NY," *Springer, New York, NY*, no. 1, 1999.

[65] A. D. Aczel, *Complete Business Statistics*. 1989.

[66] R. J. Makridakis, S., Wheelwright, S. C., Hyndman, *Forecasting methods and applications*,. New York, Wiley, 1998.

[67] E. Ostertagová, “Applied Statistic (in Slovak). Elfa Košice, Slovakia,” no. 978-80-8086-171–1, p. 161, 2011.

[68] Y. Xu, S. Member, Z. Y. Dong, S. Member, and J. H. Zhao, “A Reliable Intelligent System for Real-Time Dynamic Security Assessment of Power Systems,” *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1253–1263, 2012.

[69] A. Khamis, Y. Xu, Z. Y. Dong, S. Member, and R. Zhang, “Faster Detection of Microgrid Islanding Events Using an Adaptive Ensemble Classifier,” *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1889–1899, 2018.

[70] P. Norvig, “Applying machine learning to programs,” in *Companion to the Proceedings of the 11th International Symposium on Open Collaboration, OpenSym 2015, San Francisco, California, USA, August 19-21, 2015*, 2015, p. 11:1.

[71] P. Norvig, “Reviews of ‘Machine Learning by Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell’, Tioga publishing company, 1983, {ISBN} 0-935382-05-4,” *{SIGART} Newsl.*, vol. 90, pp. 16–17, 1984.

[72] N. Amjady and S. F. Majedi, “Transient Stability Prediction by a Hybrid Intelligent System,” vol. 22, no. 3, pp. 1275–1283, 2007.

[73] I. Kamwa, S. R. Samantaray, and G. Joos, “Catastrophe Predictors From Ensemble Decision-Tree Learning of Wide-Area Severity Indices,” *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 144–158, 2010.

[74] M. He, S. Member, V. Vittal, and J. Zhang, “Online Dynamic Security Assessment With Missing PMU Measurements : A Data Mining Approach,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1969–1977, 2013.

[75] M. He, S. Member, J. Zhang, and V. Vittal, “Robust Online Dynamic Security Assessment

Using Adaptive Ensemble Decision-Tree Learning," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4089–4098, 2013.

[76] S. J. Russell and P. Norvig, *Artificial intelligence - a modern approach, 2nd Edition*. Prentice Hall, 2003.

[77] and A. E.-S. Y. Mansour, E. Vaahedi, "Dynamic Security Contingency Screening and Ranking Using Neural Networks," *Proc. ASME Fluids Eng. Div. Summer Meet. 2006, FEDSM2006*, vol. 2 FORUMS, no. 4, pp. 427–432, 2006.

[78] L. S. Moulin, A. P. A. daSilva, M. A. El-Sharkawi, and R. J. MarksII, "Support Vector Machines for Transient Stability Analysis of Large-Scale Power Systems," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 818–825, 2004.

[79] L. Breiman, "Random Forests," pp. 1–33, 2001.

[80] G. Biau, "Analysis of a Random Forests Model," pp. 1–40, 2010.

[81] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning," *Math. Intell.*, pp. 369–370, 2008.

[82] J. Ma *et al.*, "Reducing Identified Parameters of Measurement-Based Composite Load Model," vol. 23, no. 1, pp. 76–83, 2008.

[83] J. K. Kim *et al.*, "Fast and reliable estimation of composite load model parameters using analytical similarity of parameter sensitivity," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 663–671, 2016.

[84] S. Guo and T. J. Overbye, "Parameter estimation of a complex load model using phasor measurements," *2012 IEEE Power Energy Conf. Illinois, PECI 2012*, pp. 1–6, 2012.

[85] D. Kosterev *et al.*, "Load Modeling in Power System Studies : WECC Progress Update DELAYED RECOVERY," *Power*, no. point 3, pp. 1–8, 1990.

[86] B. Choi and H. Chiang, “Multiple Solutions and Plateau Phenomenon in Measurement-Based Load Model Development : Issues and Suggestions,” vol. 24, no. 2, pp. 824–831, 2009.

[87] D. Kosterev and A. Meklin, “Load modeling in WECC,” *2006 IEEE PES Power Syst. Conf. Expo. PSCE 2006 - Proc.*, pp. 576–581, 2006.