Artificial Intelligence-Based Methods for Power System Security Assessment with Limited Dataset

by

Hoi Andy LAM

B.Eng. M.E.

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

School of Computer Science

The University of Sydney

Sydney, NSW, Australia

Year of Award 2023

Declaration

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Hoi Andy Lam
4 March 2023

Acknowledgements

I would like to express my sincerest gratitude to Prof. Zhao Yang Dong for his motivation and support along the long journey of my doctoral research. Prof. Dong left the university in 2017, but he continues to share his excellent experience and wisdom to guide me in delivering research outcomes to PhD standards.

I also would like to express gratitude to Prof. Masahiro Takatsuka for his support and assistance during my PhD candidate when I transferred to School of Computer Science.

I appreciate the faculty and my colleagues organising the independent research progress review and other support.

Finally, thanks to my family for their support during this study journey.

Hoi Andy Lam
Sydney, NSW, Australia
4 March 2023

List of Publications

Publications of journal paper relevant to the thesis

Chapter 4 of this thesis is published as H. A. Lam and Z. Y. Dong, "Transfer learning based dynamic security assessment," IET Gener. Transm. Distrib., Mar. 2021, doi: 10.1049/gtd2.12181.

I contributed to the activities, including conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, writing-original draft, writing-review & editing.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Hoi Andy Lam
4 March 2023

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Journal paper under review

H. A. Lam and Z. Y. Dong, "Transfer Learning Approach for NILM using GAN," IET Gener. Transm. Distrib.,

Abstract

Climate change, caused by the gradual rise in energy demand, has been a critical issue in recent years. Different measures have been adopted to tackle these challenges. For example, distributed generation based on renewable energy sources provides an alternative solution to reduce carbon emissions. The increased penetration of renewable types of distributed generation and the appearance of unconventional types of loads cause uncertainty in the reliable operation of electric networks. Due to the non-linear nature of the power system, artificial neural networks have been adopted to benefit the research on load modelling and online power system stability. This approach relies on communications and data to make the system work. However, sufficient training data may not be available in some situations. In addition, digital simulation, one of the essential tools for power system analysis, is heavily dependent on the accuracy of the load model. It requires understanding the concepts between the load, load model, and power system stability, and whether the traditional approach of load model accurately represents the load characteristic in the power system with high integration of renewable sources.

This thesis concerns the relationship between the load, load model, and power system stability. It investigates the possibility of developing a dynamic load model to represent the power system load characteristic during system faults when the power system operates at a high percentage of the power generation from wind farms, solar power, and vehicle-to-grid technology. Additionally, with artificial intelligence supporting the seamless integration of an increasingly distributed and multi-directional power system to unlock the vast potential of renewables, new approaches are proposed to improve the training performance for the applications of artificial neural networks in non-intrusive load monitoring and dynamic security assessment.

An improved hybrid load model is proposed to represent the load characteristics in the above power system operation. Genetic algorithms and the multi-curve identification method are applied to determine the parameters of the load model, aiming to minimize the error between the estimated and measured values. The results indicate that the proposed hybrid load model has a reasonably low fitting error to represent the load dynamics.

In addition, new approaches are proposed to tackle the challenges posed by limited data when training artificial neural networks (ANNs) for their application in power systems. The knowledge transfer approach is utilized to support the ANN training to generate synthetic data for non-intrusive load monitoring. The results indicate that this approach improves the issue of mode collapse and reduces the need for lengthy training iterations, making the ANN effective for generating synthetic data from limited data. Moreover, the knowledge transfer approach also supports ANN training with limited data for dynamic security assessment. Kernel principal component analysis is employed to eliminate the dimensionality reduction step. The results indicate an improvement in the training performance.

List of Figures

List of Tables

Contents

1 Introduction

Climate change caused by the gradual rise in energy demand has been a critical issue in recent years. According to the International Energy Outlook 2019 (IEO2019) from the U.S. Energy Information Administration (EIA), it is predicted that world energy consumption will grow by almost 50% from 2018 to 2050 [1]. In response to climate change, global leaders have set a Mandatory Renewable Energy Target (MRET) to increase the supply of green electricity from distributed renewable sources. Different measures have been adopted to tackle this challenge. For instance, distributed generation (DG) based on renewable energy sources offers an alternative solution to reduce carbon emissions. Even a 15 per cent of the reduction in electricity use among homes can generate substantial positive effects on environmental protection in the United States. This energy consumption is equivalent to 81.3 million tons of coal per year [2]. The increased penetration of renewable types (mostly stochastic) of distributed generation (wind farms and solar power farms) and the appearance of such unconventional types of loads (power electronics-based or interfaced through power electronics, e.g. electric vehicles EV) may pose important challenges for ensuring the safe and secure operation of power grids. It causes uncertainty in the reliable operation of electric networks [3].

In addition, the meteorological variables are strongly influenced by climate conditions, which in turn affect the parameter values of load model and power system stability. According to [4], relying on parameter values from the literature can result in substantial errors due to their association with loads in specific regions of the same country, which may have different load compositions. Furthermore, these parameters fail to differentiate between various time periods throughout the day and different days of the week. In [5], the author calculated the parameters of an exponential model to determine the mean values and relative

standard deviation for different times of the day and weekdays, and then compared them to the parameters obtained from existing literature. This demonstrates a basic compliance. Additionally, many sources do provide distinctions in the parameters based on seasons. The results indicate that it is possible to determine the parameters of exponential load models through high-resolution continuous measurements in a low voltage grid.

In [6], a measurement-based procedure was presented by the author to determine parameter values for a static exponential load model applicable to administrative buildings. The typical parameter values for this load model are obtained from long-term power measurements taken at the low voltage sides of the secondary 10 kV/0.4 kV transformer substations. The analysis is built upon the procedure introduced in [4] for identifying load model parameters. Average parameter values of the load model are determined for different seasons, specific days of the week, and particular time intervals within those days. Additionally, the investigation explores the possibility of establishing relationships between load model parameters and ambient temperature, considering that the ambient temperature varies based on season and time of day, rather than the day of the week. The study revealed that the variations in load model parameters during characteristic days within each season can be effectively represented using polynomial fitting functions obtained through the least squares method [7]. In addition, it was observed that incorporating the established relationships between load parameters and ambient temperature reduces the number of required identified environmental parameters for approximating the load modelling of examined demands, or the load of a similar composition to administrative buildings.

A study (referred to as [8]) was undertaken to examine the relationship between seasonal fluctuations in electricity consumption in the commercial and residential sectors based on key climatic variables in Hong Kong. Principal component analysis was utilised to explore this

connection, along with an investigation into yearly climate variations. Analysing the energy consumption trends in Hong Kong between 1979 and 2006, it was found that the residential and commercial sectors, the largest end-users of electricity, experienced average annual growth rates of 5.9% and 7.4%, respectively. The monthly consumption patterns in these sectors exhibited clear seasonal variations primarily influenced by changes in demand for air-conditioning loads, which, in turn, were influenced by prevailing weather conditions. Principal component analysis was employed to analyse five major meteorological variables, leading to the identification of two principal components as indicators of weather variations. Multiple regression techniques were utilised to correlate the normalised electricity consumption with these principal components for both the residential and commercial sectors. The study utilised 28 years of long-term weather data (1979-2006) obtained from the local observatory for the analysis. The results suggest that the regression models provided highly accurate predictions of annual electricity usage, typically within a few percentage points. However, individual monthly estimates could deviate by as much as 24%. The seasonal fluctuations in monthly electricity consumption within the commercial sector exhibited a closer alignment with the principal components compared to the residential sector. Furthermore, the analysis revealed a slight upward trend in the identified climatic indicators over the 28-year period, suggesting a gradual and subtle shift in climatic conditions.

In [9], a new method was introduced to estimate the parameters of a static exponential load model by utilising non-electrical variables, specifically meteorological variables. This approach also examines the feasibility of using meteorological variables to estimate the load model parameters for shorter time periods such as days of the week and days in different seasons. The proposed model utilises five meteorological variables, namely, ambient temperature, relative humidity, atmospheric pressure, wind speed, and wind direction, as

independent variables, and the load model parameters as dependent variables. The procedure is based on the assumption of a linear relationship between the voltage exponents of exponential load models and the values of meteorological variables during the time period when the related electrical variables are measured. Once the datasets of meteorological variable values and load model parameter values are obtained, a multiple linear regression [7, 10-12] is employed to determine the parameters of this linear model and their correlation coefficients. The analysis involves comparing the voltage exponents of the load model obtained using all five meteorological variables with progressively omitting each variable that has the lowest influence. This process allows for the adoption of simplified linear relationships that exclude dependencies on these low-influence variables.

To assess the relative influence of meteorological variables on static exponential load model parameters, partial flexibility coefficients [10] are computed. These coefficients allow for a comparison of the impact of different meteorological variables included in the linear models, separately for both active and reactive power models. The variable with the smallest absolute value of the partial flexibility coefficient is then removed from the model. This process is repeated until only one meteorological variable remains, ensuring the results of the model have the highest accuracy. Additionally, the existence of mutual correlations between the meteorological variables is examined. In order to identify multicollinearity among the meteorological variables, variance inflation factors (VIFs) are calculated for each pair of independent variables. Multicollinearity strongly exists if the values are greater than 4, especially those exceeding 10. In such instances, it is advisable to remove one of the two independent variables involved in the multicollinearity and re-examine the VIF values in the analysis. The independent variable that takes part in the multicollinearity with a smaller absolute value of partial flexibility coefficient is eliminated [10, 11].

In order to assess the accuracy of estimating static exponential load model parameters using meteorological variables, comparisons were made between the estimated parameter values and the corresponding average voltage exponents obtained from electrical variables, as well as the values derived from polynomial fitting curves in [6], taking into account different seasons, days of the week, and day intervals. The findings reveal that similar results are obtained when comparing the average voltage exponents values obtained from electrical variables and the values from the fitting curves with the estimated values derived from the application of five meteorological variables. Based on these results, it is recommended to utilise the model with the highest accuracy, which includes all five meteorological variables. However, for rough estimations of the voltage exponents of reactive power, it is possible to use a reduced number of input variables, including even just one variable. Furthermore, the aforementioned procedure was repeated by employing linear regressions for various seasons, using datasets consisting of average load model parameter values and average ambient temperatures. The author concluded that by solely considering ambient temperature, it is possible to estimate the voltage exponents of both active and reactive power when the required electrical variables for exponential load model parameter identification are unavailable.

The key challenges facing the power system today covers a wide range, for example:

- The transition towards more inverter-based resources (IBRs).
- The reduction of system inertia and system strength due to increasing IBRs and decommissioning of conventional fossil fuel fired synchronous generators.
- The lack of reliable, and updated system load models to consider the impacts of such transition.

Power systems study helps to maximise the power system availability by improving its robust performance under all operating conditions. Power blackouts may mean major losses to the power utility if improper loading conditions are ignored [13]. So, it is crucial to ensure the reliability and safety of the power system operation. The stability margins are identified through static analysis, which also investigates the impact factors on stability in different system conditions [14, 15]. However, the dynamic load model is required to analyse the short-term stability of low inertia facilities and fast-acting induction motors [16]. The power system transients are often simulated to check the system stability through a set of differential-algebraic equations (DAEs) if the constrained quantities specified in the algebraic equations have complied. The load flow study is used to analyse the loading performance of the electrical network under a steady state condition. Stability studies, e.g. frequency stability, transient stability, voltage stability, etc. also ensure the system stability during power system faults and after fault clearance. Small disturbance stability evaluates the power system performance when the power system is subject to small disturbances.

However, digital simulation, one of the essential tools for power system analysis, is heavily dependent on the performance of the load model. For example, the large-scale blackout of WSCC 1994 demands a high accuracy load model to simulate the practical running condition of the power system [17, 18]. The load characteristics are represented as an algebraic function of the frequency and bus voltage magnitude at any instant of time to approximate the system load behaviours. These characteristics also significantly influence the deceleration or acceleration rate of the individual generators during system faults [19, 20].

With the high integration of renewable energy in the power system and the active response of the demand side, these uncertainties bring unprecedented complexity to power system stability. For example, the power utility has an industrial practice for using a load model with

clear physical interpretations or reliable mathematical formulas. It may cause a challenge to such a load model to represent the load characteristics under system fault in the above operation scenarios, e.g. the power system operates at high integration of wind farms, solar power, vehicle to grid (V2G), etc. In addition, the conventional transient stability assessment (TSA) may not respond fast enough for a reliable result.

Due to the non-linear nature of the power system, artificial neural networks (ANN) have been adopted to benefit the research on load modelling and pattern recognition for online power system stability. ANNs techniques have also been successfully applied to power system problems regarding classification or estimation for fast and reliable results. An approximation problem can be solved through the ANN learning process to find a solution surface proportioned by the best adaptation of the training parameters, with the criterion of the best adaptation measured by some statistical methods.

As the smart grid technologies with advanced metering infrastructure are continued to deploy across the globe, ANN also plays a key role in non-Intrusive load monitoring (NILM) in measuring the detailed consumption of individual appliances from the aggregated power signal. It enhances the power utility to operate in demand-side management (DSM). In addition, the increasing deployment of wide-area monitoring systems (WAMS) and phasor measurement units (PMUs) makes the ANN-based online dynamic security assessment (DSA) important in early unstable event detection on power system security. It can provide an effective and timely system security assessment for fast evaluation time and high accuracy of results in real-time operation.

The ANN is hungry for labelled data, which may not be available in some situations, e.g. the power system is under topology changes. It may also be costly to obtain the labelled data. The insufficient labelled data on the training set has a negative impact on the supervised

learning algorithm on the testing set. In general, small datasets may lead to overfitting. A dataset with a wider variety can benefit the training on the deep learning network.

My PhD research describes how transfer learning is successfully employed to shorten the training time for online DSA. It also helps to improve the validation accuracy if the training dataset of the scratch ANN model is insufficient. Besides, a new approach of using the densely connected convolutional network with kernel principal component analysis (KPCA) is proposed to eliminate the traditional dimensionality reduction step. In addition, transfer learning is successfully applied to generative adversarial nets (GAN) to generate sufficient synthetic data for the training set on supervised learning algorithms of the NILM. It improves the quality of synthetic datasets and effectively tackles the challenge of mode collapse from GAN during training on limited data. The number of training iterations on GAN is also significantly reduced. Besides, the synthetic data improves the training performance of the scratch ANN model if the training dataset is insufficient.

The minimum size of a limited dataset depends on the specific AI model used by the power utility for training new operational scenarios. Since different AI models have varying numbers of parameters, these models require different minimum sizes of training datasets. Additionally, there is no available metric showing that the amount of data is sufficient to train a neural network in the relevant literature. It is evident that the power utility will conduct some running tests to determine the appropriate sample size for limited datasets based on their AI model. The findings from these tests can provide an answer to the question: Will performance continue to improve as the training data grows?

In [21], it proposed a similar idea by applying Convolutional Neural Networks (CNN) to classify medical images. In the study, the CNN was trained using six different sizes of training datasets (5, 10, 20, 50, 100, and 200) and subsequently tested with a total of 6000

medical images. The learning curve approach was employed to predict the classification accuracy at a given training sample sizes. The author utilizes this approach to determine the required size of the training dataset in order to achieve a specific classification accuracy.

According to [22], a rule of thumb for deep learning applications in computer vision suggests a minimum sample size of 1,000 images per class for image classification training. On the other hand, referring to [23], the author utilizes a limited data approach with 50 training samples per class.

1.1 Structure of the Thesis

Chapter 1. It is an introduction to discuss the challenge of increasing the deployment of renewable energy sources to the power supply system. As more renewable energy sources are deployed in response to climate change, these renewable energy sources and unconventional types of loads cause the uncertain applicability of the conventional load model to the power system stability. In addition, the ANN-based online DSA can provide a fast evaluation with high accuracy, subject to the condition that the AI model is well trained with sufficient data. Small datasets may lead to overfitting. The transfer learning technique can improve the model training for online DSA or generate synthetic data to support the model training for NILM.

Chapter 2. The load model and stability concepts are discussed. It includes the static load model, dynamic load model, composite load model, frequency stability, transient stability, small-signal stability, and voltage stability in Sections 2.1 and 2.2. Section 2.3 describes how the load model affects system stability. The model concept of system stability with artificial intelligence is discussed in Section 2.4.

Chapter 3. This chapter is extracted from a research paper. It is research to demonstrate the concept of load model in an application to the high-level V2G environment. It supposes

sufficient electric vehicles to operate as backup energy sources under an effective control mechanism to discharge battery energy into the power system. In that case, it facilitates a shift in the electric energy demand, shaves the peak, and serves as a distributed energy storage device. Also, it can improve the stability of the power grid. This paper proposes a hybrid load model to represent the dynamic load characteristics when the power system operates at a high percentage of the power generation from wind farms, solar power and V2G. This hybrid load model, which consists of components of physical meaning or reliable formulas, may have some uses in the specific configuration of the power systems in the above scenario.

Chapter 4. It is research to generate synthetic data of the NILM by generative adversarial nets for the training set on supervised learning algorithms. As NILM can recognise different appliances through signal processing and artificial neural network, it is an effective real-time approach to measuring the load. However, the labelled data may not be available in some situations to train the ANN, causing a negative effect on NILM performance.

Chapter 5. It is research to describe how transfer learning successfully shortens the training time for online DSA. The online DSA is important to predict system stability in a post-fault operation, and the timing of the DSA to detect an unstable event in the power system is critical to emergency remedial control action. Besides, a new approach of using the densely connected convolutional network with kernel principal component analysis is proposed to eliminate the traditional step of dimensionality reduction.

Chapter 6. The research is concluded. It also gives the prospects for future work.

1.2 Contribution of the Thesis

The main contributions of this thesis are listed as follows:

In Chapter 3: A hybrid load model on V2G environment

To the best of my knowledge, similar applications have not been reported in the research of load modelling. The key contribution is:

- A dynamic hybrid load model is developed through two different load models to represent the active and reactive power responses during power system fault when the power system operates at a high percentage of the power generation from wind farms, solar power and V2G. The two load models consist of components with physical meaning or reliable mathematical formulas to demonstrate the load model concepts.

In Chapter 4: Transfer Learning Approach for Non-Intrusive Load Monitoring using Generative Adversarial Networks

To the best of my knowledge, similar applications have not been reported in the field of NILM. The key contributions are:

- The synthetic data are generated from the GAN to tackle the challenge of limited data on NILM.
- The transfer learning technique is applied to the GAN to generate synthetic data from limited data.
- The transfer learning technique is applied to reduce the number of training iterations on the GAN.
- The challenge of mode collapse during the training of GAN on the limited data is effectively tackled by applying the transfer learning technique.

- The current, active power, reactive power, and harmonics are combined to form the two-dimensional (2D) data input used for GAN, DenseNet and ResNet.
- The combined pattern of the electrical features of the current, active power, reactive power, and harmonics are proposed to perform the NILM.
- A new application framework for NILM is proposed.

In Chapter 5: Transfer Learning Based Dynamic Security Assessment

To the best of my knowledge, similar applications have not been reported in the field of DSA. The key contributions are:

- The non-electrical feature dataset on ImageNet is used to improve the performance of an AI model training on the power system.
- The DenseNet with KPCA is proposed to eliminate the traditional step of dimensionality reduction.
- The feature extraction technique of transfer learning is applied to shorten the training completion time on DenseNet if the training dataset is too large.
- The finetuning technique of transfer learning is applied to improve the validation accuracy on DenseNet if the training dataset is too small.
- The 2D histogram is used to represent the feature data of the power system for the training of the AI model.

2 Load Models and Power System Stability

An electric power system refers to a network of electrical components employed for the purpose of generating, transmitting, and utilizing electric power. The electrical grid serves as a prominent illustration of such a system, as shown in Fig. 2- 1, as it facilitates the distribution of electricity to vast end users. The electrical grid is composed of a distribution system that effectively delivers power to nearby commercial users, residential houses, and industries, a transmission system that transmits the generated power from its sources to the designated load centres, and power generators responsible for electricity supply. The net amount of power consumed by the loads on a power system must equal the net amount of power produced by the supplies, minus the power lost in transmission. Supervisory control and data acquisition (SCADA) is commonly used to serve multiple functions such as switching on/off generators, controlling power output from generators, and switching in or out system elements for maintenance. The increasing deployment of wide-area monitoring systems (WAMS) and phasor measurement units (PMUs) also effectively support online monitoring in real-time operation.

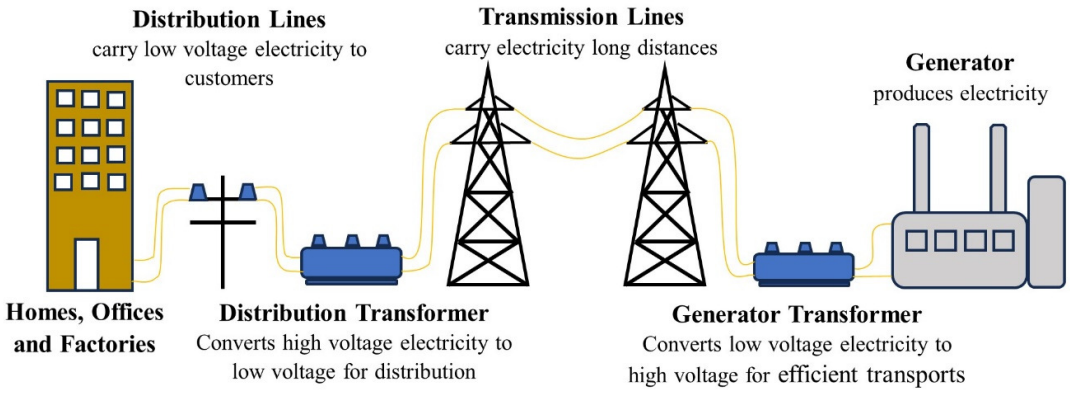

Fig. 2- 1 Transport of electricity [24]

Reference [25] is summarised as follows. In the transmission system, AC power can be converted to DC power by power electronics to make it frequency-independent for network interconnection, such as high-voltage direct current (HVDC) links, in order to improve system stability. Furthermore, transmitting DC power over long distances at very high voltages is more economical [26]. However, AC power, which can be generated by brushless generators using different energy sources such as hydropower, fossil fuel-based power, nuclear energy, and solar thermal power, is transmitted over long distances with less loss at higher voltages by stepping up the voltage at power generation sources and then stepping it down near the load. This enables the delivery of electricity through different voltage levels in the power network.

According to [27], the transmission system operates at voltages typically ranging from 66 kV to 500 kV. Transmission lines carry bulk energy at high voltages (generally exceeding 66 kV) from power plants to key distribution points across NSW and the ACT, and distribution network takes electricity from the transmission system distributes it to homes, businesses, and other end-users. Transmission lines often span vast distances and can include overhead lines, underground cables, and subsea cables where necessary.

In certain countries, electricity grid systems interconnect multiple generators operating at frequencies of 50 or 60 Hz in three phases, with a few exceptions such as gearless wind turbines or power systems with different frequencies linked by HVDC connections. All generators within a national grid rotate at sub-multiples of the same speed, generating electric current at the same frequency. Turbo generators are commonly employed for supplying alternating current power. For direct current power, photovoltaic cells, fuel cells, or batteries can be utilized.

According to [28], Ausgrid's medium/high voltage distribution network operates at several voltage ranges, typically 11kV, 22kV, 33kV, 66kV, and 132kV. The low voltage supply in the Australian Standard AS 60038 Standard Voltages falls within the voltage range of 230V, with a tolerance of +10%/-6% to account for voltage regulation within the network. The Australian Energy Market Operator (AEMO) also has electricity rules for the power systems in Australia, including the interconnected National Electricity Market (NEM) power systems, aiming to operate within a frequency range as close to 50 Hz as possible [29]. Larger centralized air conditioning systems often operate in three phases. Other appliances in residential use typically operate as single-phase systems, running at 50 Hz with a voltage of 230 volts.

The power utility is responsible for the operation of the power system network, with a primary objective of ensuring optimal power quality. This involves addressing sustained undervoltages, overvoltages, and deviations from the system frequency. Adverse effects on the power system load can arise from various incidents, including voltage sags, swells, dips, flicker, transient overvoltages, phase imbalance, poor power factor, and high-frequency noise. Additionally, reactive power, like real power, must be balanced and can be supplied by generators. Nevertheless, supplying reactive power via capacitors often proves to be more cost-effective [30].

In a typical AC power system, most of the load is inductive. Consequently, installing capacitors near the inductive loads often improves the power factor of the power system. Reactors are utilized to reduce overvoltage on long overhead transmission lines during lightly loaded conditions [31]. Also, series reactors can be used to limit fault currents. Small reactors are almost always installed in series with capacitors to mitigate the sudden surge of current associated with capacitor switching. Circuit breakers are used for switching reactors and

capacitors, resulting in substantial step changes in reactive power. One approach to address this problem is the use of static synchronous compensators, synchronous condensers, and static VAR compensators.

In high power applications, modern protective relays serve as application-specific computers that assess power system readings to determine whether to initiate a trip. Different relays are set for various protection schemes. The quality of circuit breakers is also higher, such as the utilization of sulfur hexafluoride (SF6) to fill the circuit breaker chamber for effective arc-quenching during switching. In lower voltage application, a residual current device (RCD) works by monitoring the active and neutral lines to trip the active line if the difference is higher than a specified condition [32]. Moulded case circuit breakers (MCCB) are also commonly used. Miniature circuit breakers are typically used if the power consumption is less than about 10 kW. Fuse stands as the quintessential protective device.

Power systems studies help maximise power system availability by improving its robust performance under all operating conditions. However, power blackouts can result in significant losses for the power utility if improper loading conditions are ignored [13]. Since load models are fundamental knowledge and critical to understanding the research problems in load modelling and stability, they are therefore essential to include and discuss in the following chapters.

2.1 Load Models

A composite model (CLODBL) used by PSS/E illustrates the concept of load model [33]. This load model was developed to capture the aggregate dynamic response of loads. The load model structure includes static load models, transformer saturation, discharge lighting, a small induction motor and a large induction motor. The percentage of each of these components in the load can be changed based on the load. The distribution transformer impedance can also be specified by the user. The basic structure of the model is shown in Figure 2- 2.

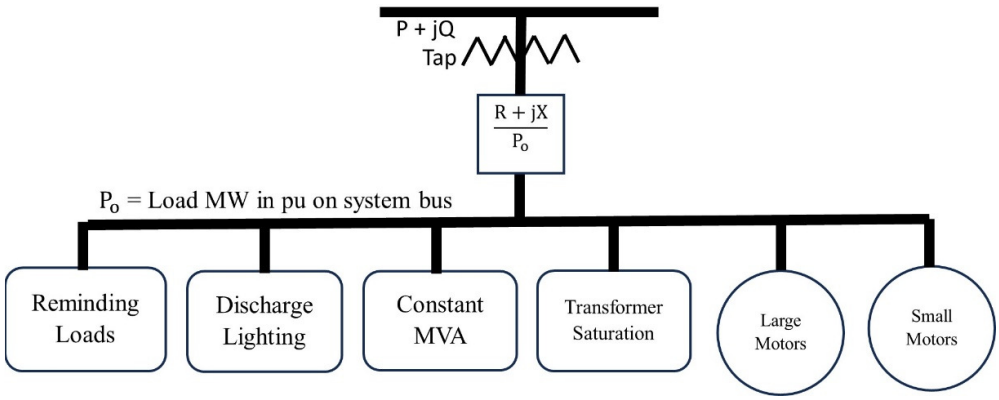

Fig. 2- 2 Schematics of composite model (CLODBL) model used by PSS/E [33]

Sections 2.1.1-2.1.3 are fundamental knowledge, critical to the understanding of the research problem in this Chapter. They are therefore very important to be included here.

In the planning stage, it is important for system planners to investigate the dynamic responses of generation equipment, active transmission equipment, and system load to severe disturbances in the power system. Thus, load modelling serves as a key function for time-domain dynamic simulations in power system planning studies. Commercial power system simulation tools are often utilized to examine the anticipated dynamic and steady-state

responses of the system to various critical contingencies and disturbances. Some applications of load models are summarized as follows:

Static load models are used to analyse steady-state operating conditions, determine voltage levels, power flows, and system losses. They are also utilized to forecast load parameters by incorporating statistical analysis of historical load data, such as meteorological variables, for different time periods, e.g., seasons, days of the week, and day intervals. The ZIP model is commonly used in both steady-state and dynamic studies [34, 35]. Although these static load models can also be used as approximations for dynamic loads, such as induction motors, their approximation is not always sufficient for certain studies, especially in stability studies. Therefore, they are often replaced by a dynamic load model.

Dynamic load models are used to analyse the dynamic behaviour of loads during transient events, such as faults or system disturbances. For example, in studies on voltage stability, the use of dynamic load models is essential for accurate representation [35].

The composite load model incorporates both the static load model and the dynamic load model. Research referenced in [36, 37] involved comparing simulation results of various load models under transient disturbances. The findings concluded that composite models offer more accurate responses. Composite load models are anticipated to be used in stability studies that demand a higher level of accuracy in representing system dynamics. Additionally, the combination of the ZIP model and the induction motor model, denoted as (ZIP + induction motor model), is widely utilized in current version of software tools for dynamic studies [35].

A survey study conducted on international practices concerning power system load modelling [38] indicates that approximately 70% of utilities worldwide solely rely on static

load models, while around 30% of utilities utilize some form of induction motor model to represent dynamic loads for power system stability studies. However, the authors were unable to draw definitive conclusions regarding the types of load models used in different dynamic studies due to the limited feedback received from respondents.

2.1.1 Static Model

In [34, 39], the static load model including the exponential load model and polynomial load model are summarized as follows.

Exponential Load Model:

The exponential load model is a generalized form of the polynomial model which the effect of frequency variation is neglected as shown in Eq. 2-1 and 2-2. This model relates the power and the voltage at a load bus by exponential equations as:

$$P_s = P_0 \left(\frac{V}{V_n}\right)^{k_{pu}} \tag{2-1}$$

$$Q_s = Q_0 \left(\frac{V}{V_n}\right)^{k_{qu}} \tag{2-2}$$

where P_s and Q_s are the active and reactive power of the load at voltage V, P_0 and Q_0 are the active and reactive power drawn by the load when it is connected to the rated voltage V_n, k_{pu} and k_{qu} represent the sensitivity of active and reactive power w.r.t. the variation of the voltage V. If the parameters k_{pu} and k_{qu} in Eq. 2-1 and 2-2 are set to 2, 1, and 0, the exponential model can be converted into constant impedance (CI), constant current (CC), and constant power (CP) load model.

If the effect of frequency variation is considered, the additional frequency exponential $(f/f_n)^{n_{pf}}$ and $(f/f_n)^{n_{qf}}$ are added to Eq. 2-1 and 2-2 to change the formulas into Eq. 2-3 and 2-4 respectively as:

$$P_s = P_0 \left(\frac{V}{V_n}\right)^{k_{pu}} \left(\frac{f}{f_n}\right)^{n_{pf}} \tag{2-3}$$

$$Q_s = Q_0 \left(\frac{V}{V_n}\right)^{k_{qu}} \left(\frac{f}{f_n}\right)^{n_{qf}} \tag{2-4}$$

where f and f_n are frequency and rated frequency, respectively.

Polynomial Load Model:

The Eq. 2-1 and 2-2 can also be extended to represent mixed loads in polynomial load model, e.g. the ZIP model. This model combines CI, CC, and CP components, as shown in Eq. 2-5 and 2-6 if the effect of frequency variation is neglected. It can be used for dynamic studies in power system stability and as constant power in steady-state for load flow calculations. The equations of Eq. 2-5 and 2-6 are given as:

$$P_s = P_{01}\left(\frac{V}{V_n}\right)^2 + P_{02}\left(\frac{V}{V_n}\right) + P_{03} \tag{2-5}$$

$$Q_s = Q_{01}\left(\frac{V}{V_n}\right)^2 + Q_{02}\left(\frac{V}{V_n}\right) + Q_{03} \tag{2-6}$$

where the coefficients, P_{0i} and Q_{0i}, i = 1, …, 3, describe the participation of the specific models including CI, CC, and CP respectively in the total load model, and

$$P_0 = P_{01} + P_{02} + P_{03} \tag{2-7}$$

$$Q_0 = Q_{01} + Q_{02} + Q_{03} \tag{2-8}$$

where P_0, Q_0 are the nominal active and reactive power respectively when $V = V_n$ in Eq. 2-5 and 2-6. The parameters P_0, Q_0 and V_n in Eq. 2-5 to 2-8 equal to the values at the initial system operating condition with the participation of P_{0i} and Q_{0i}, i = 1, …, 3 if these equations represent a bus load [40].

There is other form to include the frequency dependence by multiplying the voltage dependent load in Eq. 2-5 and 2-6 by the factor $1 + k_{pf}(f - f_0)\ and\ 1 + k_{qf}(f - f_0)$ respectively, where k_{pf} and k_{qf} are the frequency sensitivity factor of the load model to give Eq. 2-9 and 2-10.

$$P_s(V,f) = \left[P_{01}\left(\frac{V}{V_n}\right)^2 + P_{02}\left(\frac{V}{V_n}\right) + P_{03}\right](1 + k_{pf}(f - f_0)) \quad (2\text{-}9)$$

$$Q_s(V,f) = \left[Q_{01}\left(\frac{V}{V_n}\right)^2 + Q_{02}\left(\frac{V}{V_n}\right) + Q_{03}\right](1 + k_{qf}(f - f_0)) \quad (2\text{-}10)$$

In [39], it also gives the most complete static load model which is incorporated into a number of computer programs for power system computations as follows:

$$P_s(V,f) = P_0\left[P_{a1}\left(\frac{V}{V_n}\right)^{k_{pu1}}\left(1 + k_{pf1}(f - f_0)\right) + (1 - P_{a1})\left(\frac{V}{V_n}\right)^{k_{pu2}}\right] \quad (2\text{-}11)$$

$$Q_s(V,f) = P_0\left[Q_{a1}\left(\frac{V}{V_n}\right)^{k_{qu1}}\left(1 + k_{qf1}(f - f_0)\right) + (\frac{Q_0}{P_0} - Q_{a1})\left(\frac{V}{V_n}\right)^{k_{qu2}}\left(1 + k_{qf2}(f - f_0)\right)\right] \quad (2\text{-}12)$$

where V_n, P_0, and Q_0 are initial operating conditions of the aggregate load at a bus. P_{a1} is the frequency-dependent fraction of active power, $Q_{a1}$is the ratio of uncompensated reactive load to active power, k_{pf1} is the frequency sensitivity coefficients for active power load, k_{qf1} and k_{qf2} are the frequency sensitivity coefficients for uncompensated reactive power load and reactive power compensation respectively, k_{pu1} and k_{pu2} are voltage exponents for active power, k_{qu1} and k_{qu2} are voltage exponents for uncompensated reactive power load and reactive power compensation respectively.

2.1.2 Dynamic Model

Exponential Recovery Load Model:

In [41], it uses the dynamic load model to explain the exponential power recovery after a small step disturbance at the bus voltage as listed in the non-linear first-order equations Eq. 2-1 to 2-4. The exponential recovery load model was applied in tap-change action, heating load and approximation for induction motor load as follows [42]:

$$T_p \frac{dP_r}{dt} = -P_r + P_0 \left(\frac{V}{V_n}\right)^{k_{pus}} - P_0 \left(\frac{V}{V_n}\right)^{k_{put}} \quad (2\text{-}13)$$

$$P_l = P_r + P_0 \left(\frac{V}{V_n}\right)^{k_{put}} \quad (2\text{-}14)$$

$$T_q \frac{dQ_r}{dt} = -Q_r + Q_0 \left(\frac{V}{V_n}\right)^{k_{qus}} - Q_0 \left(\frac{V}{V_n}\right)^{k_{qut}} \quad (2\text{-}15)$$

$$Q_l = Q_r + Q_0 \left(\frac{V}{V_n}\right)^{k_{qut}} \quad (2\text{-}16)$$

where V_n, P_0, Q_0 are the initial values of bus voltage, active power and reactive power at the load bus respectively, P_r, Q_r are the state variables for real power recovery and reactive power recovery respectively, T_p, T_q are the real power and reactive power time constant of the exponential recovery response, k_{pus}, k_{qus} are exponents related to the steady state load response, k_{put}, k_{qut} are exponents related to the transient load response.

Induction Motor Load Model:

The reference [41] and [33] proposed to use the third-order induction motor equations for the representation of load dynamic behavior summarized as follows:

State equations:

$$\frac{dE'_q}{dt} = -\omega_0 s E'_d - \frac{1}{T_0'} E'_q + \frac{X-X'}{T_0'} i_d \quad (2\text{-}17)$$

$$\frac{dE'_d}{dt} = -\frac{1}{T_0'} E'_d + \omega_0 s E'_q - \frac{X-X'}{T_0'} i_q \quad (2\text{-}18)$$

$$\frac{d\omega}{dt} = -\frac{1}{2H}\left[\left(E'_d i_d + E'_q i_q\right) - T_L\right] \quad (2\text{-}19)$$

$$T_L = T_0(A\omega^2 + B\omega + C_o + D\omega^E) \quad (2\text{-}20)$$

$$C_o = 1 - A{\omega_o}^2 - B\omega_o - D{\omega_o}^E \quad (2\text{-}21)$$

$$i_d = \frac{1}{R_s^2 + X'^2}\left[R_s(U_d - E'_d) + X'\left(U_q - E'_q\right)\right] \quad (2\text{-}22)$$

$$i_q = \frac{1}{R_s^2 + X'^2}\left[R_s\left(U_q - E'_q\right) - X'(U_d - E'_d)\right] \quad (2\text{-}23)$$

$$T_0' = \frac{X_r + X_m}{\omega_o R_r} \quad (2\text{-}24)$$

$$X = X_s + X_m \quad (2\text{-}25)$$

$$X' = X_s + \frac{X_r X_m}{X_r + X_m} \quad (2\text{-}26)$$

Output equations:

$$P_m = \frac{X_{t1}}{G} U(U - E'_d) - \frac{X_{t2}}{G} UE'_q \quad (2\text{-}27)$$

$$Q_m = \frac{X_{t2}}{G} U(U - E'_d) + \frac{X_{t1}}{G} UE'_q \quad (2\text{-}28)$$

$$X_{t1} = \frac{R_s(X-X')}{{R_s}^2 + {X'}^2} \quad (2\text{-}29)$$

$$X_{t2} = \frac{X'(X-X')}{{R_s}^2 + {X'}^2} \quad (2\text{-}30)$$

$$G = X - X' \quad (2\text{-}31)$$

where E'_d, E'_q are the internal EMF on d-axis and q-axis respectively, I_d, I_q are the d-axis and q-axis stator currents, ω is the mechanical speed, R_s, X_s, X_m, R_r, X_r are the stator resistance, stator reactance, magnetizing reactance, rotor resistance and rotor reactance, H is inertia, ω_o is the synchronous speed, T_L, T_0 are the load torque and initial load torque respectively, all other variables are in per unit, A, B, C, D, E are load torque coefficients in term of motor speed.

Difference Equations Load Model:

Reference [33] also proposed the second order difference equations model which has no physical meaning to represent the dynamic load response summarized as follows:

$$\Delta P(k) = a_{p1}\Delta P(k-1) + a_{p2}\Delta P(k-2) + c_{p0}\Delta V(k) + c_{p1}\Delta V(k-1) + c_{p2}\Delta V(k-2) + g_{p0}\Delta f(k) + g_{p1}\Delta f(k-1) + g_{p2}\Delta f(k-2) \quad (2\text{-}32)$$

$$\Delta Q(k) = a_{q1}\Delta Q(k-1) + a_{q2}\Delta Q(k-2) + c_{q0}\Delta V(k) + c_{q1}\Delta V(k-1) + c_{q2}\Delta V(k-2) + g_{q0}\Delta f(k) + g_{q1}\Delta f(k-1) + g_{q2}\Delta f(k-2) \quad (2\text{-}33)$$

If the coefficients of g_{p0}, g_{p1}, g_{p2}, g_{q0}, g_{q1}, g_{q2} equal to zero, the frequency effect on load is ignored.

2.1.3 Composite Model

In [33], it proposed four type of the composite load model structure to represent the combination between the static load model and dynamic load model summarized as follows:

- ZIP + Induction Motor (3-state model)
- ZIP + Difference Equation (2nd order model)
- Exponential + Induction Motor (3-state model)
- Exponential + Difference Equation (2nd order model)

ZIP + Induction Motor (3-state model):

a (ZIP Model) + b (Induction Motor Model)

Refer to the Eq. 2-9 and 2-10 for the ZIP load model and Eq. 2-17 to 2-31 for the induction motor load model, and $a + \mathrm{b} = 1$.

ZIP: Refer to Eq. 2-9 and 2-10. Assume $k_{pf}\ and\ k_{qf} = 0$.

$$P_s = P_0\left[a_p\left(\frac{V}{V_n}\right)^2 + b_p\left(\frac{V}{V_n}\right) + c_p\right](1 + k_{pf}(f - f_0)) \tag{2-34}$$

$$Q_s = Q_0\left[a_q\left(\frac{V}{V_n}\right)^2 + b_q\left(\frac{V}{V_n}\right) + c_q\right](1 + k_{qf}(f - f_0)) \tag{2-35}$$

Induction Motor: Refer to Eq. 2-27 and 2-28.

$$P_m = \frac{\mathrm{X_{t1}}}{\mathrm{G}}\,\mathrm{U}(\mathrm{U} - \mathrm{E'_d}) - \frac{\mathrm{X_{t2}}}{\mathrm{G}}\,\mathrm{UE'_q} \tag{2-27}$$

$$Q_m = \frac{\mathrm{X_{t2}}}{\mathrm{G}}\,\mathrm{U}(\mathrm{U} - \mathrm{E'_d}) + \frac{\mathrm{X_{t1}}}{\mathrm{G}}\,\mathrm{UE'_q} \tag{2-28}$$

The percentage of induction motor in the composite load model is given as:

$$K_p = \frac{P_m}{P_T} \tag{2-36}$$

$$K_Q = \frac{Q_m}{Q_T} \tag{2-37}$$

The percentage for static part is given as:

$$P_s = P_T(1 - K_p) \tag{2-38}$$

$$Q_s = Q_T(1 - K_Q) \tag{2-39}$$

where P_T is total active power load, Q_T is the total reactive power load in the composite load model.

ZIP + Difference Equation (2nd order model)

a (ZIP Model) + b (Difference Equation)

Refer to the Eq. 2-34 and 2-35 for the ZIP load model and Eq. 2-32 and 2-33 for the second order difference equations load model, and $a + \mathrm{b} = 1$.

ZIP: Assume $k_{pf}\ and\ k_{qf} = 0$.

$$P_s = P_0\left[a_p\left(\frac{V}{V_n}\right)^2 + b_p\left(\frac{V}{V_n}\right) + c_p\right](1 + k_{pf}(f - f_0)) \tag{2-34}$$

$$Q_s = Q_0\left[a_q\left(\frac{V}{V_n}\right)^2 + b_q\left(\frac{V}{V_n}\right) + c_q\right](1 + k_{qf}(f - f_0)) \tag{2-35}$$

Difference Equations: Assume $g_{p0}, g_{p1}, g_{p2}, g_{q0}, g_{q1}, g_{q2} = 0$.

$$\Delta P(k) = a_{p1}\Delta P(k-1) + a_{p2}\Delta P(k-2) + c_{p0}\Delta V(k) + c_{p1}\Delta V(k-1) + c_{p2}\Delta V(k-2) + g_{p0}\Delta f(k) + g_{p1}\Delta f(k-1) + g_{p2}\Delta f(k-2) \tag{2-32}$$

$$\Delta Q(k) = a_{q1}\Delta Q(k-1) + a_{q2}\Delta Q(k-2) + c_{q0}\Delta V(k) + c_{q1}\Delta V(k-1) + c_{q2}\Delta V(k-2) + g_{q0}\Delta f(k) + g_{q1}\Delta f(k-1) + g_{q2}\Delta f(k-2) \tag{2-33}$$

The percentage of difference equations in the composite load model is given as:

$$K_p = \frac{\Delta P(k)}{P_T} \tag{2-40}$$

$$K_Q = \frac{\Delta Q(k)}{Q_T} \tag{2-41}$$

The percentage for static part is given as:

$$P_s = P_T(1 - K_p) \tag{2-38}$$

$$Q_s = Q_T(1 - K_Q) \tag{2-39}$$

where P_T is total active power load, Q_T is the total reactive power load in the composite load model.

Exponential + Induction Motor (3-state model)

a (Exponential Model) + b (Induction Motor Model)

Refer to the Eq. 2-3 and 2-4 for the exponential load model and Eq. 2-17 to 2-31 for the induction motor load model, and $a + \mathrm{b} = 1$.

Exponential: Assume $n_{pf}\ and\ n_{qf} = 0$.

$$P_s = P_0 \left(\frac{V}{V_n}\right)^{k_{pu}} \left(\frac{f}{f_n}\right)^{n_{pf}} \tag{2-3}$$

$$Q_s = Q_0 \left(\frac{V}{V_n}\right)^{k_{qu}} \left(\frac{f}{f_n}\right)^{n_{qf}} \tag{2-4}$$

Induction Motor: Refer to Eq. 2-27 and 2-28.

$$P_m = \frac{\mathrm{X_{t1}}}{\mathrm{G}}\,\mathrm{U}(\mathrm{U} - \mathrm{E'_d}) - \frac{\mathrm{X_{t2}}}{\mathrm{G}}\,\mathrm{UE'_q} \tag{2-27}$$

$$Q_m = \frac{\mathrm{X_{t2}}}{\mathrm{G}}\,\mathrm{U}(\mathrm{U} - \mathrm{E'_d}) + \frac{\mathrm{X_{t1}}}{\mathrm{G}}\,\mathrm{UE'_q} \tag{2-28}$$

The percentage of induction motor in the composite load model is given as:

$$K_p = \frac{P_m}{P_T} \tag{2-36}$$

$$K_Q = \frac{Q_m}{Q_T} \tag{2-37}$$

The percentage for static part is given as:

$$P_s = P_T(1 - K_p) \tag{2-38}$$

$$Q_s = Q_T(1 - K_Q) \tag{2-39}$$

where P_T is total active power load, Q_T is the total reactive power load in the composite load model.

Exponential + Difference Equation (2nd order model)

a (Exponential Model) + b (Difference Equation Model)

Refer to the Eq. 2-3 and 2-4 for the exponential load model and Eq. 2-32 and 2-33 for the second order difference equations load model, and $a + \mathrm{b} = 1$.

Exponential: Assume $n_{pf}\ and\ n_{qf} = 0$.

$$P_s = P_0 \left(\frac{V}{V_n}\right)^{k_{pu}} \left(\frac{f}{f_n}\right)^{n_{pf}} \tag{2-3}$$

$$Q_s = Q_0 \left(\frac{V}{V_n}\right)^{k_{qu}} \left(\frac{f}{f_n}\right)^{n_{qf}} \tag{2-4}$$

Difference Equations: Assume g_{p0}, g_{p1}, g_{p2}, g_{q0}, g_{q1}, $g_{q2} = 0$.

$$\Delta P(k) = a_{p1}\Delta P(k-1) + a_{p2}\Delta P(k-2) + c_{p0}\Delta V(k) + c_{p1}\Delta V(k-1) + c_{p2}\Delta V(k-2) + g_{p0}\Delta f(k) + g_{p1}\Delta f(k-1) + g_{p2}\Delta f(k-2) \tag{2-32}$$

$$\Delta Q(k) = a_{q1}\Delta Q(k-1) + a_{q2}\Delta Q(k-2) + c_{q0}\Delta V(k) + c_{q1}\Delta V(k-1) + c_{q2}\Delta V(k-2) + g_{q0}\Delta f(k) + g_{q1}\Delta f(k-1) + g_{q2}\Delta f(k-2) \tag{2-33}$$

The percentage of difference equations in the composite load model is given as:

$$K_p = \frac{\Delta P(k)}{P_T} \tag{2-40}$$

$$K_Q = \frac{\Delta Q(k)}{Q_T} \tag{2-41}$$

The percentage for static part is given as:

$$P_s = P_T(1 - K_p) \tag{2-38}$$

$$Q_s = Q_T(1 - K_Q) \tag{2-39}$$

where P_T is total active power load, Q_T is the total reactive power load in the composite load model.

2.2 Power System Stability

Power system stability may be defined as the ability of a power system to remain in an equilibrium operating point under normal operating conditions and to regain an acceptable state of equilibrium after being subjected to some form of disturbance [34]. In the parameter space, the stability domain is restricted by complicated surfaces, e.g. load flow feasibility and aperiodic and oscillatory boundaries. The effect of uncertainty on Net Transfer Capacity (NTC) and stability limitations needs to be considered in the power system planning and operation stages. It is crucial to obtain the system security measure by understanding different approaches to power system stability. As such, the frequency stability, transient stability, small-signal stability, and voltage stability as the common phenomenon in the power system are discussed in Sections 2.2.1-2.2.4, respectively.

As per [43], Figure [43, Fig. 9.11] illustrates an example of machine rotor angles plotted for both stable and unstable states. Under stable condition, the rotor angles exhibit oscillatory behaviour and eventually diminish to nothing due to damping effects. Additionally, the rate of change of the rotor angle will be zero at certain instants. Conversely, in an unstable condition, the angles increase with time.

In [34], the single-line diagram [34, Fig. E13.1] illustrates an example of transient stability for a generation plant comprising four 555 MVA, 24 kV, 60 Hz units that supply power to an infinite bus through two transmission circuits. A solid three-phase fault occurs on one of the transmission lines, and the fault is cleared by tripping the isolation devices. The curves depicting the rotor angle response for three different fault-clearing times (0.07 s, 0.086 s, and 0.087 s) with corresponding clearing angle values of 48.58 degrees, 52.04 degrees, and 52.30 degrees, respectively, as shown in Figure [34, Fig. E13.4], indicate that the system is stable with a fault-clearing time of 0.086 s and a clearing angle value of 52.04 degrees. However,

the system is unstable with a fault-clearing time of 0.087 s and a clearing angle value of 52.30 degrees. Therefore, the critical clearing time is determined to be 0.0865 s, and the critical clearing angle is 52.17 degrees.

In reference [44], a single-line diagram [44, Fig. 4] demonstrates another example of stability in a two-area four-machine power system. A three-phase short-circuit fault occurs at a bus positioned between the two power generation groups, with each group consisting of two generators. The fault is cleared within 100 ms by tripping one of the parallel transmission lines to disconnect the faulty part of the sectionized bus from the grid. Figure [44, Fig. 6] illustrates the dynamic oscillation process experienced by the power system between every two generators, and approximately 20.8 s later, the system enters the state of loss of synchronization.

2.2.1 Frequency Stability

Frequency stability can be defined according to the AEMO [45] as "the ability of a power system to maintain acceptable frequency following a contingency event. Typically, that contingency event causes an unbalance between generation and load in the power system. It depends on the ability of the power system to maintain or restore equilibrium between generation and load and recover the power system frequency to acceptable levels". Frequency instability means the power system is under a postfault uncontrolled sustained increase or decrease over time or sustained undamped oscillatory behaviour. So, the relevant frequency operating standards has been defined as an envelope of acceptable frequencies and the maximum durations that the power system should operate at those frequencies in various circumstances [46].

An imbalance between the load and generation causes the system to deviate from its rated frequency. The Rate of Change of Frequency (ROCOF) and frequency nadir [16, 47] are the corresponding stability indices to quantify the adverse effects of the system subjected to a severe imbalance between generation and demand. In [48], the frequency nadir represents frequency stability status as the minimum frequency value reached following a significant disturbance, the frequencies as 58 Hz in a 60 Hz system and 48 Hz in a 50 Hz system are considered. In [16], the definition of ROCOF is associated with frequency nadir, as shown in Eq. 2-44. According to [48], the design of the power system shall maintain a stable frequency status with a system imbalance split of at least 40 % between the load and the generation. Based on technology capabilities and users' operational experience, their facility shall withstand a ROCOF of at least 2 Hz/s in the power grid. The acceleration time constant $T_{A,SG,n}$ of each power plant is calculated with the nominal angular velocity Ω_N and the moment of inertia *J,* given as:

$$T_{A,SG,n} = \frac{J*\Omega^2{}_N}{P_{Nom,SG,n}} \tag{2-42}$$

where $P_{Nom,SG,n}$ is the nominal power of each synchronous generator in operation, and $T_{A,SG,n}$ is the acceleration time constant. The system inertia is contributed by the synchronous generators only. T_N is the acceleration time constant of the network to define the system inertia in the model as:

$$T_N = \frac{\sum_1^n T_{A,SG,n} * P_{Nom,SG,n}}{P_{Load}} \tag{2-43}$$

The ROCOF is inversely proportional to the overall system inertia. The smaller ROCOF means a more stable system. Its value also responds to system control during a disturbance. The ROCOF is expressed as:

$$ROCOF\ (Hz/S) = \frac{|f_{nadir} - f_{rated}|}{t_{nadir} - t_{disturbance}} \tag{2-44}$$

where f_{nadir}, f_{rated}, t_{nadir} and $t_{disturbance}$ represent frequency nadir, system nominal frequency, time of frequency nadir and disturbance occurrences, respectively [49]. The maximum ROCOF is expressed as follows:

$$ROCOF_{max} = \frac{\partial f}{\partial t_{max}} \tag{2-45}$$

$$= \frac{\Delta P_{Imbalance}}{P_{Load}} * \frac{f_o}{T_N} \tag{2-46}$$

where T_N and the imbalance $\Delta P_{Imbalance}$ related to P_{Load} define the maximum ROCOF, P_{Load} is the system load. The ROCOF can be reduced through the primary self-regulating control effect of loads and control reactions.

Reference [50] points out that the following factors determine the frequency stability if the power system is under disturbance, resulting in a sudden imbalance between generation and

load. For example, one of the large generators suddenly outages, and the load demand cannot instantaneously decrease. If the system spinning reserve is insufficient to provide the backup lost power to the load, the decline in frequency triggers the under-frequency load shedding schemes to protect the potential system blackout.

1. The amount of active power spinning reserve available across generating units.

2. The speed and reliability of the response of the turbine-governors.

It depends on the system spinning reserve and the control action of turbine-governor droop-response; the automatic generation controls respond by increasing the output power to halt the decline in frequency and resume the power system in an equilibrium operating point.

3. The total inertia of the system.

The capability of rotating machines and their associated load can be represented by the inertia as a parameter to store and inject kinetic energy into the system to reduce the transient deviation of frequency and ROCOF during a power system incident. The total inertia impacts the ROCOF in the initial frequency drop and the nature of the system response, i.e. the response for the very large system is overdamped, and the small system exhibits some oscillations in frequency due to lower inertia. In addition, the high ROCOF implies large transient voltage angle displacements in different grid zones, which may result in the tripping of the distance protection relay. Therefore, the power utility usually adopts an over-frequency gradient threshold to avoid the system operation traps into a no-return point of the frequency stability.

4. The amount of inherent load damping.

The power system loads decrease by a small percentage as the system frequency drops. It is the amount by which total system load varies as a frequency function based on the typical types of loads on the system, e.g. most mechanical load driven by the induction motor is a function of frequency. So, the mechanical load decreases if the system frequency drops. A typical number assumed in most simulations for these effects is between 1 to 1.5% change in load for a 1% change in frequency [51].

2.2.2 Transient Stability

The transient stability index represents the power system stability margin at any operating point. It quantifies the operating distance of any given state from its region of attraction (ROA) boundary in a power system.

In [34, 52-55], four transient stability indices, i.e. the rotor angle difference-based transient stability index (TSI), transient kinetic energy (TKE), transient potential energy (TPE-K), and rate of machine acceleration (ROMA) are utilised to estimate the operating distance of any given state from the points of critical operating state with high accuracy. Furthermore, the rotor angle δ w.r.t. rotor speed ω, in the phase portrait plane, determines the ROA. The above transient stability indices are mapped to ROA to establish the index-specific multidimensional stability boundaries in post-fault trajectory. These indices are listed as follows.

TSI is defined as:

$$\eta = \frac{\delta_T - \delta_{max}}{\delta_T + \delta_{max}} \tag{2-47}$$

$$If \begin{cases} \eta > 0 \rightarrow \text{stable} \\ \eta < 0 \rightarrow \text{unstable} \end{cases}$$

where δ_T is a predefined transient stability threshold and $\delta_T = 360^{\circ}$. The δ_{max} is the maximum angle separation of any two generators simultaneously in the post-fault response. The $\eta > 0$ and $\eta < 0$ correspond to stable and unstable conditions. The power system is more stable for higher positive TSI values.

TKE is defined as:

$$TKE = \sum_{i=1}^{n} \frac{1}{2} J_i \cdot \Delta\omega_i{}^2 \tag{2-48}$$

where $\Delta\omega_i$ and J_i are speed deviation and angular momentum of the rotor at a synchronous speed of *i*-th generator in a system with *n* machines, it is the transient kinetic energy of the generator immediately after the fault clearance.

TPE-K is defined as:

$$TPE - K = \sum_{i=1}^{n} P_{m_i}(\delta_{O_i} - \delta_{C_i}) \tag{2-49}$$

where δ_{O_i} and P_{m_i} are the rotor angle and mechanical power of *i*-th generator in a power system with *n* machines prior to a system fault, and the rotor angle changes to δ_{C_i} at the fault clearance time, it is the transient potential energy of the generator immediately after the fault clearance.

ROMA is defined as:

$$ROMA = \text{Max}\left(\frac{da_{PFT}}{dt}\right) \approx Max\left(\frac{\Delta a_{PFT}}{\Delta t}\right) \tag{2-50}$$

where Δt and Δa_{PFT} are the time between the fault occurrence and clearance and finite differences of rotor acceleration of all generators, respectively. The rate of rotor deceleration or acceleration of the generators immediately following a power system fault indicates its inertia and also the rate of the frequency deviation.

In [56-59], the indices of the transient rotor angle severity index (TRASI), maximum speed deviation, maximum rotor angle deviation and maximum acceleration were reviewed to assess power system transient stability. These indices, except TRASI, are specific to the individual generator, and the running characteristics of each generator are feedbacked. But they do not provide the specific instability value of the entire system. In [56], the rotor angle index (RAI), speed deviation index (SDI), acceleration index (ACI) and average acceleration

index (AAI) were used. The maximum speed deviation, SDI, maximum rotor angle deviation, RAI, maximum acceleration, ACI, AAI and TRASI are listed as follows.

Maximum rotor angle deviation and RAI are defined as:

$$\Delta\delta_i^{max}(\mathrm{t}) = \max(|\delta_i(\mathrm{t}) - \delta_i(0)|) \tag{2-51}$$

$$\mathrm{RAI}_j = \int_0^T \delta_j(t)dt \tag{2-52}$$

where $\delta_i(0)$ represents the initial rotor angle of each generator at time t =0, and this index refers to a specific generator on δ_i. The $\delta_\mathrm{j}(\mathrm{t})$ represents the rotor angle of the *j*-th generator at time instant t.

Maximum speed deviation and SDI are defined as:

$$\begin{aligned}\Delta\omega_\mathrm{i}^{\max}(\mathrm{t}) &= \max(|\omega_\mathrm{i}(\mathrm{t}) - \omega_\mathrm{i}(0)|) \\ &= \max(|\Delta\omega_\mathrm{i}(\mathrm{t})|)\end{aligned} \tag{2-53}$$

$$\mathrm{SDI}_j = \int_0^T \Delta\omega_j(t)dt \tag{2-54}$$

where $\omega_i(0)$ represents the initial rotor speed of each generator at time t = 0, and this index refers to a specific generator on ω_i. The $\Delta\omega_j(t)$ represents the speed deviation of the *j*-th generator at time instant t. Lower *T* values of RAI and SDI will cover less of the power system trajectory after the system in fault disturbance. The speed and accuracy should be compromised for the first-swing stability assessment.

Maximum acceleration deviation, ACI and AAI are defined as:

$$\begin{aligned}\Delta\mathrm{a}_\mathrm{i}^{\max} &= \max(|\mathrm{a}_\mathrm{i}(\mathrm{t}) - \mathrm{a}_\mathrm{i}(0)|) \\ &= \max(|\Delta\mathrm{a}_\mathrm{i}(\mathrm{t})|)\end{aligned} \tag{2-55}$$

$$\mathrm{ACI}_j = \int_0^T \mathrm{a}_j(t)dt \tag{2-56}$$

$$\mathrm{AAI}_j = \frac{1}{T}\cdot\int_0^T \mathrm{a}_j(t)dt \tag{2-57}$$

where $\alpha_i(0)$ represents the initial acceleration of each generator, and this index refers to a specific generator on α_i. The $\alpha_j(t)$ represents the acceleration of generator *j* at time instant *t*, and *T* is the integration period calculated from the power system fault inception at $t = 0$.

TRASI is defined as:

$$\eta = \frac{\delta_T - {\delta^{pst}}_{max}}{\delta_T + {\delta^{pre}}_{max}} \tag{2-58}$$

where δ_T is a predefined transient stability threshold and $\delta_T = 360^{\circ}$.

The ${\delta^{pst}}_{max}$ and ${\delta^{pre}}_{max}$ are the post-fault maximum rotor angle difference and pre-fault maximum rotor angle difference, respectively [60].

$$If \begin{cases} \eta > 0 \rightarrow \text{stable} \\ \eta < 0 \rightarrow \text{unstable} \end{cases}$$

The TRASI is a comparative measure of the rotor angle separation between synchronous generators following a transient fault in the power system. During transient grid fault, the maximum rotor angle difference (MRAD) between a group of generators can change from 0° to 360° depending on the network strength and fault severity observed by the synchronous generators. Therefore, the higher positive value of η between $0 < \eta < 1$ is more stable.

The TRASI and TSI are the indices to measure the whole system stability status. TRASI measures the maximum rotor angle difference before and after a system fault to determine the severity. But the research for this subject has commonly used the angle-based stability index of TSI as described in Eq. 2-47 [61]. Also, its concept is not complicated to be implemented by Power System Toolbox (PST) software. Therefore, this thesis uses TSI in the transfer learning-based dynamic security assessment in Section 5. According to the criterion of Eq. 2-47, secure and insecure labels are used to classify transient stability status.

2.2.3 Small-Signal Stability

Small-signal stability is concerned with the ability of a power system to maintain stable conditions under a small disturbance. The concept comes from a scenario of the power system stability that some parameters change the power system in small value, e.g. the power output of the generator or the load. As such, the differential and algebraic equations of the power system can be linearized around the equilibrium point. In [34], it points out that the lack of positive synchronizing torque steady increases in generator rotor angle, and lack of sufficient positive damping torque causes rotor oscillations of increasing amplitude. In [50], the post-fault time frame of a power system is expected to attenuate in under 20 s, and the small-signal instability is usually a problem of insufficient system oscillation damping. The above insufficiency causes aperiodic instability and oscillatory instability, respectively.

When the deviations of the rotor angles of synchronous machines from the equilibrium point are so small during small disturbances, the eigenvalues of the system state matrix and corresponding model analysis can analyse different oscillatory modes of the system to characterise the small-signal stability. According to [16, 50], the scenario of aperiodic instability cannot be entirely eliminated when the high-gain automatic voltage regulator (AVR) maintains constant field voltage due to the action of excitation limiters (field current limiters). The electromechanical modes of rotor angle oscillations can be grouped into inter-area modes over the frequency range of around 0.1 to 1.2 Hz and local-area modes from approximately 0.7 to 2.5 Hz [34, 62].

In [63], a general method was proposed to compute the different conditions of small-signal stability, including the Hopf and saddle-node bifurcations, maximum and minimum damping points of oscillations, and load flow feasibility boundaries.

The load flow equations and specific requirements form a set of constraints to the eigenvector and the corresponding eigenvalues of the dynamic state matrix. These constraints are used to locate the above characteristic points by optimising an eigenvalue objective function along the rays specified in the space of system parameters summarised as follows [64].

The non-linear power system model [65-72], including algebraic network equations, stator, and differential equations of the generator excitation system, can be expressed in a set of DAEs in Eq. 2-59 as:

$$\begin{aligned} \dot{x}_1 &= f(x_1, x_2, \mathrm{y}, \tau) \\ 0 &= g(x_1, x_2, \mathrm{y}, \tau) \end{aligned} \tag{2-59}$$

where vectors x_1 and x_2 are state (differential) and algebraic variables, respectively, vector y is the specified system parameter, and τ is a parameter chosen for bifurcation analysis. i.e. $\dot{x}_1 = 0$ is the load flow equation as shown in Eq. 2-61.

The equations Eq.2-59 are linearized at a system equilibrium point in load flow solution to obtain the state equation matrix and reduced system Jacobian $\tilde{J}$ in Eq.2-60 as follows:

$$\tilde{J} = J_{11} - J_{12} J_{22}^{-1} J_{21} \tag{2-60}$$

where $J_{11} = \partial F/\partial x_1, J_{12} = \partial F/\partial x_2, J_{21} = \partial G/\partial x_1, J_{22} = \partial G/\partial x_2$.

According to [63, 73], the singularity of the algebraic submatrix J_{22} induces bifurcations and instability. Saddle node bifurcations occur when J is singular, i.e. $det\frac{\partial f}{\partial x_1} = 0$, e.g. an aperiodic type of angle instability or voltage collapse. The power system Jacobian has an imaginary conjugate eigenvalues pair at the point of Hopf bifurcation with no non-negative real parts from other eigenvalues to dominate the oscillatory modes of the power system. If a

periodic type of solution moves away from an equilibrium point of the stable status, Hopf bifurcation occurs; the direction of the eigenvalue transversality condition determines whether it is a stable or unstable oscillation.

Suppose the Jacobian matrix J_{lf} on load flow calculation is singular. In that case, it represents a surface of the feasibility boundary, i.e. $det\, J(x_1, \mathrm{y}, \tau)_{lf}\big|_{x_2 \in D_{lf}} = 0$, where load flow solutions exist in the parameters change of the power system inside the space region of the surface. The load flow feasibility boundary of J_{lf} coincides with the load flow saddle-node bifurcation [68, 74].

The Hopf and saddle-node bifurcations can be located along a given direction in the space of $y_0 + \tau\Delta y$:

$$f(x, y_0 + \tau\Delta y) = 0 \tag{2-61}$$

$$\tilde{J}^t(x, y_0 + \tau\Delta y)l' + \omega l'' = 0 \tag{2-62}$$

$$\tilde{J}^t(x, y_0 + \tau\Delta y)l'' + \omega l' = 0 \tag{2-63}$$

$$l'_i - 1 = 0 \tag{2-64}$$

$$l''_i = 0 \tag{2-65}$$

where $\tilde{J}$ represents the state matrix acquired by using the linearised model in Eq. (2-59) provided that the algebraic submatrix J_{22} is non-singular, l represents the system left eigenvector with the corresponding eigenvalue $\lambda = \alpha + j\omega$, and $\alpha = 0$, i-th element of l with l'(real) and l'' (imaginary) is represented by $l'_i + l''_i$, y_0 is the current operation point, and $y_0 + \tau\Delta y$ specifies a straight line in the space of y, e.g. loading direction.

Since the DAEs are linearised, the damping of critical electro-mechanical mode α represents the system stability status under small disturbance. A negative α implies a stable system, and a positive α refers to an unstable system. The stability margin is smaller if the

negative α is closer to zero. As the corresponding eigenvalue λ has zero real part, the solutions of the power system from Eq. 2-61 to 2-65 correspond to either Hopf ($\omega \neq 0$) or saddle-node ($\omega = 0$) bifurcations. These equations become inconsistent if the above condition does not meet the ray $y_0 + \tau\Delta y$ on the feasibility boundary in load flow calculation.

The point on the feasibility boundary of the load flow calculation along the ray $y_0 + \tau\Delta y$ is given by [75-77]:

$$f(x, y_0 + \tau\Delta y) = 0 \tag{2-66}$$

$${J_{lf}}^t(x, y_0 + \tau\Delta y)l = 0 \tag{2-67}$$

$$l_i - 1 = 0 \tag{2-68}$$

where l corresponding to J_{lf} to have a zero eigenvalue.

The following constraints optimization approach in Eq. 2-69 can locate the saddle-node, Hopf bifurcation, and load flow feasibility boundary by adding the real part of an eigenvalue α with l' in Eq. 2-62 and l'' in Eq. 2-63, respectively. So, α can be the non-zero solution.

$$\alpha^2 \Rightarrow max/min \tag{2-69}$$

Subject to the constraints

$$f(x, y_0 + \tau\Delta y) = 0 \tag{2-70}$$

$$\tilde{J}^t(x, y_0 + \tau\Delta y)l' - \alpha l' + \omega l'' = 0 \tag{2-71}$$

$$\tilde{J}^t(x, y_0 + \tau\Delta y)l'' - \alpha l'' + \omega l' = 0 \tag{2-72}$$

$${l'}_i - 1 = 0 \tag{2-73}$$

$${l''}_i = 0 \tag{2-74}$$

In [64], the Lagrange function is employed in Eq. 2-75 to consider the load flow constraint in Eq. 2-70 for the optimization approach in Eq. 2-69. Refer Eq. 2-69 and 2-70.

$$\Phi = \alpha^2 + f^t(x, y_0 + \tau\Delta y)\eta \Rightarrow min_{x,\tau,\eta}$$
$$= \alpha^2 + \Phi' \Rightarrow min_{x,\tau,\eta} \quad (2\text{-}75)$$

The maximum and minimum damping points imply $\frac{d\alpha}{d\tau} = 0$. These points specify all oscillatory modes of interest in damping variations caused by the related parameters change in the power system. The Hopf or saddle-node bifurcations correspond to $a = 0$. They imply the specified loading trajectory $y_0 + \tau\Delta y$ to reach the small-disturbance stability limit. The left eigenvector $l = l' + jl''$ determines the key factors of the critical oscillatory mode, such as the excitability, participation factors, sensitivity of a w.r.t. y, shape, mode, and observability [34, 78, 79]. The constraint set Eq. 2-70 to 2-74 gives the critical oscillatory mode frequency. The points on the feasibility boundary of load flow calculation determine the maximum power transfer capabilities. The system is stable along the ray $y_0 + \tau\Delta y$ up to this feasibility boundary where the constraint Eq. 2-70 cannot be satisfied if it goes further to end the optimization procedure.

2.2.4 Voltage Stability

The concept of voltage stability in the power system is better explained by the definition of a saddle-node bifurcation point, referred to as the maximum loading point, critical loading conditions, voltage stability limit, etc. Referring to the Section on Small-Signal Stability, a general method can be used in the nonlinear equations from Eq. 2-69 to 2-74 to locate a critical loading point at a saddle-node bifurcation. When the power system operation reaches this voltage stability limit, the load flow Jacobian matrix J_{lf} is singular, i.e. $det\, J(x_1, \mathrm{y}, \tau)_{lf}\big|_{x_2 \in D_{lf}} = 0$, the two distinct solutions on the power-voltage or reactive-voltage curves emerge into one. The load flow feasibility boundary is also the critical loading condition, located at the nose point on the power-voltage or reactive-voltage curves. In practice, if the power system is under high load conditions, and the load demand increases beyond this feasibility boundary with divergence on the power flow solution, it may cause a blackout if the system control does not work. The power-voltage margin can also be calculated from an operation point along a specific ray if a locally closest saddle-node bifurcation point is located.

Voltage stability is concerned with the ability of the power system to maintain or recover voltage magnitudes to acceptable levels following a contingency event [45]. Referring to Eq. 2-59 in the Section of Small-Signal Stability, the parameter space y is composed of the system and operating parameters as follows:

$$\dot{x}_1 = f(x_1, x_2, \mathrm{y}, \tau)$$
$$0 = g(x_1, x_2, \mathrm{y}, \tau)$$

where the system parameter describes the system topography and operating parameters include load, generations and voltage set points.

The voltage collapse phenomena usually occur when the generators with voltage regulators reach their Var limit with fixed excitation that the shaft dynamics or the interaction of the regulator with the field flux cause the system instability. A single two-axis dynamic model machine and single load case are used to demonstrate such phenomena as [65, 80]:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - (X_d - X'_d)I_d + E_{fd} \tag{2-76}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d + \left(X_q - X'_q\right)I_q \tag{2-77}$$

$$\frac{d\delta}{dt} = \omega - \omega_o \tag{2-78}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - E'_d I_d - E'_q I_q - \left(X'_q - X'_d\right)I_d I_q - T_{FW} \tag{2-79}$$

where Eq. 2-76 to 2-79 model the synchronous machine, assume $X'_d = X'_q$, E'_d, E'_q are transient EMF of the d-axis and q-axis, respectively, I_d, I_q are d-axis and q-axis stator currents, X_d, X_q are d-axis and q-axis synchronous reactance, X'_d, X'_q are d-axis and q-axis transient reactance, δ is generator power angle, ω is generator rotor speed, ω_o is the rated synchronous speed, H is inertia, TM is mechanical torque, T_{FW} is friction windage torque, T'_{do}, T'_{qo} are open-circuit time constant of the d-axis and q-axis, respectively, E_{fd} is the field voltage input (from exciter), and

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right)E_{fd} + V_R \tag{2-80}$$

$$T_F\frac{dR_f}{dt} = -R_f + \frac{K_F}{T_F}E_{fd} \tag{2-81}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A R_f - \frac{K_A K_F}{T_F}E_{fd} + K_A(V_{ref} - V_t) \tag{2-82}$$

where Eq. 2-80 to 2-82 model the excitation and voltage regulator, T_A is the amplifier time constant, T_E is the exciter time constant, T_F is the stabilizer time constant, K_A is the amplifier gain, K_E self-excited dc exciter, K_F = M/R, R is resistance, M is mutual inductance of the

transformer, V_R is the exciter input, R_f is the rate of feedback, S_E is the saturation function, V_t is compensated or uncompensated terminal voltages, and

$$T_{RH}\frac{dT_M}{dt} = -T_M + \left(1 - \frac{K_{HP}T_{RH}}{T_{CH}}\right)P_{CH} + \frac{K_{HP}T_{RH}}{T_{CH}}P_{SV} \tag{2-83}$$

$$T_{CH}\frac{dP_{CH}}{dt} = -P_{CH} + P_{SV} \tag{2-84}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_o} - 1\right) \tag{2-85}$$

where Eq. 2-83 to 2-85 model the system turbine plus governor, T_M is the synchronous machine torque, P_C is a control input, P_{SV} is the effect of steam valve position, P_{CH} is the change in output power of the steam chest, T_{CH} is the time delay constant due to the steam chest and inlet piping, T_{RH} *is the* time delay constant due to reheaters, T_{SV} is the time constant due to steam valve position, K_{HP} is the fraction of output power of the steam chest to convert into torque, R_D = % droop/100.

$$\bar{E} = \left(E'_d + jE'_q\right)e^{j(\delta - \pi/2)} \tag{2-86}$$

$$\bar{I} = \left(I_d + jI_q\right)e^{j(\delta - \pi/2)} \tag{2-87}$$

Eliminate I_d, I_q, and V_1 from Eq. 2-76 to 2-85 are:

$$\mathrm{V_2 e^{j\theta_2}}\left(\mathrm{I_d} - \mathrm{jI_q}\right)\mathrm{e^{-j(\delta-\pi/2)}} = \mathrm{P_L(V_2)} + j\mathrm{Q_L(V_2)} \tag{2-88}$$

$$\mathrm{V_2 e^{j\theta_2}}\frac{\left(\left(E'_d - jE'_q\right)e^{-j\left(\delta-\frac{\pi}{2}\right)} - \mathrm{V_2 e^{-j\theta_2}}\right)}{-j\left(X'_d + X\right)} = \mathrm{P_L(V_2)} + j\mathrm{Q_L(V_2)} \tag{2-89}$$

$$\left(\mathrm{I_d} + \mathrm{jI_q}\right)\mathrm{e^{j(\delta-\pi/2)}} = \frac{\left(E'_d + jE'_q\right)e^{j\left(\delta-\frac{\pi}{2}\right)} - \mathrm{V_2 e^{j\theta_2}}}{j\left(X'_d + X\right)} \tag{2-90}$$

$$\mathrm{V_1 e^{j\theta_1}} = jX\left(\mathrm{I_d} + \mathrm{jI_q}\right)\mathrm{e^{j(\delta-\pi/2)}} + \mathrm{V_2 e^{j\theta_2}} \tag{2-91}$$

For given E'_d, E'_q and δ, it relies on a load model to represent the $\mathrm{P_L(V_2)}$ and $\mathrm{Q_L(V_2)}$ for solving $\mathrm{V_2}$ and θ_2 in Eq. 2-89. Then, $\mathrm{I_d}$, $\mathrm{I_q}$, $\mathrm{V_1}$ and θ_1 can be found in Eq. 2-90 and 2-91.

Different load model causes the power system to lose steady-state stability at different section along the PV curve.

The above equations indicate that insufficient reactive power resources and transmission facilities in locations with high electricity consumption load centres can relate to the issues of voltage instability. The power system deregulation also weakens the ability of the power system to maintain a proper balance of reactive power throughout the system. These insufficiencies may consequently lead to the power system cascading outages due to loss of transmission elements, generation units, etc. subject to the response of the protection system, and the characteristic of load model is a challenge to voltage stability.

The voltage instability can be classified as short-term instability and long-term instability. If an event, such as a contingency or rapid load increase on the power system with a large number of induction machines, bus voltages may decline sufficiently to stall many induction motors to cause short-term voltage collapse. Therefore, the load model is critical to capturing the essential motor characteristics. Since the load dynamics and system controls determine the system stability, dynamic analysis (such as time-domain simulation) is widely used for short-term voltage collapse. However, voltage instability may occur in a longer time frame to cause long-term voltage collapse if the contingency cannot stall the induction motors within a short period of time. A progressive fall of voltage on some buses in a location where there exists a reactive power imbalance between load and generating resources may occur. The system controls, i.e. generators AVRs, static VAr systems, automatic switched capacitors and automatic transformer under-load tap-changers, will continue to rectify bus voltages close to their pre-contingency levels until the additional reactive support is restricted.

2.2.5 State-Of-The-Art Technologies

The integration of converter-interfaced generation technologies (CIGs), transmission devices, and loads into bulk power systems has led to notable changes in the dynamic behaviour of the power system. These fast-response power electronic devices, while lacking short-circuit current capabilities during faults or inherent inertial response, have a significant impact on various dynamic phenomena, encompassing voltage stability, electromagnetic transients, and both large- and small-disturbance stability. The dynamic performance of CIGs predominantly relies on the effective functioning of inner current control loop, the high-level control loops, and the dynamic characteristics of the phase-locked loop (PLL). However, in connection points with a low short-circuit ratio (SCR), the response of the PLL and the inner current control loop may exhibit oscillatory behaviour, which significantly influences the system's dynamic recovery following a disturbance. Given this perspective, it is imperative to acknowledge the significance of faster dynamics that operate within electromagnetic time scales. For example, the controls of CIGs contain a range of wave and electromagnetic phenomena. In accordance with [81, Fig. 2], power system stability analysis has evolved to incorporate two novel categories: resonance stability and converter-driven stability. These novel categories and Sections 2.2.1-2.2.4 are considered as the current state-of-the-art technologies used in power system analysis. These novel categories, as summarized from [81], are described as follows:

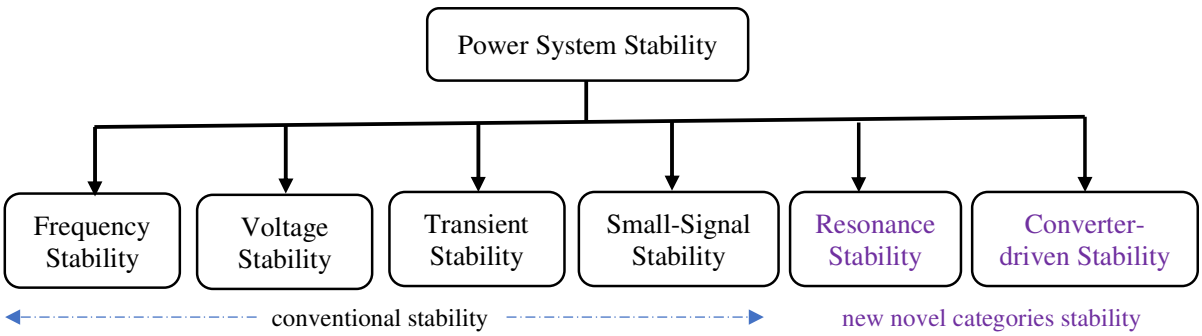

Fig. 2- 3 State of the art technologies for classification of power system stability [81, Fig. 2]

Resonance stability includes subsynchronous resonance (SSR), regardless of whether it is associated with electromechanical resonance or purely electrical resonance.

Torsional Resonance: Subsynchronous oscillations are commonly categorized as device-dependent subsynchronous oscillations (DDSSO) and SSR [82]. DDSSO occurs due to the interaction between fast-acting control devices and the torsional mechanical modes of nearby turbine-generators [82-87]. However, it is worth noting that DDSSO can have beneficial effects on torsional damping in certain scenarios [88]. SSR arises from the resonance between the series-compensated electrical network and the mechanical modes associated with torsional oscillations on the turbine-generator shaft. It involves poorly damped, undamped, or even negatively damped oscillations that increase in magnitude when the turbine-generator operates at the natural subsynchronous torsional modes of oscillation on the combined turbine-generator mechanical shaft [82, 83].

Electrical Resonance: SSR can also be associated with series compensation and the induction generator effect (IGE). However, it is exclusively observed in conventional turbine-generators within power systems. As indicated in reference [89], variable speed doubly fed induction generators (DFIGs) that have a direct grid connection are highly vulnerable to self-excited subsynchronous resonance (SSR) due to the control mechanisms of the converter linked between the rotor and stator circuits. SSR arises when a resonant circuit is formed at subsynchronous frequencies through the combination of a series capacitor and the effective inductance of the induction generator, which is affected by the inherent negative resistance of the rotor. This combination results in a negative net apparent resistance in the circuit.

The resonance phenomenon, concerning the occurrence of substantial current and voltage oscillations during subsynchronous oscillations, can be attributed to the control action of the DFIG converter. This control action plays a dominant role in generating negative damping at

electrical resonant frequencies. This phenomenon is referred to as subsynchronous control interaction (SSCI) [90-92], and the insulated-gate bipolar transistor (IGBT) has been identified as the major contributor to stability issues related to SSCI [93-100].

Voltage-source converters (VSCs) are extensively utilized in the large-scale CIGs. A typical CIG, as the VSC has a predominant role to interface with the grid, relies on control loops and algorithms with fast response times, such as the inner-current control loops and PLL [101]. On the other hand, converter-driven instability can give rise to unstable oscillations in the power system, resulting in a broad spectrum of fast and slow frequency interactions. These interactions have implications for both the electromagnetic transients and electromechanical dynamics observed in the power network. Fast-interaction converter-driven stability encompasses frequencies ranging from tens to kilohertz, whereas slow-interaction converter-driven stability primarily encompasses frequencies below 10 Hz.

Harmonic instability, also known as fast-interaction converter-driven instability, arises from the rapid dynamic interactions among the control systems of CIG and fast-response passive system components, as well as other power electronic-based devices. These interactions lead to resonance and multi-resonance effects. For instance, the operation of multiple inverters in close proximity can trigger series and parallel resonances, resulting in multiple resonance peaks [102-104]. Furthermore, it introduces mutual interaction challenges among the control loops of grid-connected converters [105, 106]. Additionally, the control of an inverter has the potential to initiate resonance within the inverter filter itself [107].

Slow-interaction converter-driven instabilities occur as a result of the slow dynamic interactions between control systems in power electronic-based devices and slow-response components within the power system. The instability arises due to the maximum power transfer between the converter and the rest of the system, especially in the case of a weak

system. As a result, converter-driven instability is associated to the control loops of power electronic converters, including the outer control loops responsible for power and voltage regulation, as well as the PLL of CIGs, which can induce unstable low-frequency oscillations [108].

According to [109, 110], oscillations in direct-drive permanent-magnet generator (PMG) wind turbines and the local grid in power systems with low SCR, particularly weak grids with an SCR below 2, have the potential to lead to growing low-frequency oscillations and instability. The crucial factors contributing to this instability include the CIG online capacity, as well as the control strategies and parameters of the converters, such as inner-current control loops or PLL gains and bandwidth imposed by the low-pass filters in weak grids [111, 112].

It is crucial to improve the ability of the PLL to synchronize with the grid, especially when encountering nearby faults, to effectively mitigate the risk of negative admittance arising in parallel with the system input admittance. This situation can lead to a high-gain PLL [113], which, in turn, can cause the converter to fail in adequately adjusting its phase for exporting generated power. Additionally, this scenario becomes problematic when the inverter reaches its current limit. To address these issues, it is essential to fine-tune the parameters of the outer loop and adjust the response time of the PLL. These adjustments can effectively improve the connection between VSC-HVDC systems and weak grids [114, 115].

As mentioned above, the power system is facing challenges, including the reduction of system inertia and system strength due to the increasing IBRs and decommissioning of conventional fossil fuel fired synchronous generators. The transition towards more IBRs leads to notable changes in the dynamic behaviour of the power system. Consequently, there is a growing application of artificial neural networks in load modelling and power system

stability. My research proposes new approaches to tackle the challenges posed by limited data when training artificial neural networks for non-intrusive load monitoring and dynamic security assessment. Additionally, it demonstrates the feasibility of developing a new hybrid load model with physical meanings or reliable mathematical formulas to represent the load characteristics of the power grid in an environment of high integration of renewable energy sources.

2.3 Load model affects system stability

In [116], the effect of load modeling on the dynamic behavior of the frequency, transient, and small disturbance stability studies regarding the model parameters of the exponential load model are investigated. Different load mixtures are obtained by varying the voltage exponent from 0 to 2, assuming a step of 0.2 on eleven load model representations.

Refer to Eq. 2-1 and 2-2 of the exponential load model in Section 2.1.1 as follows:

$$P_s = P_0 \left(\frac{V}{V_n}\right)^{k_{pu}} \tag{2-1}$$

$$Q_s = Q_0 \left(\frac{V}{V_n}\right)^{k_{qu}} \tag{2-2}$$

where the voltage exponents k_{pu} and k_{qu} equal to 2, 1, and 0 denote a CI, CC, and CP load model, respectively. Assume the effect of frequency fluctuation to be ignored.

In the work of impact evaluation on transient stability, the three-phase and self-clearing faults with a duration of 0.2 s is considered. In all cases, the fault resistance assumes to be 1 Ω in IEEE 39 bus modified network [117]. There are 20 faults on a different bus of the system in a random distribution to be simulated on each load representation. The result indicates that the number of stable cases increases if the exponent increases. 70% of the examined faults result in unstable cases on the constant power model, but the unstable cases drop to 15% on the constant impedance model.

In the work of impact evaluation on frequency stability, the active power of load demand at bus 24 on the above power network is increased by 59% as a 3% change of the total system load. The result indicates that the ROCOF index decreases if the exponent increases. It implies that the constant power load model has the worst effect on the frequency stability of the power system.

In the work of impact evaluation on small-signal stability, the system eigenvalues for the cases of the constant impedance, constant current, and constant power load models on the stability perspective are compared as referred to Section 2.2.3. The result indicates that the constant power load model leads to being cautious.

The above three evaluation report that the constant impedance load model leads to very few unstable cases and relatively low-frequency deviations than the others. On the other hand, the constant power load model leads to the most pessimistic results.

In [118], the effect of static load model as referred to Eq. 2-9 and 2-10 in Section 2.1.1 for oscillation damping and frequency on small-signal stability was investigated using the eigenvalues $\lambda = \alpha + j\omega$ of the electromechanical oscillation modes computation on a single-machine infinite bus system [119]. A three-phase short-circuit is applied at time = 1 s and a duration of 100 m s at node N-Line. The load assumes constant impedance behaviour until fault clearing at time = 1.1 s before the load model on an investigation is applied.

$$P_s(V,f) = \left[P_{01}\left(\frac{V}{V_n}\right)^2 + P_{02}\left(\frac{V}{V_n}\right) + P_{03}\right](1 + k_{pf}(f - f_0)) \quad (2\text{-}9)$$

$$Q_s(V,f) = \left[Q_{01}\left(\frac{V}{V_n}\right)^2 + Q_{02}\left(\frac{V}{V_n}\right) + Q_{03}\right](1 + k_{qf}(f - f_0)) \quad (2\text{-}10)$$

The result of time-domain simulation reveals that the damping mode of the generator is the lowest for constant power load and the highest for constant impedance load. In contrast, the oscillation frequency is the lowest for constant impedance and the highest for constant power load. Damping and oscillation frequency for constant current load lie in between the values for constant impedance and constant power load, respectively. In addition, the differences in both the damping and oscillation frequency for varying voltage dependency diminish with increasing load distance from the generator, e.g. the values of the constant impedance load as a reference, damping of constant power and constant current load are about 91% and 96%.

The eigenfrequencies are nearly 113% and 106%, respectively, at a distance of 140 km. In contrast, damping of constant power and constant current load are only 96% and 98%. The eigenfrequencies are about 106% and 103%, respectively, at a distance of 280 km.

The effect on varying the frequency sensitivity factors k_{pf} and k_{qf} reveals that $k_{pf}, k_{qf} = -1$ significantly decreases the damping and increases the eigenfrequency if the voltage dependency is ignored by setting $P_{01}, P_{02}, Q_{01}, Q_{02} = 0$. In contrast, $k_{pf}, k_{qf} = 1$ leads to the opposite effect. The differences in both the damping and oscillation frequency for varying frequency dependency also diminish with increasing load distance from the generator, e.g. $k_{pf}, k_{qf} = 0$ as a reference, the damping on $k_{pf}, k_{qf} = -1$ and $k_{pf}, k_{qf} = 1$ are about 49% and 151% respectively. In contrast, the eigenfrequencies are nearly 103.0% and 98.8% at a distance of 0 km, and damping are about 80% and 121%, whereas the eigenfrequencies account for 100.9% and 99.6% at a distance of 280 km.

The effect on typical load classes is also studied by combining the Eq. 2-1, 2-2, 2-9 and 2-10 in Section 2.1.1 to represent the load model on residential, commercial, and industrial as follows:

$$P_s(V, f) = \mathrm{P_0} \left(\frac{\mathrm{V}}{\mathrm{V_n}}\right)^{\mathrm{k_{pu}}} (1 + k_{pf}(f - f_0)) \tag{2-92}$$

$$Q_s(V, f) = \mathrm{Q_0} \left(\frac{\mathrm{V}}{\mathrm{V_n}}\right)^{\mathrm{k_{qu}}} (1 + k_{qf}(f - f_0)) \tag{2-93}$$

where $\mathrm{k_{pu}}$ = 1.4, $\mathrm{k_{qu}}$ = 2.9, k_{pf} = 0.8, k_{qf} = -1.9 represent residential load, $\mathrm{k_{pu}}$ = 0.8, $\mathrm{k_{qu}}$ = 2.5, k_{pf} = 1.5, k_{qf}= -1.4 represent commercial load, $\mathrm{k_{pu}}$ = 0.1, $\mathrm{k_{qu}}$ = 0.6, k_{pf} = 2.6, k_{qf} = 1.6 represent industrial load.

The result reveals that the damping and eigenfrequency for residential load show the lowest value, whereas they are both highest for the industrial load. Furthermore, the frequency

dependency and voltage dependency demonstrate higher effects on damping and eigenfrequency, respectively.

In [120], the dynamic load model of second-order difference equations is obtained by the data measured from the Guangdong grid, China. This dynamic load model and a BPA static load model are applied to power system simulation, i.e. two outputs are calculated by the models to compare the data obtained from field measurement in a power substation. The dynamic model is referred to the Eq. 2-32 and 2-33 in Section 2.1.2 as follows:

$$\Delta P(k) = a_{p1}\Delta P(k-1) + a_{p2}\Delta P(k-2) + c_{p0}\Delta V(k) + c_{p1}\Delta V(k-1) + c_{p2}\Delta V(k-2) + g_{p0}\Delta f(k) + g_{p1}\Delta f(k-1) + g_{p2}\Delta f(k-2) \quad (2\text{-}32)$$

$$\Delta Q(k) = a_{q1}\Delta Q(k-1) + a_{q2}\Delta Q(k-2) + c_{q0}\Delta V(k) + c_{q1}\Delta V(k-1) + c_{q2}\Delta V(k-2) + g_{q0}\Delta f(k) + g_{q1}\Delta f(k-1) + g_{q2}\Delta f(k-2) \quad (2\text{-}33)$$

The BPA static load model consists of different static load combinations in different regional areas, e.g. Hong Kong: 60% constant power, 40% constant impedance; Guangzhou: 30% constant power, 40% constant current and 30% constant impedance.

Two transient stability studies are performed with three-phase-to-ground fault and a duration of 0.5 s when load flow is within the limit. One is based on the above static load model, and the other is based on the difference equations. The short circuit simulation result indicates that the dynamic load model better fits the real currents response from the grid. The static load model does not represent the load correctly in this case. Also, the dynamic load model obtains slightly higher limit power than static load model in transient stability limit. It implies that the dynamic load model better describes the dynamic characteristic of system load than the static load model in short circuit scenarios. The dynamic load model records 6% higher stability limit with longer critical clearing time (CCT) of 0.01-0.08 s. In [121-123], the performance of dynamic and static models, e.g. induction motor dynamic and static

exponential frequency-dependent using real measurement in transient stability studies, indicated that static models reflect the system response accurately in grid-connected mode.

In [121], the software (LOADSYN) program, which was developed by General Electric (GE), was used with the mixture of data from billing information to build the static and/or dynamic load models from the component-based approach for transient stability studies on a system-wide basis of the Ontario Hydro 4000 bus base case. The static load models derived using the energy billing data suffices, and the composite load model consisting of an induction machine in parallel with a static load model represents the characteristics of the system dynamics. The comparison results of these two models indicate that the composite load model does not have significant accuracy improvement on security limit studies. However, the composite load model should be used in an islanding scenario containing heavy industrial loads with large frequency and voltage fluctuations.

Reference [123] also supports that voltage-dependent static load models work well only where the power system does not have a high penetration of electronic loads, induction motors, etc. The static load models do not effectively represent the system load dynamics, e.g. the composite load model and ZIP model are assessed in short-term voltage stability analysis. The induction motor, e.g. air-conditioner compressor, stalls to draw currents over the range of 4 to 6 times their nominal value that resist the voltage recovery in post-disturbance conditions.

In [122], the EMTP software on a component-based approach was adopted to model a mixed residential-commercial load on the Hydro-Quebec base load. The resulting composite load model is simulated on a severe single-line-to-ground fault to determine the static, dynamic, and post-fault recovery characteristics.

Refer to Eq. 2-1 and 2-2 of the exponential load model in Section 2.1.1. Assume the effect of frequency fluctuation to be ignored. The following points are concluded according to the EMTP simulation results on the composite load model by using the load coefficients (k_{pu}, k_{qu}) as functions of the voltage.

- for power flow and steady-state contingency analysis studies, a static load model with constant voltage exponent k_{pu}, k_{qu} can be conveniently used, with k_{pu} = 1.1 and k_{qu} = 1.4.
- for slow-developing phenomena such as voltage stability, a static load model with variable voltage exponent k_{pu}, k_{qu} is more suitable.
- for faster dynamic modes such as first swing and transient stability, a dynamic model including one or two equivalent motors is better developed to avoid the reactive power dynamics and large phase lead from the static load model.

2.4 The modern concept of system stability with AI

Section 2.3 has described that the choice of load models can have a remarkable impact on simulation results of system stability. As increasing load demand on power systems to operate near security limit for economic considerations and more renewable energy sources are integrated into power systems to cause large conventional power plants to be shut down, the amount of the rotating masses reduces, i.e. the system inertia to increase the influence effect of electrical load behaviour on power system stability. Due to the non-linear nature of the power system, artificial neural networks have been adopted to benefit the research on load modelling and pattern recognition for online power system stability. ANNs techniques have also been successfully applied to power system problems regarding classification or estimation for fast and reliable results. An approximation problem can be solved through the ANN learning process to find a solution surface proportioned by the best adaptation of the training parameters, with the criterion of the best adaptation measured by some statistical methods. These techniques also well support the interpolate and extrapolate approaches when the measurement-based methods lack data for dynamic load modelling.

In [124], two load models, i.e. a multi-layer feedforward neural network with backpropagation learning and difference equations with recursive extended least square identification, were compared to extrapolate and interpolate the data recorded of wide voltage variation during naturally occurring disturbances. In addition, the artificial short-circuits were carried out to investigate if the multi-layer feedforward neural network can better describe the power-voltage nonlinear relationship of load dynamic characteristics than the difference equations, which represent the interpolation and extrapolation nearly linear.

Six artificial short-circuits were tested at the Shenyang High Voltage Switch Factory nearby the HUSHTI 220 kV substation in the Northeast Power System of China. The first

three tests were three-phase short-circuits with about 8% voltage variation, and the others were two-phase-to-ground short-circuits with 2.5% voltage variation at the substation. The frequency, active and reactive powers of the distribution transformer and voltage on the 220 kV bus were measured at 6 ms intervals between two points on each test with a total of 512 to 800 points. It took about 3 to 5 s to complete the total measurement. In the load model simulation, the inputs are bus voltage and frequency, and outputs are active power and reactive power. The model represents powers as a function of frequency and voltage.

The interpolation quality of the difference equations model is verified using the data from one of the three-phase short-circuit tests to identify the difference equations model. Then, the data groups from the other three tests of two-phase-to-ground short-circuit are applied for evaluation. The calculation results in all three cases indicate that the difference equations model has much lower active power variation than the measured values.

Similarly, the extrapolation quality of the same model is verified using data from one of the two-phase-to-ground short-circuit tests to identify the difference equations model. Then, the extrapolation is carried out with the three data groups of the three-phase short-circuit tests. The results in all three cases indicate that the difference equations model has much higher active power variation than the measured values. Finally, the above approaches are applied to the multi-layer feedforward neural network model. The results indicate that the coincidence of curves calculated by the neural network model outperforms the difference equations for interpolation and extrapolation. Therefore, the proposed multi-layer neural network model is a more promising dynamic load model than the difference equations model on interpolation and extrapolation capacity.

In [125], a neural network-based composite load model was proposed. First, the simulation data is collected from a testbed of the power system for transient stability, which consists of

the nominal parameters of the loads on a system bus to the level of the distribution network. The motor loads are with associated converters for large industrial systems. This dataset of the composite load is used to train a two-layer recurrent neural network using the terminal voltage, system frequency, and load power. Since the training is based on nominal parameters of the loads, it cannot wholly represent the real load characteristics of the power system. However, it helps to set up an acceptable first-pass estimation of the complex loads in a neural network model. Then, the neural network is re-trained to update using measured data, i.e. systems disturbances such as line switching and fault, to improve the load model accuracy.

The above recurrent neural network, which estimates the load power from system frequency and terminal voltage, is developed to map the nonlinear relationship between the load power w.r.t. frequency and voltage. It reflects the changing dynamics of the loads when the system is under disturbances. A simplified testbed, which consists of an induction motor connecting through a power inverter (IGBT) and a rectifier to the power system, is set up to test the above approach. When the supply voltage and system frequency of this testbed are adjusted in the range of 0 to 1.2 p.u. and 55 to 65 Hz, respectively, the reactive and active power at the load terminals are recorded. This recorded dataset is divided into training and test data for the neural network. The results indicate the proposed recurrent neural network model successfully estimates the load dynamics with the mean of estimation error for reactive power being 4.04x10-5 p.u. and that for active power being 3.07x10-4 p.u.

In [126], an artificial neural network with feedback loops and time delay was proposed to model the dynamic load in a power system. It assumes only the measured phase voltages are available at the recall stage. Several distribution substations of Kyushu Electric PS were monitored for field measurement. The three-line currents and three-phase voltages of the

substation load were continuously monitored at a sampling frequency of 20 samples/cycle or 1.2 kHz of each quantity. The data on residential load was obtained from the substation at site A. The substation at site B measured data on residential and light industry load.

Three three-layer neural network models that differ in the number of inputs and feedback loops were proposed and tested. The output signal is either the estimated discrete reactive or real power, and the input signals are sampled phase voltages at a load bus. The first model constructs with a one-sample delay feedback block of the output power signal to the input node. The second model adds another two-sample delay feedback block of the output power signal to the second input node. The third model adds three-sample delay feedback block in a similar structure. In addition, two bias nodes are added to the input layer and the hidden layer to reduce training time, respectively.

The above three-layer neural network models were trained with the specified learning rate of 0.25 and momentum of 0.90 using the measured data from one fault at site A. Then, the sampled voltage data from the training dataset were input to the trained models to check if the estimated load power could effectively approximate the load power response during a given power system fault.

Another measured dataset at the same site and during the same season was input into the above three trained models to verify the effectiveness. The results indicate that the proposed load models have highly accurate dynamics of the active and reactive power responses for the given disturbance. In addition, the second model with one-sample and two-sample delay feedback blocks gives the best estimation result than the other two models.

The measured datasets from sites A and B during the different seasons were used to verify the above three trained models further. The results indicate that the performances of the

trained load models are dependent on location and season. The estimated output accuracies are degraded if the data from different seasons or locations are used. However, the accuracy of the load model can be significantly improved after being re-trained.

In [127], a radial basis function network (RBFN) was proposed to estimate the critical clearing time (CCT) of a single fault in a power system. CCT is the transient stability measure of power system robustness to withstand large disturbances. It is defined as the longest clearing time of system fault before the generators lose synchronism. Since the CCT value depends on the operating conditions of power systems, e.g. change of load demand and system topology, varies continuously in real-time, it requires periodically updating the CCT value by time-consuming computing solutions of nonlinear on-fault system equations for those off-line mode in utility control centres.

The RBFN has a structure of an input layer, an output layer, and the hidden layer. The radial basis functions are adopted at each node in the hidden layer to produce a localised output. An activation function maps the combinations of weighted inputs to the outputs. Each radially symmetric activation function requires a centre and a scale parameter.

The simplified Brazilian South power system with 10 generators, 45 buses, 72 lines and/or transformers was used to simulate the contingencies of a three-phase short-circuit fault at a bus and to trip one line or transformer adjacent to the faulted bus. There are 19 contingencies applied at 525 and 230 kV buses to generate 800 scenarios for each offline RBFN training. Using the ANAREDE computer package and SLEPM, the relationship between the pre-fault active power generation levels of the generators and the corresponding CCT value is obtained. As such, the active power generation levels of the generators, the value of the CCT for a given stability scenario and the number of neurons in the hidden layer (25, 50, 75, and

100 neurons) are input, output and design parameters to RBFN, respectively. Each RBFN has 3,200 training patterns (800 base cases versus 4 neurons in the hidden layer).

The sensitivity regarding the design parameters based on the 100 case tests verifies the global performance of the RBFN. The results indicate that the relative mean errors using a different number of neurons in the hidden layer are lesser than 3.65%, and the approximation capacity is relatively good.

Reference [128] presented various artificial neural networks such as adaptive neuro-fuzzy inference systems (ANFIS), general regression neural network (GRNN), radial basis neural network (RBNN) and multi-layer perceptron (MLP) for estimating CCT. Since the common approaches, including time-domain simulations (TDS) and direct methods based on equal area criteria and transient energy functions, have shortfalls in high computational effort and low applicability for real power systems, respectively, an effective method is required to obtain a high level of CCT accuracy in faster calculation.

This study used a single ANN structure to estimate the minimum CCT value for all contingencies, topologies, and loading conditions. First, the power flow computations and dynamic security assessment via time-domain simulations at each operation point (OP) were carried out by the software DSAToolsTM on a 16-generator 68 bus test system [129, 130]. The simulation includes 8 credible topologies, and the loading level on each topology is increased from 60 to 110 %. Then, the three-phase-to-ground faults were simulated to trip a single line in 5 cycles on critical contingencies.

Decision Trees (DTs) were used as a feature selection tool of 410 features from the dataset of 1344 OPs, which is classified as insecure and secure, based on their CCT values on each OP. As such, the phase angle and magnitude of the bus voltages, active and reactive power

flows at each transmission line, active and reactive power outputs of each generator, active and reactive power demands at all load buses, and CCT value for each critical contingency are selected to reduce the ANN complexity and training time. The 1344 OPs are randomly divided into training data (80%) and testing data (20%). As each OP has the same number of CCT values and critical contingencies, the minimum value of CCTs of each OP is selected when the class of each OP is detected, i.e.

If $CCTi > T$ cycles, sample i is secure; otherwise, insecure.

where T is the threshold value determined for each contingency considering the clearing time values of circuit breakers, and $CCTi$ is the minimum CCT value for i-th OP considering all the critical contingencies.

According to [128], The size of the input layer significantly affects the performance of ANFIS. RBNN and GRNN structures. The cluster size for the data features also affects ANFIS structure and variances of activation functions for RBNN and GRNN structures, respectively. A performance evaluation by calculating the root mean square error (RMSE), mean absolute error (MAE) and correlation (C) between real values calculated by TDS and estimated CCTs are carried out on each critical parameter (10, 15, 20, 25, 30 features) of ANNs. The average results on repeated 5 times evaluations indicate that each ANN structure gives good classification results. Still, the MLP approach has a most reliable effect than the others for the CCT estimation due to high correlation and low error values.

3 A hybrid load model on V2G environment

NOMENCLATURE

P_L	Active power of the composite load model
P_d	Active power of the dynamic part in composite load model
P_s	Active power of the static part in composite load mode
P_z	Constant impedance part of the ZIP model
P_i	Constant current part of the ZIP model
P_p	Constant power part of the ZIP model
P_0	Active load in the steady state before the disturbance
V_n	Nominal rated bus voltage
K_{pm}	Initial active power proportion of the equivalent motor in the composite load model
E_m, δ_m	Voltage magnitude and angle behind transient reactants respectively
X_m	Transient reactance
G_s	Constant of the static load in parallel to third order induction motor
E_d	Exponential coefficient for the dynamic part of the load during load shedding
E_s	Exponential coefficient for the static part of the load during load shedding
T_1	The time when disturbance begins to cause the voltage drops below critical voltage under which the load is cut off for load shedding
T_2	The time when the voltage rises above critical voltage after disturbance
Q_0	Reactive power measured at the bus in steady state before the disturbance
Q_d	Reactive power calculated by static model
Q_L	Reactive power calculated by hybrid load model
β	Exponent related to the static model

RC_1 Constant in capacitor

RC_2 Constant in capacitor

RC_3 Constant in capacitor

K_1 Scaling constant

K_2, K_c Constant

3.1 Introduction

Global leaders have set a Mandatory Renewable Energy Target to increase the green electricity supply from distributed renewable sources in response to climate change. As such, distributed generation based on renewable energy sources provides an alternative solution to reduce carbon emissions. With the inclusion of DG by injecting the real power flow near the loads, the voltage values at almost all the buses significantly improve [131]. For example, suppose there are sufficient electric vehicles (Vehicle to Grid, V2G) to operate as backup energy sources under an effective control mechanism. In that case, it can improve the stability of the power grid. Also, it facilitates a shift in the electric energy demand, shaves the peak, and serves as a distributed energy storage device. On the other hand, distributed generation is intermittent based on these energy sources, e.g. wind farms and solar PV farms. Also, the intermittent load pattern from many electric vehicles (Grid to Vehicle, G2V) weakens the smart grid network. Therefore, these renewable energy sources are running at low-capacity factors unless the technology of energy storage improves the reliability of these sources.

The appearance of such unconventional types of loads (power electronics-based or interfaced through power electronics, e.g. electric vehicles) and increased penetration of renewable types (mostly stochastic) of distributed generation (wind farms and solar power farms) may challenge the operation of power grids safely and securely. It causes uncertainty in the operation of electric networks [3]. However, digital simulation, one of the essential tools for power system analysis, is heavily dependent on the performance of the load model. For example, the large-scale blackout of WSCC 1994 demands a high accuracy load model to simulate the practical running condition of the power system [17, 18]. Therefore, a new power system modelling approach is required to meet the stability criteria, and it is a challenge to formulate an accurate dynamic load model.

On the other hand, the power utility has an industrial culture of using a load model with clear physical interpretations or reliable mathematical formulas. They do not accept a black box load model [132]. According to the Electric Power Research Institute (EPRI) 2020, 82% of the respondents from 31 distribution utilities across states in the U.S., Canada, South Africa, and South Korea frequently use the constant power models [133]. In [134], it also states that the load model should be physically based. Therefore, it has value to investigate the improved dynamic load model with physical interpretations or reliable mathematical formulas. This research proposes a new approach to develop a dynamic load model with physical interpretation and reliable formula to meet the above challenge. It is a measurement-based dynamic hybrid load model. This dynamic load model has some uses for the specific power system configuration, as shown in Figure 3- 2. Real data from field measurements with noise can be tested in realistic applications.

Reference [135] validated that a measurement-based load model has a suitable generalization of load dynamics under artificial short-circuit disturbances in large-scale power grids. It is a top-down technique. It analyses disturbance data from a fault event to determine the load characteristics. First, the fault data is collected at the measured bus. Then, a load model is proposed, and its parameters are estimated via optimization by minimizing the error between the measured data and simulated output. It is a mathematical identification process. Inputs are voltage and frequency. Outputs are active and reactive power. With the increasing knowledge of power systems available in a smart grid environment, there is a better chance to build more accurate dynamic load models for simulating the load characteristic in the transient response through a measurement-based load modelling approach [136].

A new hybrid model which separates the conventional load model into two different composite load models, i.e. a composite load model of active power response and a composite load model of reactive power response, is proposed. The composite load model of active power response consists of the components of the equivalent induction motor (IM), capacitor, and static load model (ZIP). The ZIP consists of constant impedance, constant current and constant power. The composite load model of reactive power response consists of the capacitor and the modified formula of the static exponential load model.

As previously mentioned, the appearance of unconventional types of loads and the increased penetration of renewable types of distributed generation introduces operational uncertainties within electric networks [3]. Nonetheless, the power utility expresses a preference for using a load model with clear physical interpretations or reliable mathematical formulas. Therefore, it has value to investigate the improved dynamic load model that incorporates these attributes in the aforementioned operational environments. In this regard, a novel hybrid model is proposed, which divides the conventional load model into two different composite load models: one for active power response and another for reactive power response.

3.2 System Configuration on Testbed

In [137], the power control on V2G is briefly referred to here. The arrangement of the aggregator groups to the electrical branches of the V2G is illustrated in Figure 3- 1. The aggregator controls different electric vehicle groups to measure the state of charge (SOC) of each battery. The battery voltage is continuously monitored by mapping the SOC to typical battery charge characteristics to determine the battery charge percentage. The power utility monitors these aggregators in real-time to keep track of the battery energy availability status on each EV group.

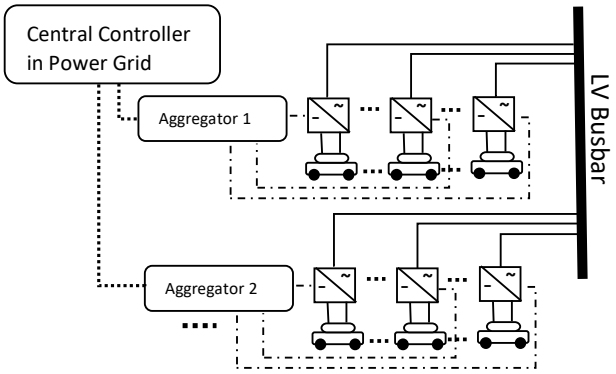

Fig. 3- 1 Configuration of the proposed V2G system

The system configurations of wind farms and solar power farms are shown in Figure 3- 2. Each electrical branch connects the general load, industrial motor, and a large group of EVs. The DIgSILENT PowerFactory software program is used to develop this testbed.

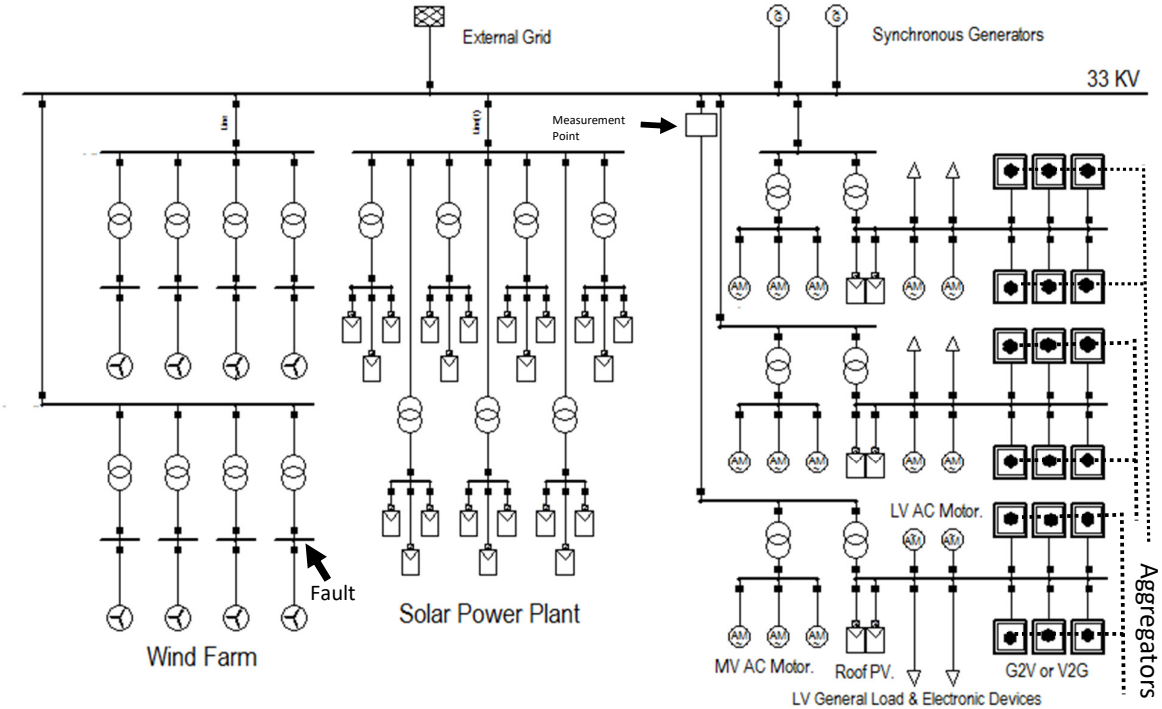

Fig. 3- 2 Single line diagram of the testbed

The nonlinear global optimization technique of Genetic Algorithms (GA) is used to estimate the load model parameters by comparing the measured output and the estimated output until the minimum error complies with the criteria. It is a heuristic probabilistic optimization technique based on the chromosome approach [138-142]. This technique does not require differentiation to generate efficient solutions in complex and mathematically sophisticated environments. It continues to produce new generations from the initial parameter range until the final solution locates the individual parameters of the load model with the best fitness of the selection criteria.

3.3 Hybrid Load Model Development

The power system stability analysis relies on the choice of load models, which can be developed through pseudo-measurement to simulate the practical running conditions of the power grid. The DIgSILENT PowerFactory software is used for pseudo-measurement on the testbed to generate synthetic data during a system fault. The priori method is adopted to determine the structure of the hybrid load model. The parameters of this model are estimated by GA to compare the estimated output power and the measured values. If it is a suitable model, the input-output system behaviour of the testbed can be described.

The structure of the proposed composite load model of active power response is shown in Figures 3- 3 and 3- 4. The dynamic portion of the equivalent induction motor load model (IM) as Differential-Algebraic Equations are referred to [33, 143-146] for the composite load model of active power response. The static part of the ZIP load model is expressed as algebraic functions of the bus voltage.

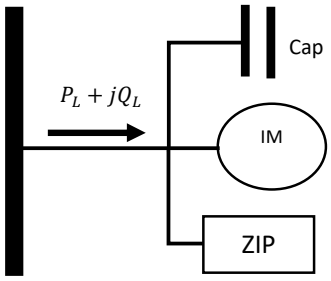

Fig. 3- 3 (ZIP + IM + Capacitor) load model diagram

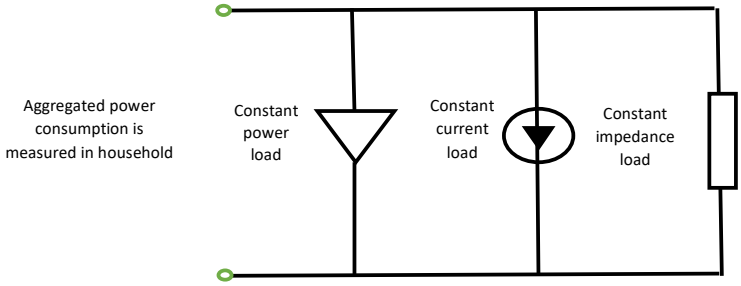

Fig. 3- 4 The equivalent circuit of the ZIP load

The equations represent the proposed composite load model of active power response, briefly described as follows:

$$P_d = \left[G_s \left(\frac{V}{V_n}\right)^2 - \frac{V}{X_m} E_m \sin \delta_m\right] e^{-E_d(T_2 - T_1)} \tag{3-1}$$

$$Cap = \frac{K_c - e^{\frac{-t}{RC_1}}}{K_c} \tag{3-2}$$

$$K_{pm} = \frac{P_{motor}}{P_0} \tag{3-3}$$

$$P_i = 1 - K_{pm} - P_Z - P_p \tag{3-4}$$

$$P_s = P_0 \left[P_Z \left(\frac{V}{V_n}\right)^2 + P_i \left(\frac{V}{V_n}\right) + P_p\right] e^{-E_s(T_2 - T_1)} \tag{3-5}$$

$$P_L = (P_d + P_s) \frac{K_c - e^{\frac{-t}{RC_1}}}{K_c} \tag{3-6}$$

where P_{motor} is the initial motor load, and P_0 is the initial active power load at the measurement point [33, 132].

Figure 3- 7 illustrates that the measured active power does not quickly rise to the previous level after the voltage recovers from a large disturbance. It indicates that load shedding may occur during the process of system fault and the voltage recovery period. A minus exponential function can be applied to simulate the load shedding operation in the load model [147].

The proposed composite load model of the reactive power response consists of the capacitor and the static exponential load model. The static load model is expressed as the algebraic function of the bus voltage as described in Eq. 3-7. The equation consists of an exponential form for the reactive load [34, 148, 149]. The value of β determines the load characteristic of the model.

The following approach expresses this composite load model of reactive power response.

Step 1:

The difference between the estimated reactive power from the static load model in Eq. 3-7 and a measured reactive power from the testbed after a reference fault disturbance, as shown in Figure 3- 2 is expressed in Eq. 3-9.

$$Q_d = Q_0 \left(\frac{V}{V_n}\right)^{\beta} \tag{3-7}$$

$$Cap = \left(1 - e^{\frac{-t}{RC_2}}\right) e^{\frac{-t}{RC_3}} \tag{3-8}$$

$$Diff = K_1 \left(\frac{V}{V_n} - 1\right)\left(1 - e^{\frac{-t}{RC_2}}\right) e^{\frac{-t}{RC_3}} - K_2 \tag{3-9}$$

Step 2:

$$Q_L = Q_d - R_1 K_1 \left(\frac{V}{V_n} - 1\right)\left(1 - e^{\frac{-t}{RC_2}}\right) e^{\frac{-t}{RC_3}} + K_2 \tag{3-10}$$

3.4 Identification of Load Model Parameter

The multi-curve identification method [132] is modified to identify the estimation errors of the model parameters for the proposed composite load models. Figure 3- 5 illustrates the identification process of the hybrid load model parameters.

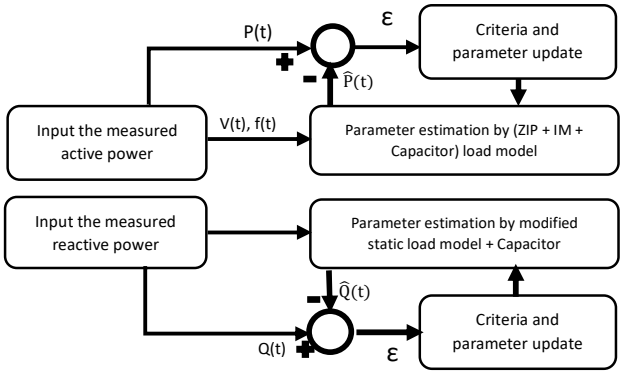

Fig. 3- 5 Identification process of the hybrid load model parameter

Assume that there are M sets of record data of the power system response collected from the measurement point, as shown in Figure 3- 2. The proposed composite load models of active and reactive power response are estimated in two individual paths to identify the best-fit load model parameters. They have the same identification procedure for the model parameters.

The estimation error of the model parameters for the proposed composite load model of active power response is illustrated as follows.

Referring to the procedure in [150], it develops the estimation error Eq. 3-11 to 3-15. The root-mean-square error between the estimated active power output of the (ZIP + IM + Capacitor) model and the measured active power output is expressed as:

$$e_{pi} = \sqrt{\frac{\sum_{k=1}^{N} W_p(k)[\hat{P}(k) - P(k)]^2}{N}} \tag{3-11}$$

where N is the number of sampling points of the active power output from the measurement, $\hat{P}(k)$ is the estimated active power output by MATLAB simulation, whereas $P(k)$ is the measured active power output at the k-th steps. The value 'n' of the weighting factor W_p is based on industrial practice.

$$W_p(k) = 0.4 \frac{\left[U(k) - \overline{U}\right]}{\sum_{k=1}^{n} \left[U(k) - \overline{U}\right]} + \frac{0.6}{n} \tag{3-12}$$

The relative fitting error is defined as follows:

$$\tilde{e}_{pi} = \frac{e_{pi}}{\overline{P}} \tag{3-13}$$

$$\overline{P} = \frac{\sum_{k=1}^{N} P(k)}{N} \tag{3-14}$$

The loss function is formulated as follows:

$$min \frac{1}{M} \sum_{i=1}^{M} W_i(\tilde{e}_{pi}) \tag{3-15}$$

where W_i is the weighting coefficient for a specific measurement.

3.5 Testing of the Proposed Hybrid Load Model

The proposed hybrid load model is tested by MATLAB, based on the simulation for a single-phase to earth fault scenario at the busbar of the doubly-fed induction generator (DFIG) in 0.25 s, as described in Figure 3- 2 [151, 152]. Figure 3- 6 shows that the voltage dropped significantly below the nominal voltage at the measurement point.

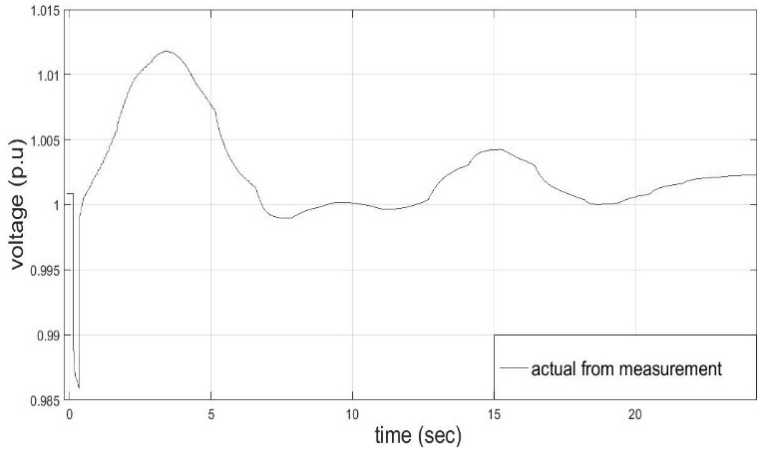

Fig. 3- 6 Voltage response during a fault disturbance of the testbed

Figures 3- 7 and 3- 8 illustrate that the fluctuation in both active power and reactive power, caused by the voltage dip during a fault disturbance, is much shorter than the voltage recovery period. Therefore, it does not have a significant effect on calculating the accuracy of the load model. In addition, the fault clearing time can affect power system stability if the tripping of the protection relay is not fast enough to clear the fault. The results indicate that the hybrid load model has good accuracy in reproducing the load dynamics of the testbed.

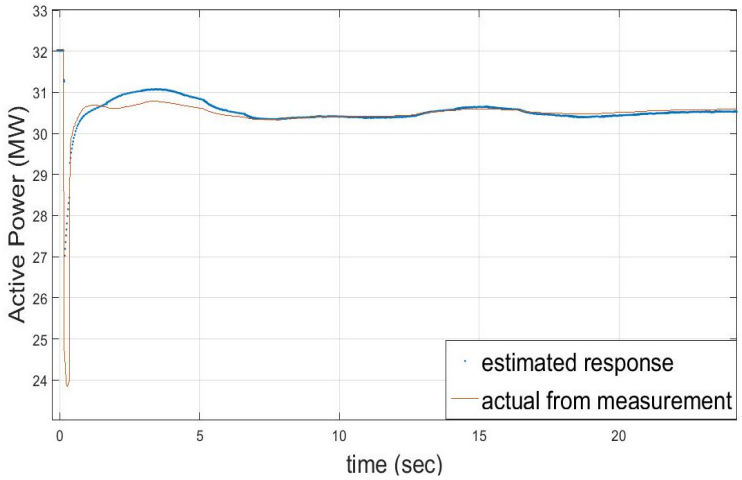

Fig. 3- 7 Comparison between the estimated active power output of the proposed composite load model and measured active power output

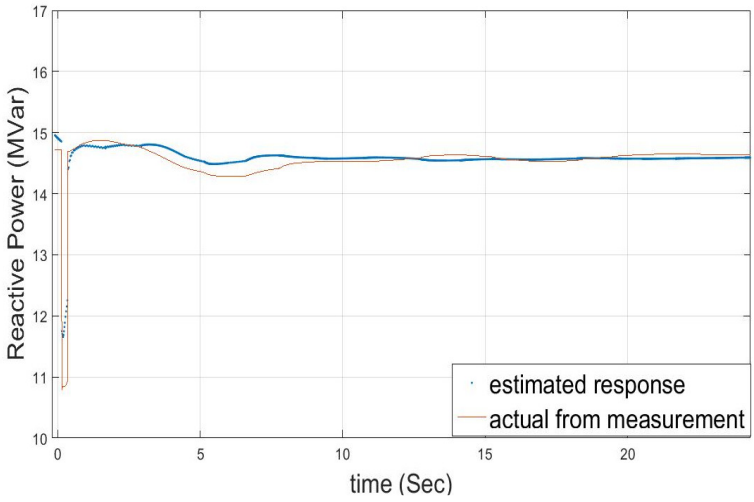

Fig. 3- 8 Comparison between the estimated reactive power output of the proposed composite load model and measured reactive power output

The estimated active power response and measured active power response are compared, as shown in Figure 3- 7. It indicates that load shedding may occur during the fault simulation. Referring to Eq. 3-11 to 3-14, the relative fitting error is 0.00054. The estimated reactive power response and measured reactive power response are compared, as shown in Figure 3-8. Referring to Eq. 3-11 to 3-14, the relative fitting error is 0.00019.

Figure 3- 9 illustrates that the conventional composite load model has a low comparative accuracy in reproducing the load dynamics of the test network if the drop of the reactive power is high, e.g. the relative fitting error of the reactive power is 0.0028. The dynamic part of the conventional load model does not respond well to the reactive power change during fault disturbance.

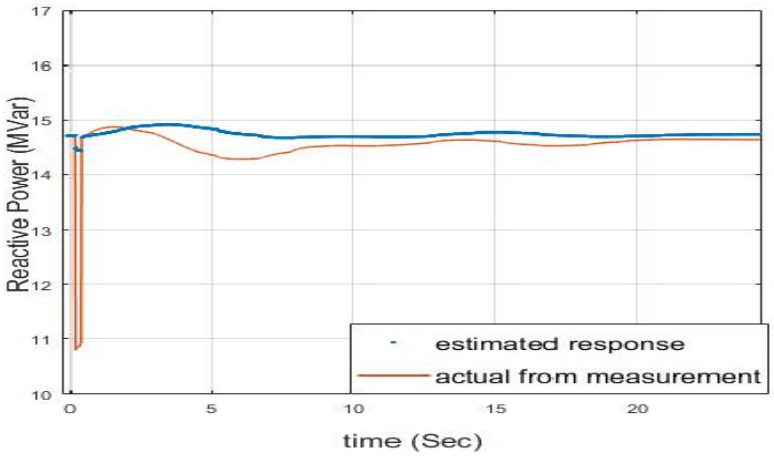

Fig. 3- 9 Comparison between the estimated reactive power output of the conventional composite load model and measured reactive power output

3.6 Conclusion

This research investigates the possibility of developing a new hybrid load model with physical meanings or reliable mathematical formulas to represent the load characteristics of the power grid in an environment of high integration of renewable energy sources, e.g. wind farms, solar power, and electric vehicles EV. The result indicates that the new hybrid load model has a reasonably low fitting error than the conventional load model if the drop of the reactive power is high after fault disturbance.

4 Transfer Learning Approach for Non-Intrusive Load Monitoring using Generative Adversarial Networks

4.1 Introduction

Climate change and a gradual rise in energy demand have been critical issues in recent years. According to the U.S. Energy Information Administration (EIA) and International Energy Outlook 2019 (IEO2019), it is predicted that world energy consumption will grow by almost 50% over 2018 to 2050 [1]. Different measures have been adopted to tackle these challenges, e.g. the renewable energy development strategy for clean energy, resource planning, and demand-side management. Even a 15 per cent reduction in electricity use among homes can generate substantial positive effects on environmental protection in the United States. This energy consumption is equivalent to 81.3 million tons of coal per year [2]. So, it is required to better analyse the customers' electricity consumption behaviours by collecting disaggregated power consumption data of the different appliances. Such a result will benefit the demand-side management with a consequential price reduction in electricity bills.

Appliance load monitoring (ALM) techniques, including both intrusive and non-intrusive load monitoring (ILM and NILM), have been employed to measure the disaggregated data of electricity usage [153]. The generated dataset can support the power utility companies to re-schedule the high-power demand to off-peak periods. Also, the utility provides a specific incentive scheme of different charge rates for electricity bills to motivate the users to change their electricity consumption habits in demand-side management [154]. The ILM is a traditional load monitoring method, e.g. supervisory control and data acquisition (SCADA) requires sensors to be installed at different load points for communicating with the monitoring and control networks. Its shortcomings are the concerns of compromising system

reliability for higher hardware requirements and communication platforms. In addition, such an ILM network is not practical for household users at a cost consideration [155, 156].

In [157], NILM was introduced to measure the detailed consumption of individual appliances from the aggregated power signal in a house. The raw signal of total power consumption is disaggregated by software programming to recognise different appliances through signal processing and mathematical modelling. The NILM can be divided into two techniques: event-based and non-event-based [96]. The event-based technique captures the transient features during the switching events in high sampling frequency, e.g. the state change from on to off [158-160]. The wavelet transform (WT) is used for non-periodic and non-stationary signals, and the discrete wavelet transform (DWT) is a popular model in this area [161]. In [156] and [162-165], short-time Fourier transform (STFT) and wavelet transform were used to identify the turn-on time of load and the duration of transients for frequency components as power signatures. In [166], support vector machines (SVM) were used for pattern recognition on load signatures after the S-Transform extracts the feature. This approach performs similarly or even superior to the wavelet packet transform (WPT). In [167], load transient state and steady-state characteristics were used as signature features to identify the multiple appliances using a multiple-layer feedforward neural network. However, the transient analysis approach has the drawback of using a high sampling rate and large memory size to capture the signal sequences. It incurs extra costs and algorithm complexity. The business requirements work group of the Australian National Smart Meter Program (NSMP) has set up the smart metering infrastructure minimum functionality specification. As a simplified version, the smart meter shall calculate active power every 5 seconds [168]. In [169], the author also supported that the available smart meters did not provide the standard functionality of high-frequency measurement. However, the Queensland Competition Authority reported that all benefits of engineering data, including power,

current, harmonic, etc. to be a potential benefit of advanced digital meters were considered [170].

The non-event-based technique takes all power states into account. It does not detect any event transition. According to [169, 171], this technique typically works with data sampled at 1 Hz or even lower with multi-states of an appliance. In this paper, the field measurement on the appliances is sampled in 3-second intervals. To make the NILM work properly, it is a challenge to develop an effective algorithm that can recognise the power signatures of each appliance by measuring the raw signal from the utility service entrance. The hidden Markov model (HMM) and its variants are popular solutions in real power measurement on low-resolution data [171-176]. Its modelling dynamics on time series observation refers to power readings or differential power readings, while the hidden states represent the appliances. However, this model does not work well for a large number of appliances which increases the complexity of the algorithm in time and space [177]. In [178], the author demonstrated that the fusion of two deep learning-based models could outperform the HMM. First, the user's appliance usage habits are used to train one model for predicting the next appliance usage. Second, another model is trained using the characteristics of appliances with temporal and electrical information.

Karhunen loeve expansion (KLE) [179] and Bandt-Pompe method [180] are also common choices to improve the performance of NILM using the sliding window for feature extraction in the time series. In [169, 179, 181], the KLE was proposed to extract a set of power signal features optimal for representing a pattern class of appliance features whose observable is a random process. But KLE has difficulty tackling the presence of unknown or non-target appliances. The Bandt-Pompe method transforms time series into a sequence of patterns using a non-parametric transformation. The Shannon entropy from the histogram of causal

patterns is then computed to detect relevant causal information related to the unobserved variables that control the system. In [180], the Bandt–Pompe methodology combined with the causality complexity–entropy plane was applied to characterise the regimes and electrical load behaviours over the data. It showed that this characterisation provides a valuable insight into the underlying dynamics that govern the electric load. In [182], an image encoding method of triadic state-space construction (TSSC) was proposed to generate images for classifying the chaotic time series with a convolutional neural network (ConvNet). The triadic state-space extends the ordinal orders of triadic time series motifs to outperform the Bandt-Pompe method.

In [183], the convolutional sequence-to-sequence (seq2seq) model was proposed to generate the sequences of disaggregated power consumption for a single appliance. The convolutional model takes the input x of the length of the sequences l_{in} and maps it to the output y of the length of the sequences l_{out}. The window of each y is aligned to the centre of each x. The windows of the output and input sequences are moved with a small step size sΔ, which is much smaller than l_{out} on each time step. Then, the output is calculated by averaging the values on all the time steps. Gated linear unit convolutional layers and Residual blocks are employed to extract useful information from the aggregated electricity consumption sequence and refine the output of the neural network, respectively. In [184], the convolutional sequence-to-point model (seq2point) proposed that the input sequence is a sliding window of the mains, and the network is trained to predict the output signal at the midpoint of the corresponding window of the target appliance. So, W is the size of the sliding window and T is the length of the input signal; the seq2point method predicts the T outputs.

In [185], the current harmonics were used as a signature feature to identify multiple appliances using a feed-forward neural network with the Levenberg-Marquardt

algorithm. In [186], the different power states of individual appliances were clustered using the *k*-means algorithm for combinatorial optimization (CO) and a binary on/off vector for the manual labelling of recurrent neural networks (RNN).

In [187], a NILM architecture focus on current waveform images of the consistent shape loads that the same type of loads to be highly similar was proposed. It distinguishes the different types of loads from different shapes of load waveform images. After detecting events, the steady-state current of different events is extracted under the same initial phase angle of voltage sampled at high frequency to satisfy the superposition of currents. A CNN is used to identify the current waveforms of consistent shape loads, which are in a low correlation with sinewaves. These are 2D waveforms drawn from the current amplitude as the ordinate and the sampling point as the abscissa.

In [188], a multi-layer perceptron (MLP) approach with the kurtogram technique was proposed. The MLP is adjusted based on batch size and nodes to recognise the load signatures as image patterns extracted from the kurtogram algorithm. It is the fusion of multi-level kurtosis values to convert 5,000 sampling data into one image.

In [189], a Siamese network for few-shot learning was proposed to calculate the similarity of V-I trajectories of different electrical appliances and the database from the feature library for load identification. The active power and V-I trajectory are used as the load signatures to identify the electrical appliances with similar V-I trajectories but different active power.

In [190], an approach using two deep convolutional generative adversarial networks (DCGANs) was proposed. The low-frequency sampling data at 1 Hz is transformed using an algorithm of vector-to-matrix time series into the images whose vector of length is 64 is equivalent to 1 minute plus 4 seconds of the data measurement. One of the DCGANs is

trained on the aggregated power data for the appliance at on-state. Another is trained for the appliance off-state. Therefore, there are two determiners to classify the generated image and the real image. After the two DCGANs are trained separately, the two discriminators compare individual predictions. If the output of the discriminator trained on the off data is greater than that of the discriminator trained on the on data, the appliance is judged to be off status. Otherwise, it is in the on status. However, the accuracy of identifying devices with limited data for on-off or off-state is not very high.

In [191, 192], an approach for inputting the grayscale images into CNN was proposed for identifying the linear and nonlinear loads. The frequency invariant transformation of the periodic signal (FIT-PS) method is used for the detection of the zero-crossing point of the voltage waveform, calculation of the voltage zero offsets, position of the sampling point in each cycle, and interpolation pre-processes the signal from the power main with a high sampling frequency. Then, the Euclidean distance between the current waveform and the sine waveform determines whether it is a nonlinear load from the pre-set threshold. The data of 1D current waveforms from the linear load is used for training the CNN classifier. The transient power signal of the nonlinear load is converted into 2D L × L size grayscale images for another CNN classifier training. The grayscale between black and white represents the current amplitude values. L cycles of the current waveform with L sampling points are arranged into a 2D array to produce a grayscale image.

In [193], a multi-task federated learning algorithm (MTFed-NILM) for NILM was proposed to preserve the users' privacy by exploiting the dependencies among multiple tasks. A FedAvg strategy [194] is applied in that the clients transmit the updated gradients of the DNN models to the central server after the models are trained with local data. Then, the server calculates the weighted average of the gradients and broadcasts it for synchronous

training to each client. The U-Net is used to identify different load usage patterns of multiple appliances by extracting the shared features. Its encoder-decoder architecture uses the 1D convolutions and 1D transposed convolutions to refine high-level semantic features and perform feature localisation, respectively. These local features are concatenated with the high-level feature maps to generate a multi-scale feature map. Then, a stacked 1D CNN block extracts the latent features of the signal to produce a latent feature map, which contains representations of the non-temporal features and appliance usage behaviours shared by multiple tasks. The multi-task learning consists of multi-label classification and multi-target quantile regression. The latent feature map is further converted into a lower-dimensional representation by the fully connected layers in the multilayer perception block to diverge into two task-specific layers that conduct state recognition and power estimation separately.

To conclude from the above algorithm development, the ANN application significantly enhances the performance of NILM through an adequate mathematical model. However, the ANN is hungry for the labelled data of NILM, which may not be available in some situations, e.g. new appliances. The insufficient labelled data on the training set has a negative impact on the performance of the supervised learning algorithm on the testing set. It may also be costly to obtain the labelled data. In general, a larger dataset with a wider variety can benefit the training on the deep learning network. On the other hand, small datasets may lead to overfitting. It may work with a simple machine-learning algorithm with low complexity. The existing works of AI models in NILM cannot manage model training in insufficient data. As such, a new approach is required to make a difference.

Data augmentation (DA) is a common practice to increase the dataset size of limited training data in deep learning [195]. This technique includes colour modification and affine image transformations by rotation, reflection, scaling, shearing, geometric distortion,

deformation, etc. to make the classifier training model more robust. But this technique has a deficiency of data bias that the augmented data distribution can be quite different from the original one. Moreover, it is not practical in some cases, e.g. mnist dataset. So, it depends on how the data are presented.

On the other hand, GANs have a well-designed adversarial game. It is one of the most popular unsupervised methods to be considered as an alternative solution [196-198]. The ability of extrapolation is the advantage of traditional DA over GANs, which cannot extend the data distribution exceeding the limit of the available training data. But GANs can analyse the underlying distribution of training data to fill in gaps in the discrete data distribution. It also helps to augment the sources of variance for additional synthetic data on real training data. In this paper, GAN is proposed to solve the problem of limited data on NILM by generating synthetic data.

The proposed NILM approach is different to the existing works. For example, referring to [184], the convolutional sequence-to-point model proposed the input sequence of aggregate power consumption to be a sliding window of the mains. The network is trained to predict the output signal value at the midpoint of the corresponding window of the target appliance, e.g., active power. So, it is a NILM approach for one-dimensional sequence input and signal value prediction output. It requires checking the accuracy between the estimated output value of an appliance and the real value by metric, e.g. MAE, RMSE, comparing the results of signal disaggregation from each appliance, i.e. the generated and real datasets, etc. It is an approach for applying prediction with one-dimensional input in a neural network. However, the proposed approach employs a different methodology. The data of each appliance is presented in a two-dimensional (2D) image of a categorical histogram measured at each point in time. GAN is used to generate the 2D data for training on DenseNet and ResNet

classifications from a limited dataset of specified features. Also, the NILM framework is new.

The NILM technique can recognise different appliances through ANN and signal processing. However, the ANN is hungry for labelled data of NILM, which may not be available in some situations. My research objectives are to apply transfer learning to GANs to generate sufficient synthetic data for the training set in supervised learning algorithms with limited data. This aims to address the issues of training performance in the scratch ANN model when the training dataset is insufficient, as well as improve mode collapse and long training iterations during GAN training.

4.2 Generative Adversarial Nets

In the deep learning era, the hierarchical artificial intelligence (AI) model was developed to represent the data probability distributions during the training stage. This process is primarily based on backpropagation and dropout algorithms with good gradient linear units. As a result, it obtains excellent outcomes for discriminative classification problems. However, as discussed, one of the biggest problems facing the deep learning application is the lack of sufficient training data or uneven class balance within the labelled datasets due to high-cost consideration, expert knowledge, etc. to obtain the data. This issue may cause the ANN model to fall into overfitting during the training stage, e.g. the discriminator in GANs may fall into divergence, and then the model collapses. Furthermore, if the discriminator fails to generalise as it only memorises the same training data from a limited data dataset, the validation accuracy of the discriminator falls dramatically as little feedback is received to improve the generated images.

The basic structure of GANs is summarised here [196].

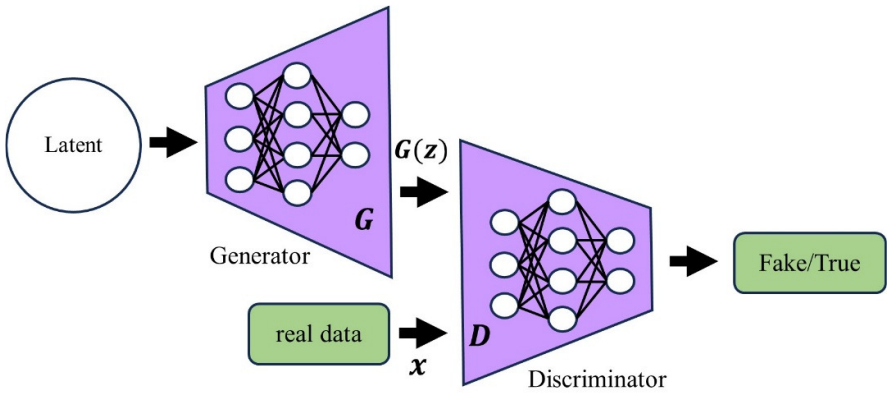

Fig. 4- 1 The structure of generative adversarial network (GAN) model [199]

A minimax two players game model was proposed to estimate the generative models via an adversarial process: a game player of generative model G is trained to recover the training distribution of real data, and another game player of discriminative model D to determine if a sample comes from synthetic data or real data. The training process with backpropagation in Eq. 4-1 continues to learn the hyperparameter until real and synthetic data distributions are indistinguishable.

$$\lim_{n\to 0} \nabla_x E_{\epsilon\sim N(0,\sigma^2 I)} f(x+\epsilon) = \nabla_x f(x) \tag{4-1}$$

In GANs, the data of high dimensional vector generated by model G is input to the model D that a saddle point finally reaches between this minimax two players game model to end the training in Eq. 4-2:

$$\min_G \max_D V(D,G) = E_{x\sim P_{data}(x)}[\log D(x)] + E_{z\sim P_z(z)}[\log(1-D(G(z)))] \tag{4-2}$$

where $p_z(z)$ is input noise variables, G is a differentiable function represented by a multilayer perceptron with parameters θ_g, $G(z;\theta_g)$ is data space, $D(x)$ is the probability that x came from the real data, $D(x;\theta_g)$ is data space, $V(D,G)$ is a value function. D is

trained to maximise the probability of assigning the correct label that the training of G makes $\log(1 - D(G(z)))$ minimum.

In [200], a neural NILM model was proposed that the load profile of mains was split into mains sequences to feed the appliance-specific disaggregators. Then, the outputs from the disaggregator were merged to form the appliance load. In [201], auxiliary classifier GAN (AC-GAN) was proposed to enhance the neural NILM approaches of low-quality shaped appliance load profiles from reference [200] by generating the load sequences of appliances from disaggregator outputs. In [202], an EnerGAN model, which consists of three core components, i.e. seeder S, generator G and discriminator D, was proposed to support the fine-tuned dissagregators for producing accurate load estimations of a specific appliance. The seeder encodes the input noise aggregate time-series data to reduce the non-linear dimensionality. Then, the data is mapped to a compact and informative subspace to produce the required latent vectors to the generator G as a sense for guiding the generation process for the power waveform of an individual appliance. Generator G and discriminative D are used to capture the active power signal distribution of the appliance and estimate the probability that the extracted appliance power signal is real, respectively.

Some efforts have already been made to combine traditional DA techniques and GANs to generate image data, e.g. the training of ANN model for research on the human body in medical science. In [203], the experimental result indicates that the traditional data augmentation method has better classification accuracy on diabetic retinopathy (DR) than without augmenting any data. In [203, 204], a training scheme was proposed that classical data augmentation was used to increase the dataset size before applying GAN to produce more data and diversity in the dataset for additional data augmentation to a classifier. The

combined effects of the traditional data augmentation method and GAN outperform the above two approaches separately.

The above findings have supported that progressive growing of GANs (PGGANs) is used to derive synthetic data for three datasets, i.e. computed tomography (CT) images with manually delineated cerebrospinal fluid (CSF) labels segmentation, fluid-attenuated inversion recovery (FLAIR) magnetic resonance (MR) images with manual WMH segmentation, and DeepMedic by both GAN and rotation augmentation. The simulation result indicates that the insufficient size of a real dataset for AI training under the most limited conditions can be significantly improved in Dice Similarity Coefficient (DSC) [205].

In [206-208], it further points out that if the concept of DA is only applied to the real data in GAN, it misleads the generator to produce synthetic images which match the augmented image distribution P_{data}^{T} rather than original real images P_{data}. It is because the real data distribution has already been significantly altered after augmentation. Reference [209] also supports that directly applying DA to GAN does not improve the baseline.

In [206], the method of differentiable augmentation (DiffAugment) was proposed to apply the same differentiable augmentation on both real and generated synthetic data. It improves the data efficiency for both generator and discriminator training. Also, it avoids breaking the exact balance between the generator and the discriminator. So, if the augmentation operator T is differentiable with respect to the input, the gradients can be backpropagated through the augmentation operator T to update the generator G.

Similar concepts are found in [207, 208]. In [207], a principled framework of data augmentation optimised for GAN (DAG) was proposed to improve the learning of distribution from original data P_{data} by using an augmented dataset X^{T} with samples

transformed by a transformation T, $\{T_k\} = \{T_2, T_3, \dots, T_K\}$. The discriminator D_k is trained on real samples and fake samples to be invertible transformed by the same T_k that different discriminators D of $\{D_k\} = \{D_2, D_3, \dots, D_K\}$ with weight sharing process of different transformed samples to preserve the jensen-shannon (JS) divergence of original GAN with invertible transformations. It ensures the training of D_k to achieve the same optimal as D in the distribution of original data P_{data}.

In [208], stochastic discriminator augmentation was proposed to use the augmented images for evaluating the discriminator on the leaking issues. It also applies to generator training. According to [210], if the GAN is trained under a corrupted data distribution, and the corruption process is represented by an invertible transformation of probability distributions over the data space, GAN can achieve the correct data distribution in non-leaking. As a maximal set of various augmentations for every image to discriminator inputs supposed to be helpful to the GAN training, the method of adaptive discriminator augmentation (ADA) was proposed to control the degree of overfitting by incrementing/decrementing the augmentation strength p, in which the same value of p is applied for all invertible transformations.

4.3 A Transfer Learning Algorithm

However, the approach that applies the exact differentiable augmentations on both real and synthetic data for all invertible transformations to GAN is not a unique solution to improve the efficiency of convergence properties, stability issues and mode collapse on GAN. Transfer learning is also a possible alternative solution. In reference to [211], the concept of transfer learning is explained in Section 5.3.

In [212], a transfer learning-based active learning framework was proposed to maintain high classification accuracy if the size of required labelled training data for the scratch ANN model is insufficient. The technique of finetuning can be applied to GAN to enhance the quality of synthetic images on domain adaptation with limited training data by using prior knowledge. In [213], a residual neural network (ResNet) structure was used in both generator and discriminator with 1000 target samples to verify the effect of domain transfer for GAN. The results show that the transfer learning technique applied to the discriminator performs better than the generator, and the best is the domain transfer for both the discriminator and the generator. However, this technique may lead to the mode collapse if all the pre-trained parameters of the generator adapt to the limited training data [214]. Furthermore, the input noise distribution must be cut short when the information is inferred to generate realistic images if only the relationship between the vectors and limited training data is considered in model training.

In [215], a minor network MineGAN, which is less robust than the pre-trained GAN, was proposed to steer the latent space of generator G to a specific region of limited data distribution. Then, the weight of generator G is updated by finetuning. When only the learnable parameters to the batch normalisation are finetuned, a smaller size of trainable parameters can control the overfitting. This approach is to find a new target generative distribution by locating those regions in the generative distribution of pre-trained GAN that better approximate the target distribution of the limited data $P_{T\ data}(x)$. It basically identifies a narrower latent distribution region $P_{T\ z}(z)$ that samples $G(z)$ with target $z \sim P_{T\ z}(z)$ are just like samples collected from the target distribution of limited data $P_{T\ data}(x)$. So, the original input noise variable $u \sim P_z(u)$ is transformed to follow the $P_{T\ z}(z)$. The variables of multilayer perceptron M for the miner network play an interface between the input noise u and generator $G(M(u))$. Then, the gradients are backpropagated to M to finetune its

parameters during model training. This strategy can gear the miner network to align to the most relevant input regions of prior $P_z(u)$ and generate images close to the target domain D_T. It also avoids mode collapse.

In [216], a transfer learning method based on the efficient-net model and time-series generative adversarial network was proposed to tackle the severely imbalanced dataset of industrial rotating machinery in machine fault diagnosis. The proposed model uses a time-series GAN based on an adversarial network and an auto-encoder to extend the imbalanced fault diagnosis of rolling bearings. It synthesises the dynamic and static features of time-series vibration signals of imbalanced categories. Then, transfer learning is applied to the balanced vibration signals to train the efficient net on classification.

In [217], a GAN-NILM model was proposed as an end-to-end load disaggregation approach to improve the accuracy of load disaggregation. It consists of a generator G and a discriminator D. The generator G takes the input x of each aggregate measurement sequence to generate the disaggregated appliance measurement $G(x)$ by matching feature. The $G(x)$ is the appliance consumption. Then, $G(x)$ is the output input concatenated with the corresponding real aggregate measurement and classified by the discriminator. The selective model parameters of GAN-NILM between the source domain and target domain are shared to maximise the generalizability. The author also proposed a TrGAN-NILM model to minimise the statistical distance between the source domain D_S and the target domain D_T in feature space by learning the common feature representations between the two domains. The generator G consists of G_e and G_d. G_e maps the aggregate measurements for the source and target domains to two sets of compact representations, respectively. Another auxiliary discriminator D_{tr} is used to achieve a minimum statistical distance between these two

encoded representations for the same distribution. Then, G_d maps the encoded compact representations to the corresponding appliance measurements.

Some advantages of transfer learning are listed as follows:

- It offers the potential to apply pre-trained models to new tasks, particularly when there is limited data available. By utilizing the knowledge learned from a pre-trained model, pre-trained models can enhance performance in the new task [218]. This approach enables developers to avoid the need for large amounts of new data, as a model trained on a task with abundant labeled data can handle similar tasks with less data [219]. Additionally, transfer learning allows developers to reuse and adapt pre-trained models, saving time and reducing costs in the development of AI solutions [220]. It plays a critical role in making sophisticated models like deep neural networks more accessible, as they typically require significant resources and computing power [219].
- It addresses the issue of limited or scarce data for a new task by utilizing pre-existing weights and biases, adapting them to the new data. This approach offers the advantage of reducing the amount of data and annotation needed to train a model. By leveraging pre-trained weights from a related task, it saves time and resources, particularly for complex tasks or when working with extensive data [218].
- It utilizes pre-trained models to reduce training time and achieve faster and more efficient learning. By leveraging existing knowledge from a related task, the model starts with pre-trained weights, allowing it to converge faster and achieve better results in less time [218, 219, 221].
- It enhances the accuracy and generalization of neural networks by utilizing knowledge from a larger and more diverse dataset. transfer learning improves

model accuracy by leveraging pre-trained models and helps address the issue of overfitting by providing supplementary training data [220].

- It improves model performance by using pre-trained weights from a related task after summarising the above advantages. Transfer learning enables faster convergence, better results, and the ability to apply pre-trained models to new tasks, especially when data is limited [218, 222]. By leveraging the knowledge acquired from a related task, transfer learning boosts model performance, providing a better starting point and a higher learning rate [221]. This approach allows models to start with higher performance levels, benefitting from a pre-trained source model before further refinement [221, 222].

4.4 A New Application Framework on Non-Intrusive Load Monitoring

In reference to Section 4.1, the traditional augmentation technique may not work well in some situations, e.g. the data are represented in a histogram. In [169], the author proposed a data augmentation solution to use customers' aggregated power consumption profiles. These data are readily available from the power utility. Such data with a similar distribution of limited data datasets can be pre-trained both generator and discriminator of GANs for future use in transfer learning.

The transfer learning-based approach to the NILM framework is illustrated in Figure 4- 2. The power profile of the new appliance is uploaded from the smart meter to the pool of database in the power utility. This dataset is converted into 2D form to be accessed if the meaningful data is sufficient to continue the training on the AI model. If the dataset is classified as a limited data dataset, a pre-trained GAN model, available from the database, is used for transfer learning with the uploaded limited data dataset. Then, GAN generates

sufficient synthetic data, which is used to support the training of the AI model to operate on NILM. Finally, the updated AI model is downloaded to the smart meter after being trained.

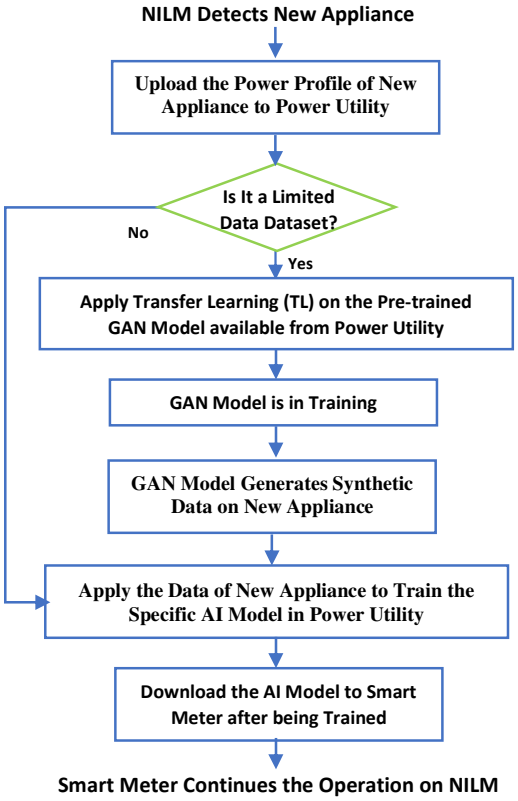

Fig. 4- 2 Flow chart of the proposed transfer learning framework on NILM

In this paper, DenseNet [223] and ResNet [213] are used to demonstrate the effectiveness of the transfer learning approach on multi-class classification. These neural networks are very powerful for the feature analysis of image patterns. This strength is employed for 2D histogram analysis for different image inputs.

ResNet has a skip connection to bypass the non-linear transformation $H_\iota(*)$ with an identity function in a feed-forward fashion. The identity function makes the gradient flow from later layers to earlier ones, as shown in Eq. 4-3. DenseNet shares a similar approach to connect each layer to all preceding convolutional blocks so that high-level features may be re-used in the deeper final layers. Therefore, the DenseNet differs from ResNet through

different layers of learning the concatenating feature maps. The Dense blocks connection exists in the network. It improves the performance of DenseNet by bypassing settings of information flow, shown in Eq. 4-4.

Assume X_0 to be a photo input for a convolutional feed-forward network. The network comprises L layers, and the feature output of X_ι at the ι^{th} layer for ResNet and DenseNet are formulated as:

$$X_\iota = H_\iota(X_{\iota-1}) + X_{\iota-1} \tag{4-3}$$

$$X_\iota = H_\iota([X_0, X_1, \ldots, X_{\iota-1}]) \tag{4-4}$$

where $H_\iota(*)$ is a non-linear transformation at the X layer such as Pooling, BN, ReLU or Conv. [] is the concatenation of all preceding feature maps that the feature size is the same in the same Denseblock.

DenseNet has a very narrow layer structure. A small growth rate of κ makes it outperform the other neural network. If κ feature maps are the $H_\iota(*)$ output, there are $\kappa_0 + (\iota - 1) \times \kappa$ feature maps input at the ι^{th} layer.

4.5 Database Generation

In reference to the available data, current, real power, reactive power, and harmonics are selected to represent any appliances listed in Tables 4- 1 to 4- 3. As GANs are unstable to train, a deep convolutional GAN (DCGAN) with 36M parameters for generating images of 64 x 64 is more stable in training [224]. Also, reference [225] reported that the results produced by different GAN architectures were, on average, not significantly different. In DCGAN, both the generator and the discriminator are based on convolutional neural networks (CNNs). This paper uses it to generate three thousand images on each appliance at

an acceptable computational cost. The database is generated by field measurement at a 3-second interval on different appliances. According to the field measurement data, these appliances are grouped into three categories, i.e. four for limited data, seven for one-state datasets, and another three for multi-state datasets, as shown in Tables 4- 1 to 4- 3, respectively. Those datasets grouped into limited data datasets are also one-state datasets. The heater is chosen to represent the electrical profile of the one-state dataset, as illustrated in Figures 4- 3 and 4- 4. Also, the refrigerator is chosen to represent the electrical profile of the multi-state dataset, as illustrated in Figures 4- 5 and 4- 6.

Table 4- 1 Limited data datasets

Appliances	Microwave (ME)	Vacuum (VM)	Kettle (KE)	Air Purifier (DN)
Data size	76	66	79	99

Table 4- 2 One-state datasets

Appliances	Television (TV)	Hi-Fi Audio Amplifier (AP)	Desktop Computer (DC)
Data size	1097	495	724

Appliances	Oven (ON)	Dryer (DR)	Heater (HR)	Freezer (FR)
Data size	862	2040	722	1869

Table 4- 3 Multi-state datasets

Appliances	Air Conditioner (AC)	Cooktop (CT)	Refrigerator (RR)
Data size	3920	1463	3718

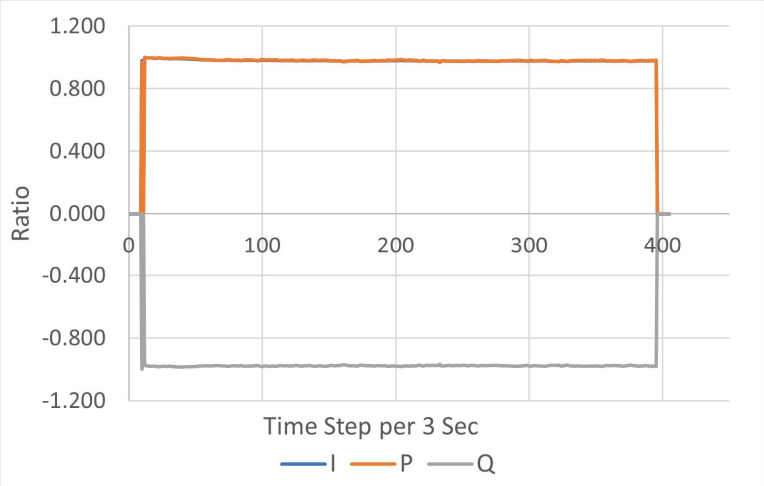

Fig. 4- 3 Electrical profile of current (I), active power (P) and reactive power (Q) on heater (HR)

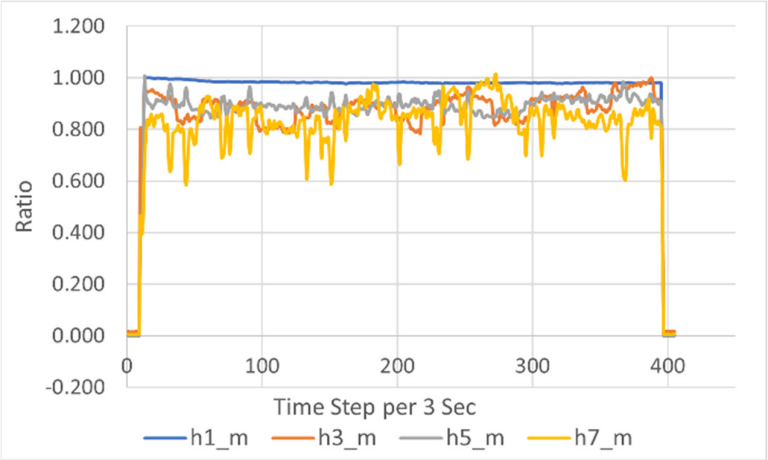

Fig. 4- 4 Harmonics (h1, h3, h5, h7) on heater (HR)

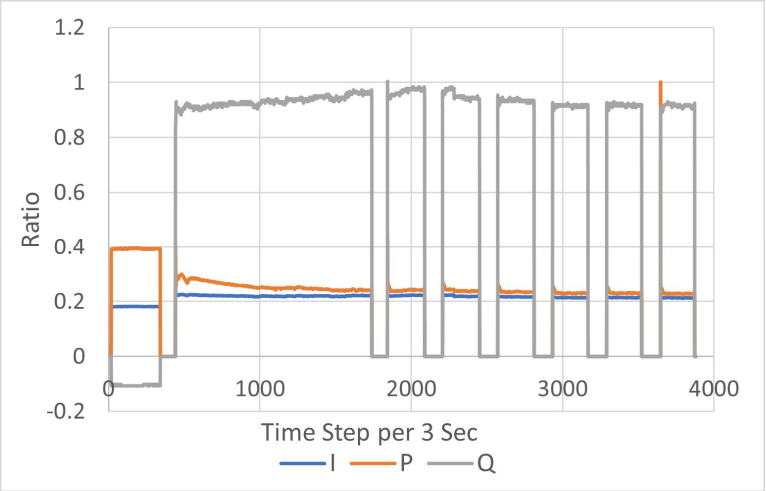

Fig. 4- 5 Electrical profile of current (I), active power (P) and reactive power (Q) on the refrigerator (RR)

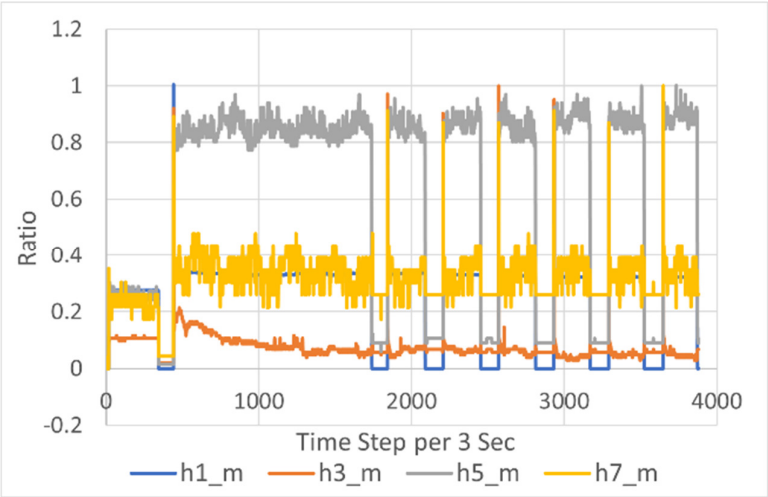

Fig. 4- 6 Harmonics (h1, h3, h5, h7) on refrigerator (RR)

4.6 NILM Implementation Architecture

In [169], the author proposed to retrieve the working power profile of the appliances by smart meter measurement from the pool of target appliances specified by the subscriber in the cloud of the power utility. This concept is modified to re-apply to the total aggregated power consumption profile as demonstrated in AP, FR and TV, shown in Figures 4- 7 to 4- 8.

Stage 1: NILM subscribers registered the brand and model for appliances AP and FR in a household to the pool of appliances database in the power utility. The trained model is downloaded from the cloud of the power utility to the smart meter to perform NILM, shown in Figure 4- 7.

Stage 2: The smart meter measures the total aggregated power consumption of appliances AP, FR and TV to upload the total aggregated power consumption profiles to the cloud of the power utility. As the datasets of AP and FR have already been stored

in the appliance pool of power utility, the disaggregated power profile of the new appliance TV can be extracted for training the AI model. Then, the trained AI model is downloaded to the smart meter, shown in Figure 4- 8.

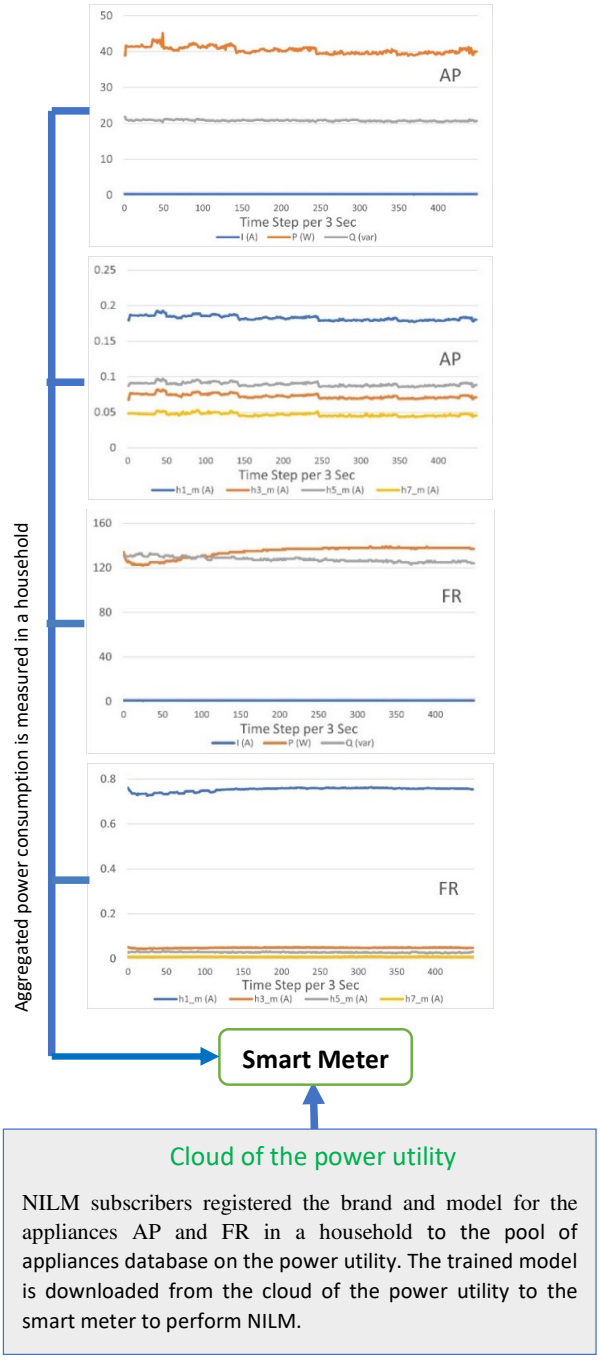

Fig. 4- 7 The registration for appliances of the hi-fi audio amplifier (AP) and freezer (FR) in a household to the pool of appliances database on the power utility

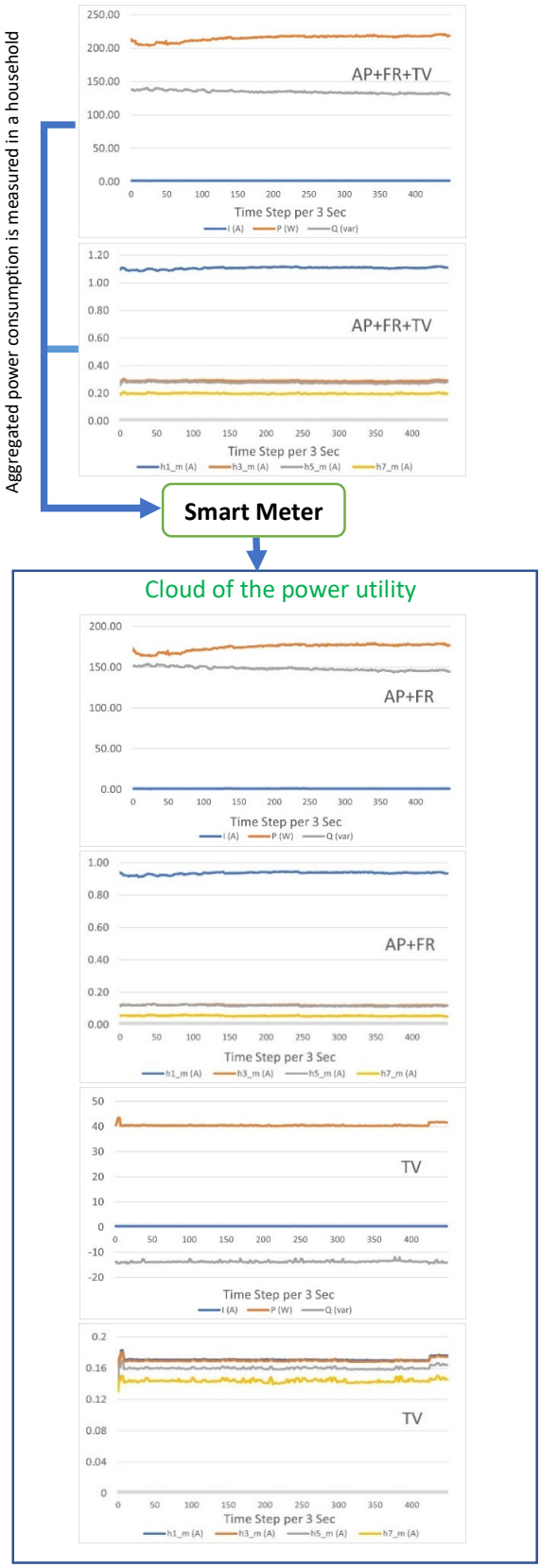

Fig. 4- 8 The extraction of the power profile of the new appliance of television (TV) from the total power consumption of the AP, FR and TV for the AI training in the cloud of the power utility

NILM subscribers registered the brand and model for appliances in a household to the pool of appliances database in the power utility. It helps to identify the power profile of a new appliance once it is loaded into the power utility. The data is presented using a categorical histogram of feature samples with normalised weight properties to represent the data distribution measured at a specific time. The pattern of feature data distribution can be detected.

The whole process is required to extract the disaggregated power profile of the new appliance for training the AI model in the cloud. If the dataset is classified as a limited data dataset, a standard pre-trained GAN model available from the database is used for transfer learning with the uploaded limited data dataset. Then, GAN generates sufficient synthetic data, which is used to train the AI model for operating on NILM. Finally, the updated AI model is downloaded to the smart meter after training. The database is updated with the electrical feature of the new appliance. As such, this process also facilitates the power utility to build its database.

4.7 Numerical Simulation

The feature data on each appliance obtained in Section 4.5 is converted into a 2D image input to be tested on DenseNet and ResNet. The test data are presented using a categorical histogram of feature samples with normalised weight properties to represent the data distribution. DenseNet has fewer parameters than ResNet, making it less likely to fall into overfitting, gradient vanishing and model degradation [213]. Also, it reported that DenseNet performed better than ResNet50, Xception and VGG19 on biomedical image classification with a small dataset [226]. DenseNet is used to analyse the 2D graph for multi-class classification based on its strong feature analysis capacity on different image patterns.

The simulation results of the limited data datasets are illustrated in Table 4- 4, Figures 4- 9 and 4- 10. Dataset A is the real data. The best result is Val loss: 0.2275 and Acc: 1.0000. Dataset B is the mix of real and three hundred synthetic data. The best result is Val loss: 0.0192 and Acc: 1.000. The simulation results of the one-state datasets are illustrated in Table 4- 5, Figures 4- 11 and 4- 12. The best result of dataset A is Val loss: 0.0532 and Acc: 0.9810. The best result of dataset B is Val loss: 0.0328 and Acc: 0.9995. The simulation results of the multi-state datasets are illustrated in Table 4- 6, Figures 4- 13 and 4- 14 below. The best result of dataset A is Val Loss: 0.1851, Acc: 0.9143. The best result of dataset B is Val Loss: 0.1844, Acc: 0.9218.

The DCGAN has suffered the problem of mode collapse to generate low-quality data from the limited data datasets. This challenge has been resolved by applying the transfer learning technique to generate 3000 good-quality synthetic data in 1200 training epochs. The simulation of DCGAN by applying the transfer learning on a limited data dataset of KE, a one-state dataset of TV and a multi-state dataset of CT are shown in Figures 4- 15 to 4- 17. Furthermore, it demonstrates that the number of training iterations on GAN is significantly reduced, as in Tables 4- 7 to 4- 9.

Table 4- 4 Simulation results with limited data datasets obtained from real data and GAN

Data source (Real data)	Best Val loss: 0.2275, Acc: 1.0000.
Data source (Real data plus 300 synthetic data)	Best Val loss: 0.0192, Acc: 1.0000.

Table 4- 5 Simulation results with one-state datasets obtained from real data and GAN

Data source (Real data)	Best Val loss: 0.0532, Acc: 0.9810.
Data source (Real data plus 300 synthetic data)	Best Val loss: 0.0328, Acc: 0.9995.

Table 4- 6 Simulation results with multi-state datasets obtained from real data and GAN

Data source (Real data)	Best Val loss: 0.1851 Acc: 0.9143.
Data source (Real data plus 300 synthetic data)	Best Val loss: 0.1844 Acc: 0.9218.

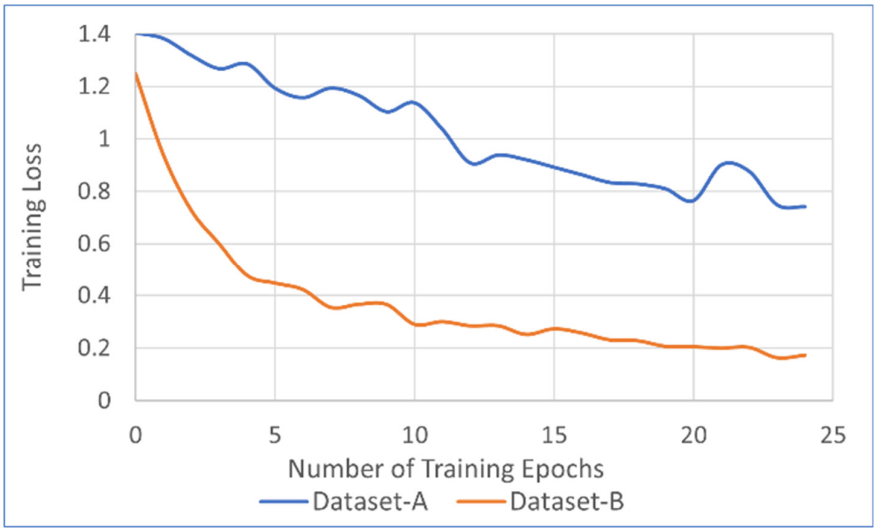

Fig. 4- 9 Simulation for the training loss with limited data datasets of microwave (ME), vacuum (VM), kettle (KE) and air purifier (DN) for real data on dataset A and real data plus 300 synthetic data on dataset B

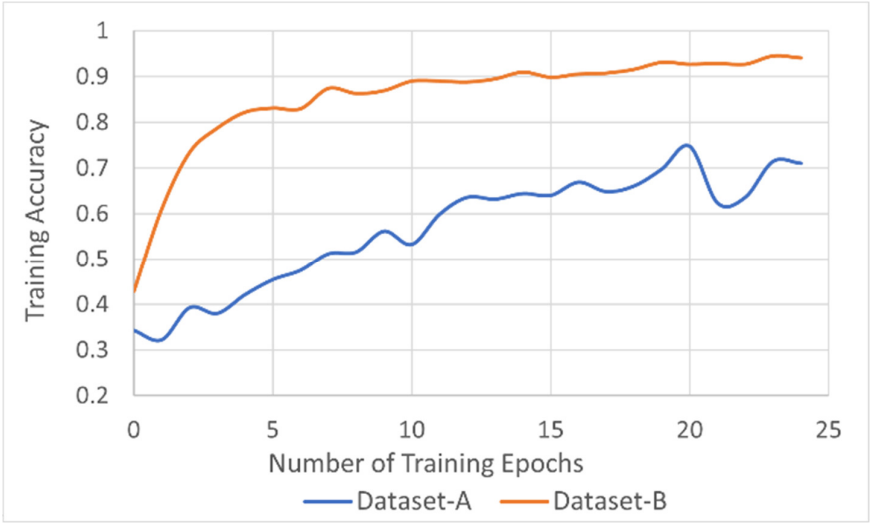

Fig. 4- 10 Simulation for the training accuracy with limited data datasets for real data on dataset A and real data plus 300 synthetic data on dataset B

In reference to Figures 4- 11 to 4- 14, the training loss and accuracy curves indicate that dataset B performs better than dataset A. Dataset A is set up from real data on both one-state and multi-state datasets. Dataset B is set up from the mix of real and three hundred synthetic data. The simulation results in Figures 4- 9 to 4- 10 further verify that the training

loss and accuracy are significantly improved by adding the synthetic data into real data on the limited data datasets to find the best weights during DenseNet training. The best values of validation loss and accuracy of dataset B also outperform dataset A for all training.

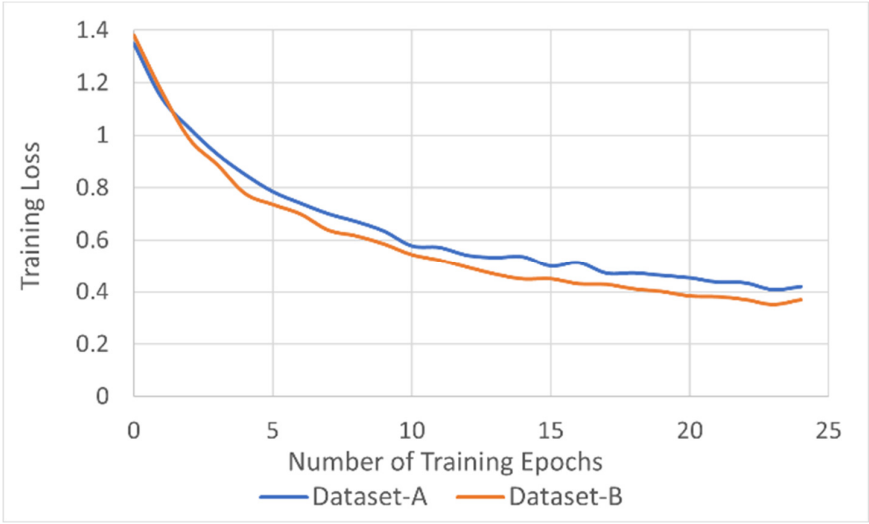

Fig. 4- 11 Simulation for the training loss with one-state datasets of the television (TV), hi-fi audio amplifier (AP), desktop computer (DC), oven (ON), dryer (DR), heater (HR) and freezer (FR) for real data on dataset A and real data plus 300 synthetic data on dataset B

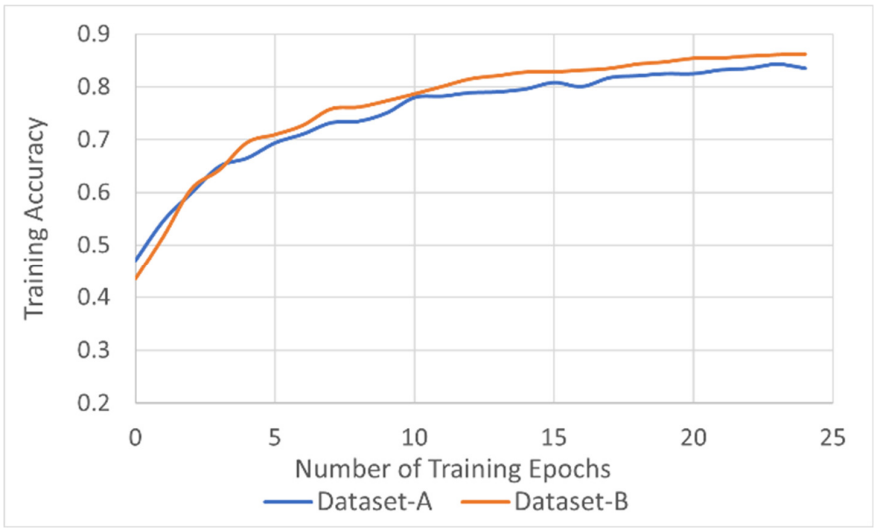

Fig. 4- 12 Simulation for the training accuracy with one-state datasets for real data on dataset A and real data plus 300 synthetic data on dataset B

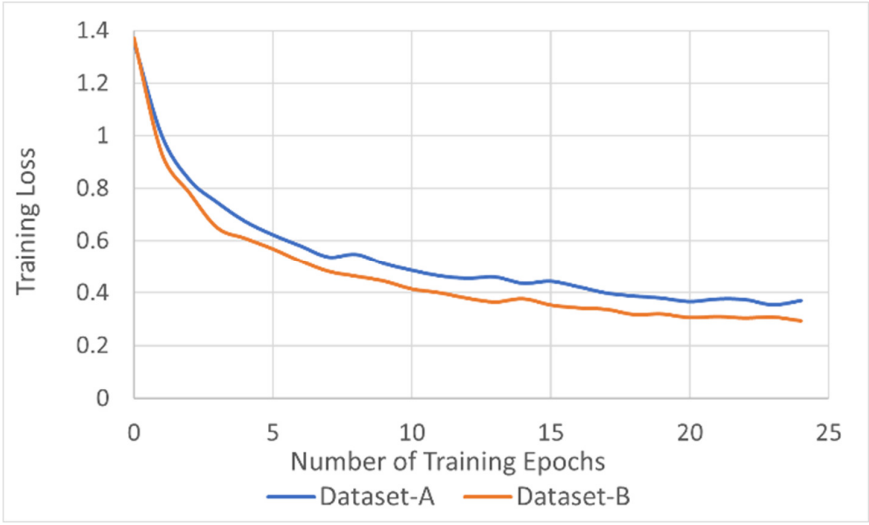

Fig. 4- 13 Simulation for the training loss with multi-state datasets of the air conditioner (AC), cooktop (CT) and refrigerator (RR) for real data on dataset A and real data plus 300 synthetic data on dataset B

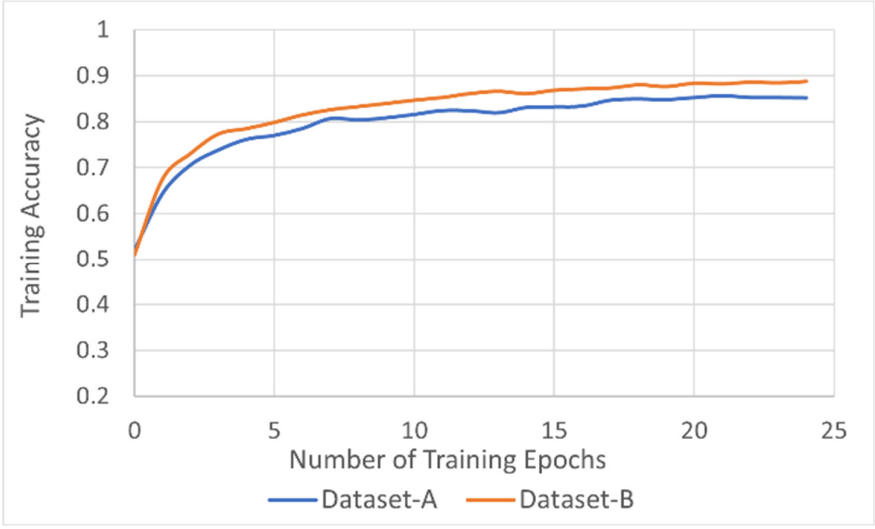

Fig. 4- 14 Simulation for the training accuracy with multi-state datasets for real data on dataset A and real data plus 300 synthetic data on dataset B

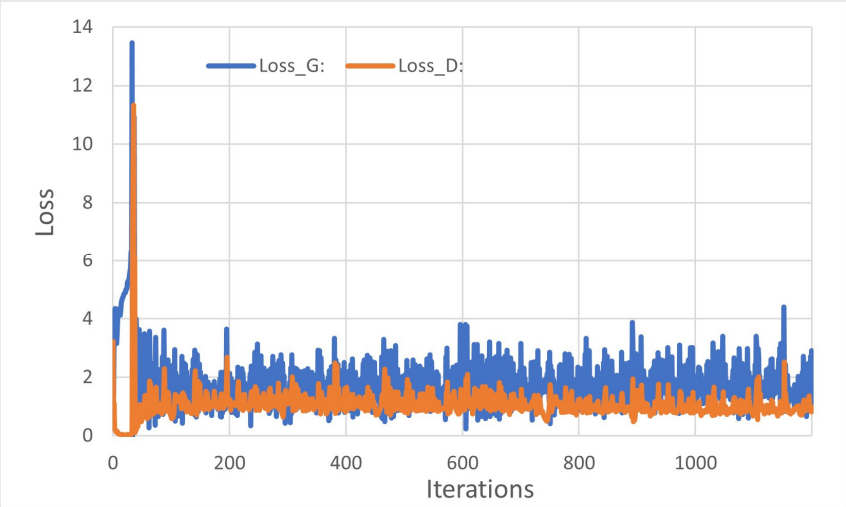

Fig. 4- 15 GAN simulation for transfer learning with limited data dataset of the kettle (KE) for 3000 synthetic data generation in 1200 training epochs

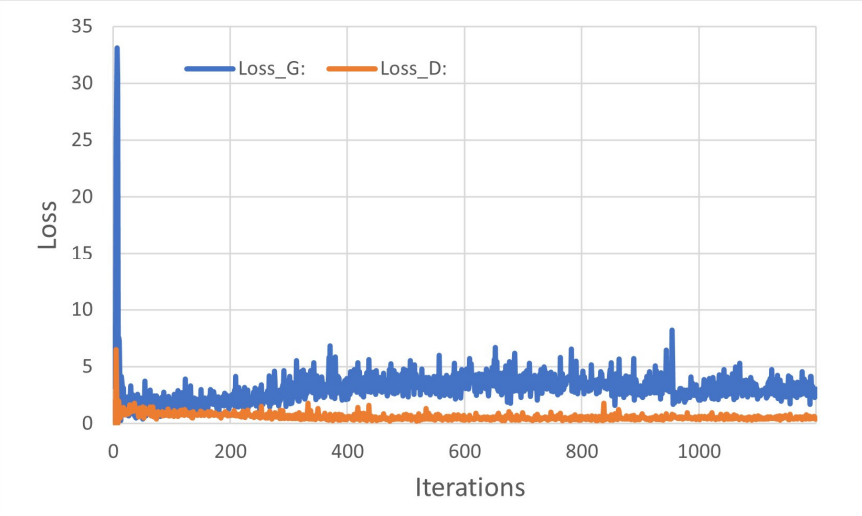

Fig. 4- 16 GAN simulation for transfer learning with a one-state dataset of the television (TV) for 3000 synthetic data generation in 1200 training epochs

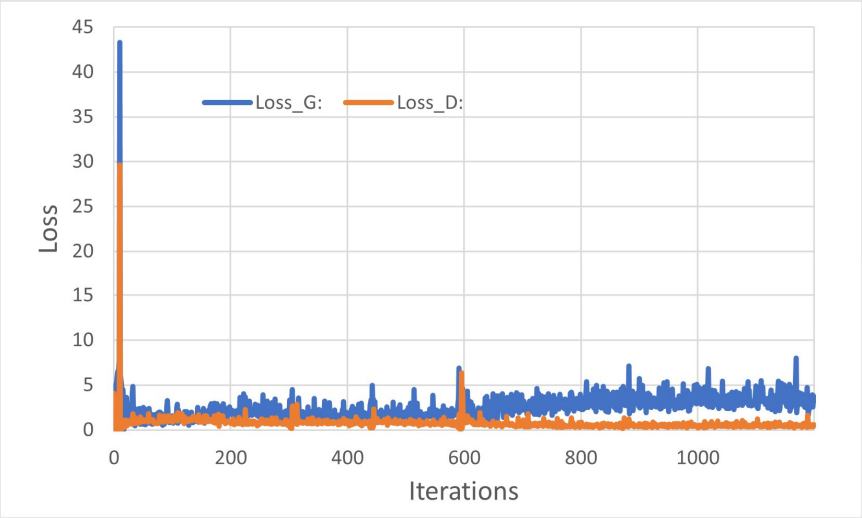

Fig. 4- 17 GAN simulation for transfer learning with a multi-state dataset of the cooktop (CT) for 3000 synthetic data generation in 1200 training epochs

Table 4- 7 GAN simulation for 3000 synthetic data generation from scratch by limited data datasets

Appliances	ME	VM	KE	DN
Number of Training Epochs	16000	12400	22875	21700

Table 4- 8 GAN simulation for 3000 synthetic data generation from scratch by one-state datasets

Appliances	TV	AP	DC	
Number of Training Epochs	32250	16250	28300	
Appliances	ON	DR	HR	FR
Number of Training Epochs	58250	7860	36500	28250

Table 4- 9 GAN simulation for 3000 synthetic data generation from scratch by multi-state datasets

Appliances	AC	CT	RR
Number of Training Epochs	38500	16500	15875

A ResNet [213] is used to further check the proposed approach to classification performance in four test scenarios, i.e. scenario A, scenario B, scenario C and scenario D, to unseen data points described in Section 4.7. The dataset size of appliances listed in Tables 4-

1 to 4- 3 is reduced to 20 real data samples of the training dataset with 40 synthetic data samples each. The test dataset size is selected to be 15 samples. Referring to Tables 4- 10 to 4- 21, the evaluation results of classification metrics, i.e. accuracy, precision, recall and F1-score, indicate that the synthetic data effectively improves the performance when the training data is insufficient.

SCENARIO A: Simulation of the ResNet with test data of air conditioner (AC), hi-fi audio amplifier (AP), cooktop (CT), air purifier (DN), desktop computer (DC) and refrigerator (RR)

Table 4- 10 Classification metric – real limited datasets – Scenario A

	accuracy	precision	recall	F1-Score
AC	0.8333	0.5000	1.0000	0.6667
AP	0.9889	0.9375	1.0000	0.9677
CT	1.0000	1.0000	1.0000	1.0000
DN	0.8333	0	0	0
DC	1.0000	1.0000	1.0000	1.0000
RR	0.9889	1.0000	0.9333	0.9655

Table 4- 11 Classification metric – synthetic datasets – Scenario A

	accuracy	precision	recall	F1-Score
AC	0.8222	0.4839	1.0000	0.6522
AP	0.8333	0.5000	1.0000	0.6667
CT	0.8333	0	0	0
DN	0.8333	0	0	0
DC	0.8333	0	0	0
RR	0.8444	0.5172	1.0000	0.6818

Table 4- 12 Classification metric – real + synthetic datasets – Scenario A

	accuracy	precision	recall	F1-Score
AC	1.0000	1.0000	1.0000	1.0000
AP	1.0000	1.0000	1.0000	1.0000
CT	1.0000	1.0000	1.0000	1.0000
DN	1.0000	1.0000	1.0000	1.0000
DC	1.0000	1.0000	1.0000	1.0000
RR	1.0000	1.0000	1.0000	1.0000

SCENARIO B: Simulation of the ResNet with test data of air conditioner (AC), cooktop (CT), dryer (DR), freezer (FR), kettle (KE) and refrigerator (RR)

Table 4- 13 Classification metric – real limited datasets – Scenario B

	accuracy	precision	recall	F1-Score
AC	0.9889	0.9375	1.0000	0.9677
CT	1.0000	1.0000	1.0000	1.0000
DR	0.9444	1.0000	0.6667	0.8000
FR	1.0000	1.0000	1.0000	1.0000
KE	0.9556	0.7895	1.0000	0.8824
RR	1.0000	1.0000	1.0000	1.0000

Table 4- 14 Classification metric – synthetic datasets – Scenario B

	accuracy	precision	recall	F1-Score
AC	0.9111	0.8889	0.5333	0.6667
CT	0.8333	0	0	0
DR	0.6000	0.2941	1.0000	0.4545
FR	0.9444	1.0000	0.6667	0.8000
KE	0.8333	0	0	0
RR	0.9444	0.7500	1.0000	0.8571

Table 4- 15 Classification metric – real + synthetic datasets – Scenario B

	accuracy	precision	recall	F1-Score
AC	1.0000	1.0000	1.0000	1.0000
CT	1.0000	1.0000	1.0000	1.0000
DR	0.9889	1.0000	0.9333	0.9655
FR	1.0000	1.0000	1.0000	1.0000
KE	0.9889	0.9375	1.0000	0.9677
RR	1.0000	1.0000	1.0000	1.0000

SCENARIO C: Simulation of the ResNet with test data of air conditioner (AC), cooktop (CT), microwave (ME), heater (HR), refrigerator (RR) and television (TV)

Table 4- 16 Classification metric – real limited datasets – Scenario C

	accuracy	precision	recall	F1-Score
AC	0.9889	1.0000	0.9333	0.9655
CT	0.9889	0.9375	1.0000	0.9677
ME	1.0000	1.0000	1.0000	1.0000
HR	1.0000	1.0000	1.0000	1.0000
RR	1.0000	1.0000	1.0000	1.0000
TV	1.0000	1.0000	1.0000	1.0000

Table 4- 17 Classification metric – synthetic datasets – Scenario C

	accuracy	precision	recall	F1-Score
AC	0.8444	1.0000	0.0667	0.125
CT	0.8333	0	0	0
ME	0.8333	0	0	0
HR	1.0000	1.0000	1.0000	1.0000
RR	0.8333	0	0	0
TV	0.3444	0.2027	1.0000	0.3371

Table 4- 18 Classification metric – real + synthetic datasets – Scenario C

	accuracy	precision	recall	F1-Score
AC	1.0000	1.0000	1.0000	1.0000
CT	1.0000	1.0000	1.0000	1.0000
ME	1.0000	1.0000	1.0000	1.0000
HR	1.0000	1.0000	1.0000	1.0000
RR	1.0000	1.0000	1.0000	1.0000
TV	1.0000	1.0000	1.0000	1.0000

SCENARIO D: Simulation of the ResNet with test data of air conditioner (AC), cooktop (CT), dryer (DR), oven (ON), refrigerator (RR) and vacuum (VM)

Table 4- 19 Classification metric – real limited datasets – Scenario D

	accuracy	precision	recall	F1-Score
AC	1.0000	1.0000	1.0000	1.0000
CT	0.9667	1.0000	0.8000	0.8889
DR	0.9778	0.8824	1.0000	0.9375
ON	0.9111	1.0000	0.4667	0.6364
RR	1.0000	1.0000	1.0000	1.0000
VM	0.9	0.625	1.0000	0.7692

Table 4- 20 Classification metric – synthetic datasets – Scenario D

	accuracy	precision	recall	F1-Score
AC	0.8556	1.0000	0.1333	0.2353
CT	0.7222	0.3611	0.8667	0.5098
DR	0.8333	0	0	0
ON	0.6000	0.2162	0.5333	0.3077
RR	1.0000	1.0000	1.0000	1.0000
VM	0.8333	0	0	0

Table 4- 21 Classification metric – real + synthetic datasets – Scenario D

	accuracy	precision	recall	F1-Score
AC	1.0000	1.0000	1.0000	1.0000
CT	0.9889	0.9375	1.0000	0.9677
DR	0.9778	0.8824	1.0000	0.9375
ON	0.9667	1.0000	0.8	0.8889
RR	1.0000	1.0000	1.0000	1.0000
VM	1.0000	1.0000	1.0000	1.0000

4.8 Conclusion

The simulation results indicate that GAN can generate synthetic data to tackle the challenge of limited data on NILM effectively. Also, the transfer learning technique can be applied to resolve the mode collapse problem when GAN generates synthetic data. Furthermore, it helps to stabilise the training on GAN in a shorter number of training iterations.

As more new types of electrical devices will be connected to the electricity network, the proposed methodology works for domestic houses, and the procedure is also applicable to the application of NILM in the industrial sector.

5 Transfer Learning Based Dynamic Security Assessment

5.1 Introduction

In recent years, major power blackouts have occurred in various countries. Together with the increasing integration of renewable energy in the power supply system and the active response of the demand side, these uncertainties bring an unprecedented level of complexity to power system stability. Traditional DSA technology has not been able to provide an effective and timely system security assessment to meet this challenge.

The conventional DSA analysis methods are based on time domain simulation (T-D) techniques and energy functions. However, it is not efficient to meet the high-speed online assessment if a large-sized model parameter and contingency set is encountered. The corresponding remedial control actions will no longer be reliable to satisfy the dynamic system condition.

With the wide application of technologies such as WAMS and PMU in power system, fast online DSA has been adopted as an effective means of detecting the unstable status of power system stability by continuous inputting the features data of operating parameters of the power system in real-time [227-229].

The DSA is applicable to the assessment of the power system security status for both pre-fault and post-fault stages. It monitors if the system will continue to be stable with respect to an anticipated but not yet occurred fault conditions through evaluating the pre-fault steady-state of the power system features, and also predict the system security status under post-fault ongoing disturbance. As such, the appropriate remedial control action is activated to prevent and avoid the power system operating at insecure regions when it encounters an unexpected fault condition or an ongoing disturbance.

In a smart grid, the data-driven method based on artificial intelligence (AI) architecture and data mining technology can effectively achieve rapid evaluation for online DSA [230]. It is a mapping relationship to correlate the input of power system features and the corresponding secure/insecure operating regions of the power system. This method can rapidly assess potential faults as a secure category or an uninsured category.

In [231, 232], decision tree (DT) and random forest (RF) were used for online DSA. In [233], support vector machine (SVM) was developed for rotor angle stability analysis. In [234, 235], artificial neural network was proposed to perform DSA. In [236], SVM and different neural network (NN) including probabilistic, multilayer perceptron (MLP), recurrent with a long short term memory (RNN-LSTM), extreme learning machine (ELM) and stacked auto-encoders (SAE) and convolutional neural network (CNN) were tested in power systems with very high accuracy. The author could not conclude which method was more suitable for transient stability assessment. In [237], a binary classification CNN model was generated for transient stability prediction. And a multi-class classification CNN model was built for unstable generator identification. The author demonstrated that the proposed CNN models provided more accurate results than DT, MLP, RF and *k*-nearest neighbor (KNN) for transient stability prediction and unstable generator identification. In [238], transfer learning technique was employed to minimise the marginal and conditional distribution differences between the trained data and unknown data in power system for utilising one trained model to assess unlearned faults.

The AI is capable of predicting the security status of the power system in real-time after being trained by a dynamic security database of the power system. However, if the training data set is too small, it results in a poor approximation. An over-constrained model will underfit the small training dataset. If the training data set is too large, it will slow down the

rapid update of the online AI model to identify the potential risk of the power outage. The timing of an insecure event identification is crucial to the effectiveness of emergency remedial control action. The earlier the unstable status of the power system is detected, the greater the likelihood that the power system will return to a stable condition. Transfer learning technology can tackle this challenge to improve the validation accuracy by finetuning technique if the training dataset is too small and rapidly train the AI model to shorten its online update time by feature extraction technique if the training dataset is too large.

Using ImageNet pre-trained CNN features, impressive results were obtained on several image classification datasets, as well as object detection, action recognition, human pose estimation, image segmentation, optical flow, image captioning, and others. In [47], the author demonstrated that training of finetuning on ImageNet classes that were not present in PASCALDET was sufficient to learn features that were also good for PASCAL classes. This concept is employed in this paper, that is, the pretrained non-electrical features from ImageNet on an AI model are transferred to improve the performance on the same model which is based on the training dataset obtained from the simulation of the power system. As such, the feature data obtained from the simulation of stability studies after being selected by KPCA or ReliefF is converted into the two-dimensional (2D) image input to train the scratch AI model. It is in the form of a two-dimensional graph whose two coordinates are feature samples and normalised weight properties of the feature sample to represent the distribution of data in a categorical histogram. This simpler approach has the dimension lower than the feature dataset selected by KPCA and ReliefF.

In [226], the author reported that DenseNet outperformed the VGG19, ResNet50 and Xception on a biomedical image binary classification in an application of transfer learning

method with a small dataset. In this paper, a binary classification DenseNet model is used to demonstrate the effectiveness of transfer learning approach even though the DenseNet can be used for multiclass classification. As DenseNet has a powerful capacity for feature analysis on different image patterns, this advantage is applied to analyse the 2D histograms of simulated OPs as different image inputs.

The pre-fault input data used in the work is converted to 2D histogram image data that is different but related to the dataset on ImageNet. The inductive transfer learning is appliable to this scenario. Then, the performance of the transfer learning approach with ImageNet is compared to scratch DenseNet trained by the operation points (OPs) obtained from the Western System Coordinating Council (WECC) 3-machine, 9 bus system [239-241]. Other smaller datasets from three scenarios on New England 39 bus system [242] are tested for further validation.

In addition, the transient stability status of a power system following a fault can be predicted early based on the measured post-fault values of the feature data obtained from the simulation of stability studies [233]. The feature data collected immediately after clearing a fault is used as inputs to an SVM classifier, predicting the transient stability status. As these feature data are used as training data for SVM, the proposed approach from this paper can convert the training data to be 2D data. Then, transfer learning techniques can be applied.

The current challenges on DSA are as follows:

- The increasing integration of renewable energy sources into the power supply system, coupled with the active response of the demand side, introduces a heightened level of complexity to power system stability due to the associated

uncertainties. Traditional DSA technology has not been able to provide an effective and timely system security assessment to meet this challenge.

- The AI is capable of predicting the real-time security status of the power system after being trained on a dynamic security database. The timing of identifying an insecure event is crucial for effective emergency remedial control action. Detecting the unstable status of the power system at an earlier stage significantly increases the likelihood of restoring it to a stable condition.
- If the training data set is too large, it will slow down the rapid update of the online AI model to identify the potential risk of the power outage.
- If the training data set is too small, it results in a poor approximation. An over constrained model will underfit the small training dataset.

Transfer learning technology can address the challenges mentioned above in AI model training. It achieves this by using the finetuning technique, which improves validation accuracy when the training dataset is small. Additionally, it utilises the feature extraction technique to rapidly train the AI model when the training dataset is too large.

5.2 Common Learning Algorithms

In deep learning, CNN as one of the major computational applications for image recognition and image classification is widely used in different fields [243-247]. As CNNs draw representational power from extremely deep or wide architectures, it emerges the new problems of gradient vanishing and model degradation. The recent network frameworks of ResNet [248], Highway Networks [249] and Fractal Nets [250] vary in the network topology to tackle these drawbacks; they all contain the same core ideas that they create short paths from early layers to later layers.

DenseNet [223] has deviated from the above concepts that the network performance is improved through feature reuse and bypass settings of information flow. It connects each layer to every other layer in a feed-forward fashion that some high-level features may be produced in the deeper final layers of the network. This approach makes it fewer parameters than ResNet for easier training and alleviates gradient vanishing and model degradation.

Assume a photo of X_0 to be input on a traditional convolutional feed-forward network which comprises L layers, ι indexes the layer for implementing a non-linear transformation $H_\iota(*)$. $H_\iota(*)$ can be a composite function of operations such as BN, ReLU, Pooling or Conv. The feature output at the ι^{th} layer is denoted as X_ι. Refer to Eq. 4-4 in Section 4.4:

$$X_\iota = H_\iota([X_0, X_1, \ldots, X_{\iota-1}]) \tag{4-4}$$

where [] refers to the concatenation of the feature-maps produced from layer 0 to layer ι-1. The input of the ιth layer is not only related to the output of the ι-1th layer but also accessible to the output of all preceding layers. The feature size must be the same in the same Denseblock for performing concatenation of feature maps on the different layers. Transition layers used for indirect information flow are set between different Denseblocks to achieve down-sampling. A single classifier on top of the network supervises all layers through transition layers. It makes the model structure and gradient calculation simpler than deeply-supervised nets (DSN) which have classifiers attached to every hidden layer.

If the feature maps are viewed as the global state of the network, the training goal of each layer is to determine the updated value that needs to be added to the collective knowledge of its Denseblock. Assuming that the output of each non-linear transformation H is κ feature maps, the ι^{th} layer has $\kappa_0 + (\iota - 1) \times \kappa$ input feature maps. The growth rate κ also determines how much information each layer updates to the global state. DenseNet can have very narrow layers that a relatively small growth rate is sufficient to achieve state-of-art

performance. This characteristic makes it significantly different from the other neural network.

The DenseNet-B uses a 1 × 1 Conv (Bottleneck) before each 3 × 3 Conv to reduce the computational burden that the improved non-linear transformation of input feature maps becomes BN-ReLU-Conv(1 × 1)-BN-ReLU-Conv(3 × 3). If a Denseblock contains m feature maps, the transition layer generates θm output feature maps. It is called DenseNet-C if the compression factor $\theta < 1$. If the Bottleneck is also used, it is called DenseNet-BC. DenseNet-BC is the most parameter efficient variant of DenseNet. It requires only about one-third of the number of parameters of ResNet with the same model accuracy. Comparing ResNet with 1001 layer over 10 M parameter and DenseNet-BC with the only 0.8 M parameter for 100 layers in training, they converge on the same training epoch.

The learning algorithm has several compelling advantages [223], and the DenseNet fits the challenges of both DSA and load modelling problems very well.

- It substantially reduces the number of parameters, enhance feature reuse, strengthen feature propagation, and alleviate the vanishing-gradient problem.
- It has better parameter efficiency by improving the flow of information and gradients throughout the network, resulting in easy training.
- It allows layers to access feature-maps from all of their preceding layers, facilitating feature reuse.
- It achieves significant improvements over the state-of-the-art on most tasks while requiring fewer computational resources to achieve high performance.
- It successfully tackles overfitting by reducing the number of parameters and enabling feature reuse to manage redundancy, enhancing the model's capacity to generalise to unknown data.

5.3 A Transfer Learning Based Transient Stability Assessment Framework

Transfer learning is not a specific new concept for deep learning. Instead, it is entirely different from the traditional approach of machine learning to train a model. Users can re-use knowledge (i.e. features, weights, etc.) from another pretrained model for a related task in a new model. Reference [211] gave the definitions of "Transfer Learning" as follows.

Given a source domain D_S and learning task $\mathcal{T}_S$, a target domain D_T and learning task $\mathcal{T}_T$, transfer learning aims to improve the learning of the target predictive function $\mathfrak{f}_T(\cdot)$ in D_T using the knowledge in D_S and $\mathcal{T}_S$, where $D_S \neq D_T$ or $\mathcal{T}_S \neq \mathcal{T}_T$. If the target task $\mathcal{T}_T$ is different from the source task $\mathcal{T}_S$, it is the definition of inductive transfer learning no matter when the source and target domains are the same or not.

A domain D consists of two components: a feature space $\mathcal{X}$ and a marginal probability distribution $P(\mathrm{X})$, where $\mathrm{X} = \{x_1, \ldots, x_n\} \in \mathcal{X}$, $D = \{\mathcal{X}, P(\mathrm{X})\}$.

A task consists of two components: a label space $\mathcal{Y}$ and an objective predictive function $\mathfrak{f}(\cdot)$ (denoted by $\mathcal{T} = \{\mathcal{Y}, \mathfrak{f}(\cdot)\}$), which is not observed but can be learned from the training data of pairs $\{x_i, y_i\}$, where $x_i \in \mathrm{X}$ and $y_i \in \mathcal{Y}$. The function $\mathfrak{f}(\cdot)$ can be used to predict the corresponding label $\mathfrak{f}(x)$ of a new instance x. From a probabilistic viewpoint, $\mathfrak{f}(x)$ can be written as $P(y|x)$, and $\mathcal{T} = \{\mathcal{Y}, P(\mathrm{Y}|\mathrm{X})\}$.

When the learning tasks $\mathcal{T}_S$ and $\mathcal{T}_T$ are different, then either (1) the label spaces between the domains are different, i.e. $\mathcal{Y}_S \neq \mathcal{Y}_T$, or (2) the conditional probability distributions between the domains are different, i.e. $\mathrm{P}(\mathrm{Y}_S|\mathrm{X}_S) \neq \mathrm{P}(\mathrm{Y}_T|\mathrm{X}_T)$, where $\mathrm{Y}_{S_i} \in \mathcal{Y}_S$ and $\mathrm{Y}_{T_i} \in \mathcal{Y}_T$.

When the domains are different. Then either (1) the feature spaces between the domains are different, i.e. $\mathcal{X}_S \neq \mathcal{X}_T$ or (2) the feature spaces between the domains are the same, but the

marginal probability distributions between domain data are different. i.e. $P(\mathrm{X}_S) \neq P(\mathrm{X}_T)$, where $\mathrm{X}_{S_i} \in \mathcal{X}_S$ and $\mathrm{X}_{T_i} \in \mathcal{X}_T$.

When the target and source domains are the same, i.e. $D_S = D_T$, and their learning tasks are the same. i.e. $\mathcal{T}_S = \mathcal{T}_T$, it becomes a traditional machine learning problem.

The inductive transfer learning can be considered as multi-task learning if there are ample labelled data available from the source domain for target labelled data to induce an objective predictive model in the target domain. It helps to achieve high performance in the target task by transferring knowledge from the source task.

The feature-representation approach to the inductive transfer learning problem aims at finding good feature representations to minimise domain classification error. If a lot of labelled data in the source domain are available, supervised learning methods can be used to construct a feature representation. In [251], a two-stage approach was presented to domain adaptation. A generalisable feature representation with appropriate weights across different domains is selected to train a general classifier. Those features specifically useful for the target domain is picked up by employing semi-supervised learning.

The feature-projection mapping approach is used to map the data of each domain from the high-dimensional feature space to the low-dimensional latent feature space. After projection mapping, the marginal distributions of data between the source domain and target domain are close to each other in the low dimensional latent feature space. In this way, the tagged source domain sample data can be used to train the classifier and predict the target test data. The training model of a supervised learning algorithm can be used in this case. In [252], a new dimension reduction method was proposed to solve the dimensionality reduction by minimising the Maximum Mean Discrepancy of the source domain data and the target

domain data in the latent space. In [253], a structural corresponding learning (SCL) algorithm was proposed to learn a projection mapping of pivot features from the feature spaces of the source domain and target domain to a shared, low-dimensional real-valued feature space through defining a set of pivot features on the unlabelled data from both domains.

The idea of supervised feature construction methods for the inductive transfer learning setting is similar to multi-task learning. A low-dimensional representation that is shared across related tasks is learned. The common features can be learned by solving an optimisation problem as follows:

$$\begin{aligned} &\arg\min_{A,U} \sum_{t\in\{T,S\}} \sum_{i=1}^{n_t} \mathrm{L}\left(y_{t_i}, \langle a_t, U^T x_{t_i}\rangle\right) + \gamma \|A\|_{2,1}^2 \\ &\text{s.t.} \quad \boldsymbol{U} \in \mathbf{O}^d \end{aligned} \tag{5-1}$$

In this equation, S and T denote the tasks in the source domain and target domain, respectively. $A = [a_S, a_T] \in R^{dx2}$ is a matrix of parameters. U is a d x d orthogonal matrix for mapping the original high-dimensional data to low-dimensional representations. The learned new representation can reduce the classification error of each task. The (r, p)-norm of A is defined as:

$$\|A\|_{r,p} \coloneqq \left(\sum_{i=1}^{d} \left\|a^i\right\|_r^p\right)^{\frac{1}{p}} \tag{5-2}$$

The above optimisation problem estimates the low-dimensional representations $U^T X_T$, $U^T X_S$ and the parameters, A, of the model at the same time. It can be further transformed into an equivalent convex optimisation formulation and be solved efficiently.

5.4 A Scratch DenseNet Model for Transient Stability Assessment

The later the training on the AI model is completed, the less the likelihood that the risk of blackouts can be detected earlier. In practice, it is common to pretrain a DenseNet on a very

large dataset and then use it either as an initialisation or a fixed feature extractor. It helps to tackle the challenge if a scratch deep neural network of stability prediction takes too long to complete the training from a large feature dataset of the power system. Also, if the training data is not sufficient, e.g. the power system is under the situation of topology changes, it results in poor performance on validation accuracy.

In this paper, the ImageNet is used to pretrain the proposed DenseNet. Then, the techniques of transfer learning are applied to improve the performance that it outperforms the scratch DenseNet model to be trained by the dataset obtained from the simulation of stability studies. The training dataset is described in Section 5.5. The two major transfer learning scenarios are given as follows:

a. DenseNet as a fixed feature extractor: The weights for all the DenseNet are frozen except that of the final fully connected layer. This last fully connected layer is replaced with a new one with random weights, and only this layer is trained. The feature codes for all images are extracted to train this layer as a classifier. The rest of the DenseNet is treated as a fixed feature extractor for the new dataset.
b. Finetuning the DenseNet: Instead of random initialisation, it finetunes the weights of the pretrained DenseNet using data from ImageNet by continuing the backpropagation. As the earlier layers of a DenseNet contain a more generic feature; it is an option on finetuning some higher-level portion of the network rather than all the layers. Rest of the training is running as usual.

As described in Section 5.3, if the feature dataset of the power system is used to pretrain the DenseNet before feature extraction of transfer learning is performed, it becomes a traditional machine learning approach. If the DenseNet has already been trained by a feature dataset, the proposed transfer learning approach provides flexibility to re-use this DenseNet

model as a pretrained deep neural network. Then, the weights of the DenseNet are updated from its last checkpoint through finetuning or feature extraction techniques to save for later use.

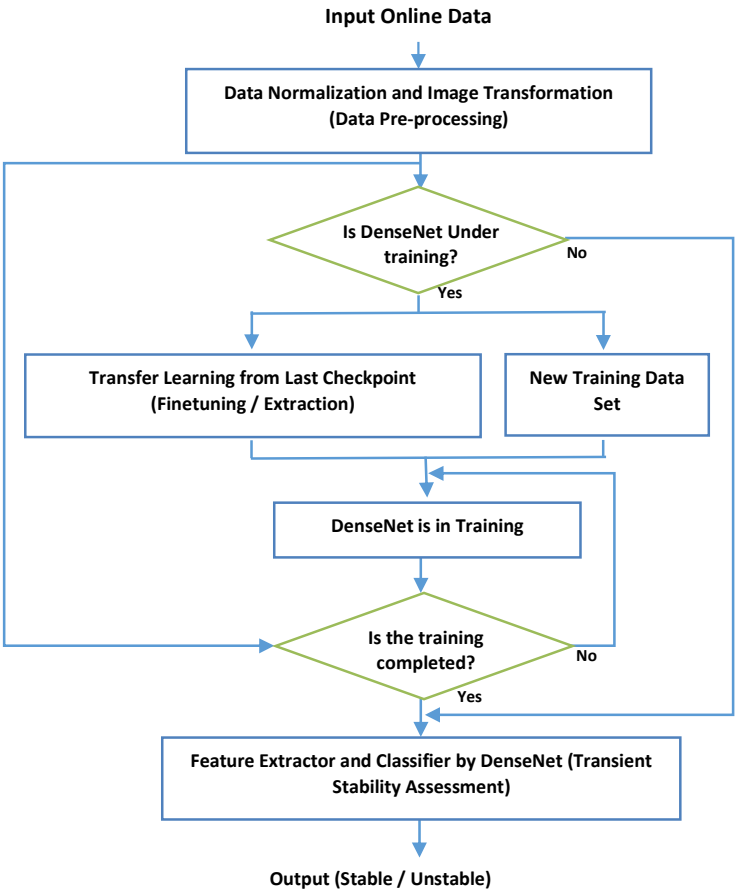

Fig. 5- 1 Flow chart of the proposed transfer learning-based transient stability assessment framework

Since the target dataset is similar in context to the original dataset of the pretrained DenseNet at the last checkpoint, the feature extraction technique can rapidly train the DenseNet with validation accuracy close to the pretrained model. If the new dataset is very small, training a DenseNet on a very small dataset from scratch may affect its ability to generalise, often cause overfitting concerns. The finetuning technique can tackle this problem to get a better result. The flow chart of transfer learning-based framework is shown in Figure 5- 1.

5.5 Database Generation

In [254], five indices were reviewed for the assessment of power system transient stability. They are the maximum rotor angle deviation, maximum speed deviation, maximum acceleration, TRASI and transient stability index (TSI). The indices of maximum rotor angle deviation, maximum speed deviation and maximum acceleration are specific to the individual generator that the running characteristics of each generator are feedbacked. But they do not provide the specific instability value of the entire system. The TRASI and TSI are the indices to measure the whole system stability status. TRASI measures the maximum rotor angle difference before and after a system fault to determine the severity. But the research for this subject has commonly used the angle-based stability index of TSI as described in Eq. 5-3 [254]. Also, its concept is not complicated to be implemented by Power System Toolbox (PST) software. Therefore, it is used as the transient stability index in this paper. According to the criterion of Eq. 5-4, secure and insecure labels are used to classify transient stability status.

The popular WECC 3-machine, 9 bus system and New England 39 bus system are considered in this paper. The simple 9 bus system well demonstrates how the proposed DenseNet and transfer learning techniques are tested for transient stability classification. The generator at bus #2 is assigned as a wind farm, and the power system network and parameters are obtained from [255]. Then, the effectiveness of the proposed approach is further validated by New England 39 bus system.

Power angle-based stability index is defined as follows:

$$\eta = \frac{360-\delta_{max}}{360+\delta_{max}}\ x\ 100 \qquad -100 < \eta < 100 \tag{5-3}$$

$$If \begin{cases} \eta > 0 \ \rightarrow \ \text{stable (secure label)} \\ \eta < 0 \ \rightarrow \ unstable\ (insecure\ lebal) \end{cases} \tag{5-4}$$

where δ_{max} is the maximum angle separation of any two generators at the same time in the post-fault response. The $\eta > 0$ and $\eta < 0$ correspond to stable and unstable conditions [61].

Under the above transient stability criterion, 4,900 OPs are sampled evenly distributed across the variation of the generator output power at bus #2 and load demand that 3,500 pre-fault secure OPs and 1,400 pre-fault insecure OPs of the 9 bus test system are produced, respectively. It covers various load patterns ranging from 60% to 180% and the electricity output capacity of the wind farm ranging from 0% to 100% of its base OP. A three-phase fault is simulated at bus #4 by PST software, and the fault is cleared after 0.05 s.

Other three smaller datasets of 500 OPs are generated from New England 39 bus system by DIgSILENT PowerFactory software as shown in Table 5- 1. Monte Carlo technique is applied to generate the random load ratio within the pre-determined range between 0.52 to 1.2.

Table 5- 1 records the protection measures on three scenarios that the three-phase fault occurs on buses at 0.5 s to obtain different stability dataset. In the scenario of bus #3, the fault is clear at 0.75 s. It generates 207 pre-fault secure OPs and 293 pre-fault insecure OPs. No transmission line is tripped. On scenarios of bus #19 and #22, the faults are clear at 0.7 and 0.6 s. The transmission lines between bus #19 - #33 and #21 - #22 are tripped respectively on each scenario. The scenario of bus #19 generates 280 pre-fault secure OPs and 220 pre-fault insecure OPs. The scenario of bus #22 generates 217 pre-fault secure OPs and 283 pre-fault insecure OPs. The validation data is selected from 25% of the above datasets.

Table 5- 1 Short circuit scenarios for three phase faults on New England 39 bus power system

Fault location	bus #3	bus #19	bus #22
Fault duration (s)	0.25	0.2	0.1
Tripped line	N/A	19-33	21-22
Secure OPs	207	280	217
Insecure OPs	293	220	283

5.5.1 Dimensionality Reduction

The advantage of using the feature selection technique has recently attracted much attention within the deep learning studies that it helps to reduce training time, overfitting and complexity of the feature datasets, and also improve data understandability and validation accuracy. The cost of data acquisition may be significantly reduced if the right subset is chosen.

The power system has a high dimension of features in nature. Only those features which can characterise the system stability status are served as useful inputs. ReliefF algorithm [256, 257], which is well known applied to genetic analyses where epistasis is common is used in this paper to identify feature interactions without having to check every pairwise interaction exhaustively. It significantly reduces the initial dimensionality of the feature datasets with less calculation time than exhaustive pairwise search. The features are assigned with different weights according to the relevance of each feature and category. Also, KPCA for dimensionality reduction in non-linear space is used to compare the effectiveness of the ReliefF algorithm through the validation accuracy of the DenseNet model for transient stability prediction.

Dimensionality reduction has been studied widely for image processing in the deep learning community. Traditional dimensionality reduction methods try to project the original data to a 2D coordinate while preserving the critical properties of the original data for a CNN model. In [258], the maximum variance unfolding (MVU) technique was proposed to maximise the variance of the embedding while preserving the local distances between neighbouring observations for extracting a low-dimensional representation of the data. The MVU applies principal component analysis (PCA) [259, 260] to kernel matrix K to choose

the base eigenvectors for mapping data on KPCA. The concepts of PCA and KPCA are summarised as follows.

5.5.2 PCA

It is a transformation to diagonalise the covariance matrix of the feature data $x_j, j = 1, \dots, m$ for $x_j \in R^d$ that the data is centred $\sum_{j=1}^{m} x_j = 0$ to be defined as:

$$C = \frac{1}{m} \sum_{j=1}^{m} x_j \, x_j^T \tag{5-5}$$

The above problem can be solved by the following eigenvalue equation [261]:

$$\lambda v = Cv \tag{5-6}$$

where $\|v\| = 1, \lambda \geq 0$ are eigenvalues and $v \in R^d$ are eigenvectors.

The orthogonal projection on the eigenvector v^k of the principle component for a test point x can be found as:

$$x^k = v^k \cdot x \tag{5-7}$$

$$x_j' = \sum_{k=1}^{n'} a_{k,i} \, v^k \tag{5-8}$$

where $n' \leq m$ and x_j' are vectors after PCA transforms in n' lower-dimensional space.

Refer to [262, 263]. The scatter matrix S is defined as a statistical expectation operator:

$$S = E\left(x_j , x_j^T\right) \tag{5-9}$$

where $j = 1, \dots, m$ for $x_j \in R^d$, combine Eq. 5-8 and 5-9 that $a_{k,i}$ are the projections of the x_j on v^k.

$$\sigma^2(v^k) = E[a_{k,i}{}^2] = \left(v^k\right)^T S v^k \tag{5-10}$$

Given by [263], the variance σ^2 is a function of eigenvector v^k. Eq. 5-10 can be represented as an eigenvalue problem with nontrivial solutions at local maximum or minimum as shown in Eq. 5-11:

$$Sv^k = \lambda v^k \tag{5-11}$$

$$E_{n'} = \frac{1}{2}\sum_{j=n'+1}^{m} \lambda_k \tag{5-12}$$

where $\lambda_1 > \lambda_2 > \cdots > \lambda_k > \cdots > \lambda_m$. $E_{n'}$ is the error in the representation of original data x_j compared to data x_j'after dimensionality reduction by PCA, and λ_k is the eigenvalue corresponding to v^k. As the eigenvectors $v_1, v_2, \ldots, v_m$ are corresponding to $\lambda_1 > \lambda_2 > \cdots > \lambda_k > \cdots > \lambda_m$, the lowest order component of eigenvectors with the largest eigenvalues gives the smallest error in representation. It implies the variance to be maximum in the direction of the eigenvectors.

5.5.3 KPCA

The dynamic security problem is in the form of a large number of non-linear differential-algebraic equations. PCA is appliable for linear correlation among the variables that its performance is degraded in a non-linear problem. Kernel PCA is more suitable to transform data which are confined to low dimensional non-linear subspace in power system [34, 264]. It is a non-linear extension of PCA [265] to be described as follows.

First step: The $x_i \in R^d$ which is non-linearly separable is mapped to the feature space $\mathcal{F}$, dimensionally higher $(> d), \Phi : R^d \rightarrow \mathcal{F}$, so that the new non-linear subspace is linearly separable to perform PCA in $\mathcal{F}$ for a specific choice of Φ [266]. As such, $x_i \in R^d$ are mapped

through $\Phi(\mathrm{x_i}) \in \mathbb{F}$. The Mapping $\Phi: R^d \rightarrow \mathbb{F}$ is defined as dot product $K(x, y) = \Phi(x) \cdot \Phi(\mathrm{y})$, which is similar to Eq. 5-7. The dot product of Gaussian kernel is:

$$K_{\mathrm{ij}} = K(\mathrm{x_i}, \mathrm{x_j}) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\gamma^2}\right) \gamma > 0, \gamma: bandwidth \tag{5-13}$$

Second step: Suppose the feature space F which is defined by the vectors $\{\Phi(\mathrm{x_1}), \Phi(\mathrm{x_2}), \ldots, \Phi(\mathrm{x_m})\}$ to be centralised, i.e. $\sum_{j=1}^{m} \Phi(\mathrm{x_j}) = 0$, covariance for these vectors is shown as:

$$\hat{c} = \frac{1}{m} \sum_{j=1}^{m} \Phi(\mathrm{x_j})\, \Phi(\mathrm{x_j})^{\mathrm{T}} \tag{5-14}$$

The eigenvalues $\lambda \geq 0$ and nonzero eigenvectors $v \in \mathbb{F}$ that satisfy $\lambda v = \hat{c} v$ is required to do PCA for Eq. 5-14:

$$\lambda(\Phi(\mathrm{x_k}) \cdot v) = (\Phi(\mathrm{x_k}) \cdot \hat{c} v), \forall k = 1, 2, \ldots, m. \tag{5-15}$$

where all eigenvectors v in the feature space $\mathbb{F}$ can be expressed as a linear combination of $\Phi(\mathrm{x_k})$ in Eq. 5-15. There exist coefficients α_{i} where i = 1, …, m, in such, $v = \sum_{i=1}^{m} \alpha_{\mathrm{i}}\, \Phi(\mathrm{x_i})$. From Eq. 5-15:

$$\lambda \sum_{i=1}^{m} \alpha_{\mathrm{i}}\, \Phi(\mathrm{x_k}) \cdot \Phi(\mathrm{x_i}) =$$
$$\frac{1}{m} \sum_{i=1}^{m} \alpha_{\mathrm{i}}\, \{\Phi(\mathrm{x_k}) \cdot [\sum_{j=1}^{m} (\Phi(\mathrm{x_j}) \cdot \Phi(\mathrm{x_i}))\Phi(\mathrm{x_j})]\}$$
$$\forall k = 1, 2, \ldots, m. \tag{5-16}$$

The Eq. 5-16 is simplified by defining the m x m Gram matrix K for $K_{\mathrm{ij}} = \Phi(\mathrm{x_i}) \cdot \Phi(\mathrm{x_j})$. It is also known as Kernel Matrix. we obtain,

$$\lambda K \alpha = \frac{1}{m} K^2 \alpha \tag{5-17}$$

$$m\lambda\alpha = K\alpha \tag{5-18}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)^T$. α is not considered if $\lambda = 0$. PCA can be performed in the feature subspace $\mathbb{F}$ by solving the Eq. 5-18 for eigenvectors $\alpha_1, \alpha_2, \ldots, \alpha_m$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$. If the eigenvector solutions from the range of p+1 to *m* are reduced, the dimensionality reduction for PCA can be operated for the eigenvectors $\alpha_1, \alpha_2, \ldots, \alpha_p$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$, where $\lambda_p \geq \lambda_m$. The $\alpha_1, \alpha_2, \ldots, \alpha_p$ are normalised if the corresponding vectors $\nu_k \in \mathbb{F}$ be normalised, i.e. $\nu_k \cdot \nu_k = 1$ and $\nu_k = \sum_{i=1}^{m} \alpha^k{}_i \, \Phi(x_i)$. Apply the Gram matrix *K*,

$$1 = \sum_{i,j=1}^{m} \alpha^k{}_i \alpha^k{}_j (\Phi(x_i) \cdot \Phi(x_j)) \tag{5-19}$$

$$= \alpha^k \cdot K\alpha^k \tag{5-20}$$

$$= \lambda_k(\alpha^k \cdot \alpha^k) \qquad \text{where k = 1, …, p.} \tag{5-21}$$

The projection on the normalised eigenvector ν_k in feature subspace $\mathbb{F}$ for a test point x can be found as:

$$\nu_k \cdot \Phi(x) = \sum_{i=1}^{m} \alpha^k{}_i (\Phi(x_i) \cdot \Phi(x)) \tag{5-22}$$

The kernel function $\Phi(x_i) \cdot \Phi(x_j) = K(x_i, x_j)$, which acts as a dot product to satisfy the Mercer's theorem for mapping into the feature space $\mathbb{F}$, can be used to reduce the computational cost for non-linear information of process. Using the kernel function of the Gaussian kernel in Eq. 5-13, the dot product in feature space can be reduced to a function in input space.

Referring to Eq. 5-13, the performance of a Gaussian kernel is highly dependent on the choice of bandwidth. As such, the simulations on different bandwidths of the Gaussian kernel were carried out using the stability data of different operating scenarios as described in Section 5.5. Nine bandwidths sizes were selected as 0.05, 0.75, 1, 10, 20, 40, 60, 80 and 100.

Using ReliefF, the values of all the features from the simulation of the 9 bus system were calculated and shown in Table 5- 2 in descending order. Their distinct values imply the ability to classify stability instance.

Table 5- 2 Features and their ReliefF weights on 9 bus system

P_{G2}	Q_{T7}	P_{T4}	θ_2	θ_8	θ_3	θ_7	θ_6	P_{T2}	Q_{T2}
0.3234	0.3233	0.3233	0.2576	0.2432	0.2327	0.2312	0.2306	0.2106	0.2068
Q_{T5}	P_{T1}	Q_{T1}	P_{G1}	P_{T3}	θ_5	Q_{L7}	θ_4	P_{T5}	Q_{T9}
0.1984	0.1976	0.1976	0.1975	0.1952	0.191	0.1904	0.1880	0.1792	0.1774
θ_9	Q_{G2}	P_{L7}	Q_{G1}	Q_{G3}	Q_{L5}	Q_{T8}	P_{L5}	P_{T8}	P_{T3}
0.1721	0.1381	0.135	0.1328	0.1306	0.1178	0.1052	0.1022	0.1017	0.1003
Q_{T3}	Q_{L9}	P_{T6}	Q_{T6}	V_5	V_7	V_9	V_4	V_8	V_6
0.0978	0.0782	0.0533	0.0517	0.0220	0.019	0.0188	0.0156	0.0132	0.0115

5.6 Numerical Simulation

5.6.1 *WECC 3-machine 9 Bus Power System*

The proposed scratch DenseNet model for transient stability prediction was trained using the data obtained from the dimensionality reduction by ReliefF and KPCA method as described in Section 5.5.

To verify the benefits of the proposed transfer learning approach, the techniques of feature extraction and finetuning were repeated with the same training and validation datasets for comparison with the scratch model. The classification results of different techniques for the above test data are shown in the Case 5-1 and 5-2.

As compared to the effectiveness of ReliefF, the data obtained by using KPCA method for non-linear transformation is used to train the scratch model of DenseNet.

Table 5- 3 Selected features

Components	Operating Features
Buses	Bus voltage magnitude (V_i)
	Bus voltage angle (θ_i)
Transmission lines	Real power flow (P_{Ti})
	Reactive power flow (Q_{Ti})
Generators, loads	Net real power injection (P_{Gi} , P_{Li})
	Net reactive power injection (Q_{Gi}, Q_{Li})

Also, the selected features as listed in Table 5- 3 are reduced to the same dimensions as ReliefF did. The respective results of validation accuracy are shown in Table 5- 4. It indicates that the bandwidth parameter γ = 20 has the best validation accuracy of 0.990283. This value is higher than the best validation accuracy of 0.982843 obtained from the simulation data of ReliefF, as shown in Table 5- 5. The bandwidthγ = 20 was selected to proceed with the transfer learning approach.

Table 5- 4 Respective best validation accuracy for the scratch model of DenseNet based on bandwidth selection of KPCA

Bandwidth (γ)	0.05	0.75	1
Best Val Acc	0.715909	0.955482	0.972403
Bandwidth (γ)	10	20	40
Best Val Acc	0.978896	0.990283	0.989448
Bandwidth (γ)	60	80	100
Best Val Acc	0.984578	0.982941	0.977273

According to Section 5.4, the simulation results of the scratch DenseNet model with data from ReliefF and KPCA are compared with the transfer learning approach, as shown in Figures 5- 2, and 5- 3 and Tables 5- 5 to 5- 7.

The DenseNet obtained scratch train and validation results with data from ReliefF and KPCA at Epoch 14/14, as shown in Table 5- 5. It is the scratch model that does not use the transfer learning technique during the model training.

Table 5- 5 Simulation results of scratch DenseNet model with data obtained from ReliefF and KPCA

Data source (ReliefF)	Train Loss: 0.1466, Acc: 0.9409. Val Loss: 0.0537, Acc 0.9796. Training complete in 633m 24s. Best Val Acc: 0.982843.
Data source (KPCA, $\gamma = 20$)	Train Loss: 0.3953, Acc: 0.8100. Val Loss: 0.1187, Acc: 0.9822. Training complete in 802m 54s. Best Val Acc: 0.990283.

As described in Section 5.3, the proposed transfer learning approach can improve validation accuracy or reduce training completion time compared to the simulation result of scratch DenseNet model, which does not use the transfer learning as illustrated in Figures 5-2 and 5- 3. The following two case studies of transfer learning demonstrate the effectiveness on 9 bus system if the original pretrained dataset comes from ImageNet.

Case 5-1: Feature extraction

Apply feature extraction technique of transfer learning to verify if training completion time is reduced. The DenseNet obtained train and validation results with data from ReliefF and KPCA at Epoch 14/14 are shown in Table 5- 6.

Table 5- 6 Simulation results of feature extraction technique with data obtained from ReliefF and KPCA

Data source (ReliefF)	Train Loss: 0.3116, Acc: 0.8700. Val Loss: 0.1337, Acc: 0.9453. Training complete in 313m 47s. Best Val Acc: 0.962418.
Data source (KPCA, $\gamma = 20$)	Train Loss: 0.6100, Acc: 0.7108. Val Loss: 1.0730, Acc: 0.3247. Training complete in 310m 39s. Best Val Acc: 0.854251

So, the feature extraction technique reduces training completion time from 633 m 24s to 313 m 47s and 802 m 54s to 310 m 39s for data obtained from ReliefF and KPCA, respectively. As previously described, if the feature dataset of the power system is used to pretrain the DenseNet before feature extraction of transfer learning is performed, it becomes a traditional machine learning approach. The validation accuracy is close to the pretrained DenseNet.

Case 5-2: Finetuning

Apply the finetuning of transfer learning to verify if validation accuracy is improved. The DenseNet obtained train and validation results with data from ReliefF and KPCA at Epoch 14/14, as shown in Table 5- 7.

Table 5- 7 Simulation results of finetuning technique with data obtained from ReliefF and KPCA

Data source (ReliefF)	Train Loss: 0.0994, Acc: 0.9634. Val Loss: 0.0627, Acc: 0.9796. Training complete in 626m 1s. Best Val Acc: 0.995098.
Data source (KPCA, $\gamma = 20$)	Train Loss: 0.263, Acc: 0.8794. Val Loss: 0.0192, Acc: 0.9943. Training complete in 801m 26s. Best Val Acc: 0.995951.

So, the best validation accuracies are increased from 0.982843 to 0.995098 and 0.990283 to 0.995951 for data obtained from ReliefF and KPCA, respectively.

Figures 5- 2 and 5- 3 illustrate the effects if the DenseNet model applies transfer learning techniques.

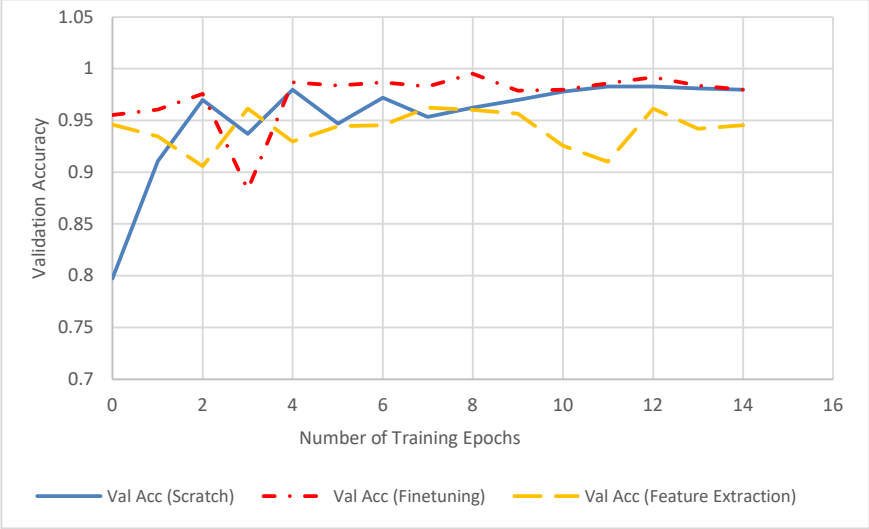

Fig. 5- 2 Simulation for transfer learning of feature extraction and finetuning, and scratch model with the dataset from the ReliefF technique

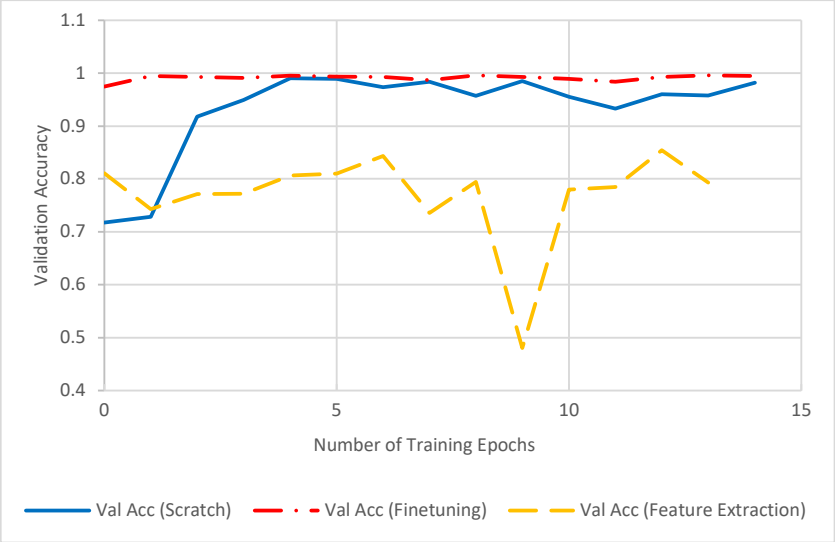

Fig. 5- 3 Simulation for transfer learning of feature extraction and finetuning, and scratch model with the dataset from the KPCA technique

The dark blue line shows that the scratch DenseNet model, which does not use the transfer learning technique, is trained to classify the pre-fault stability status, as described in Section 5.5. It attains validation accuracy 0.9796, or 2.04% error for source data from ReliefF and validation accuracy 0.9822, or 1.78% error for source data from KPCA, respectively, at Epoch 14/14.

The orange dash line shows the effect if DenseNet model applies the feature extraction technique. The training completion time drops from 633 m 24s to 313 m 47s or 49.54% for source data from ReliefF and 802 m 54s to 310 m 39s, or 38.69% for source data from KPCA, respectively, at Epoch 14/14. Its performance is dropped below the dark blue line of the scratch DenseNet model, as shown in Figures 5- 2 and 5- 3. The feature transfer from ImageNet trained model to the layer of DenseNet model reflects the specificity of representation to dominate the drop in performance. However, when the target and source domains are the same, and their learning tasks are the same, it becomes a traditional machine learning problem to resolve the drop-in performance.

The red dash line shows a very positive effect if the ImageNet trained model transfers features to DenseNet model and finetune them. It generalises better performance than the scratch DenseNet model, which does not use the transfer learning technique. It has the advantage not only to enable training without overfitting on small target datasets but also boost generalisation performance even if the target dataset is small as 4900 OPs.

5.6.2 *New England 39 Bus Power System*

The proposed transfer learning approach which uses the ImageNet as pretrained dataset is further validated on New England 39 bus power system. The simulation results are shown in Tables 5- 1, 5- 8 to 5- 10 and Figures 5- 4 to 5- 6. Tables 5- 8 to 5- 10 and Figures 5- 4 to 5- 6 illustrate the simulation results to validate the findings on Cases 5-1 and 5-2 that the feature extraction technique reduces training completion time, and finetuning technique improves validation accuracy.

In the scenario of bus #3, the feature extraction technique reduces the training completion time from 373 m 27s to 163 m 7s or 43.68% at Epoch 29/29. The best validation accuracy also increases from 0.616 to 0.696. The finetuning technique improves validation accuracy from 0.616 to 0.824, as shown in Table 5- 8 and Figure 5- 4.

Table 5- 8 Simulation results of scratch, feature extraction and finetuning techniques with data obtained from ReliefF for contingency of three phase fault on bus # 3

Scratch	Train Loss: 0.7035, Acc: 0.519. Val Loss: 0.6813, Acc: 0.52. Training complete in 373m 27s. Best Val Acc: 0.616.
Feature extraction	Train Loss: 0.722, Acc: 0.5521. Val Loss: 0.6005, Acc: 0.648. Training complete in 163m 7s. Best Val Acc: 0.696.
Finetuning	Train Loss: 0.5891, Acc: 0.7299. Val Loss: 0.6227, Acc: 0.688. Training complete in 376m 12s. Best Val Acc: 0.824.

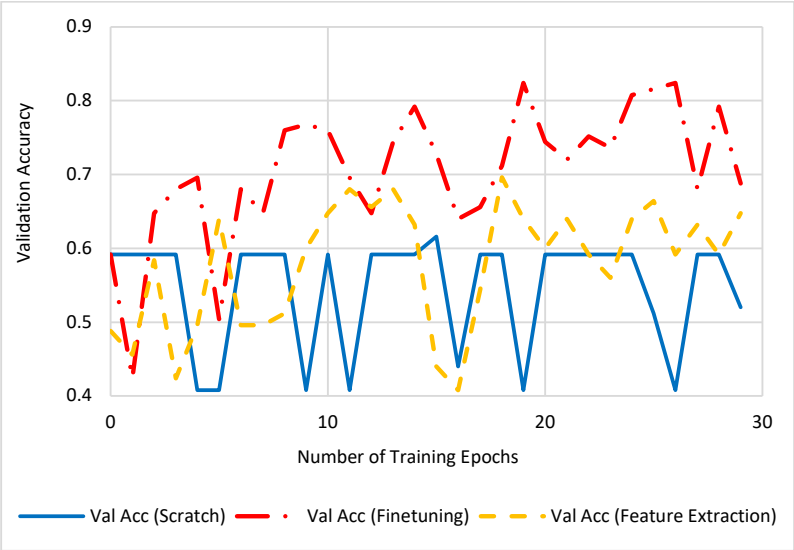

Fig. 5- 4 Simulation for transfer learning of feature extraction and finetuning, and scratch model with the dataset from the ReliefF technique for the contingency of three phase fault on bus #3

In the scenario of bus #19, the feature extraction technique reduces the training completion time from 354 m 50s to 155 m 16s, or 43.76% at Epoch 29/29. The finetuning technique improves validation accuracy from 0.936 to 0.952, as shown in Table 5- 9 and Figure 5- 5.

Table 5- 9 Simulation results of scratch, feature extraction and finetuning techniques with data obtained from ReliefF for contingency of three phase fault on bus #19

Scratch	Train Loss: 0.4726, Acc: 0.7473. Val Loss: 0.2403, Acc: 0.9355. Training complete in 354m 50s. Best Val Acc: 0.936.
Feature extraction	Train Loss: 0.6895, Acc: 0.621. Val Loss: 0.6451, Acc: 0.621. Training complete in 155m 16s. Best Val Acc: 0.871.
Finetuning	Train Loss: 0.365, Acc: 0.8414. Val Loss: 0.1284, Acc: 0.9435. Training complete in 352m 25s. Best Val Acc: 0.952.

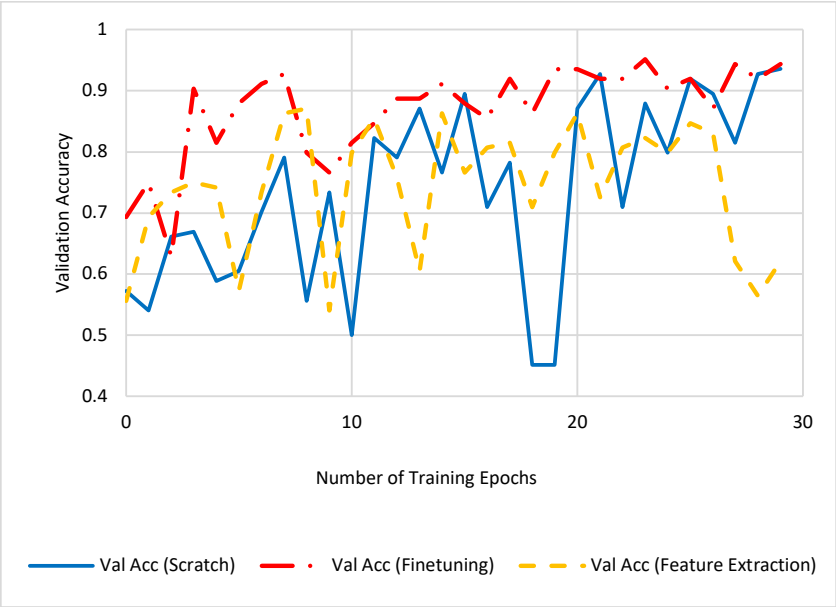

Fig. 5- 5 Simulation for transfer learning of feature extraction and finetuning, and scratch model with the dataset from the ReliefF technique for the contingency of three phase fault on bus #19

In the scenario of bus #22, the feature extraction technique reduces the training completion time from 347 m 18s to 158 m 47s or 45.72% at Epoch 29/29. The finetuning technique improves validation accuracy from 0.848 to 0.872, as shown in Table 5- 10 and Figure 5- 6.

Table 5- 10 Simulation results of scratch, feature extraction and finetuning techniques with data obtained from ReliefF for contingency of three phase fault on bus #22

Scratch	Train Loss: 0.6089, Acc: 0.6927. Val Loss: 0.4189, Acc: 0.808. Training complete in 347m 18s. Best Val Acc: 0.848
Feature extraction	Train Loss: 0.6468, Acc: 0.6739. Val Loss: 0.4278, Acc: 0.808. Training complete in 158m 47s. Best Val Acc: 0.824.
Finetuning	Train Loss: 0.6275, Acc: 0.7089. Val Loss: 0.488, Acc: 0.824. Training complete in 287m 52s. Best Val Acc: 0.872.

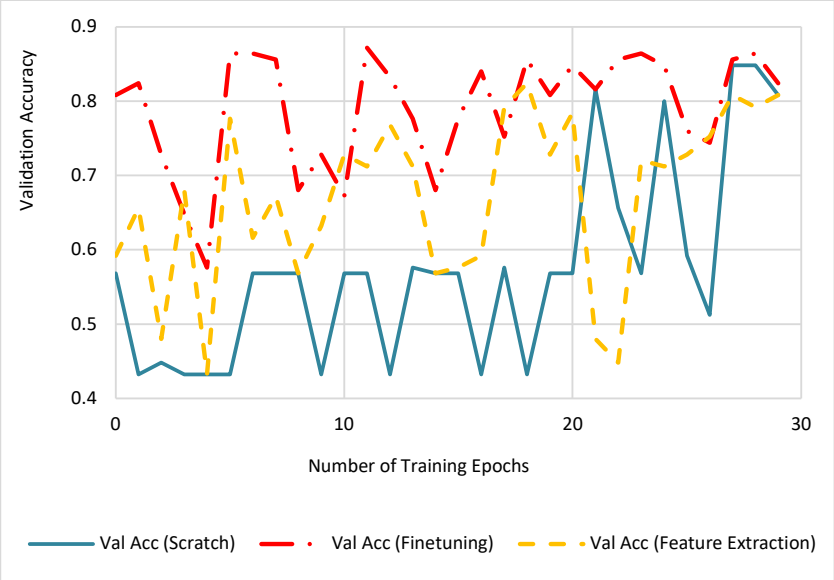

Fig. 5- 6 Simulation for transfer learning of feature extraction and finetuning, and scratch model with the dataset from the ReliefF technique for the contingency of three phase fault on bus #22

The simulation results using the dataset obtained from New England 39 bus system as shown in Tables 5- 8 to 5- 10 and Figures 5- 4 to 5- 6 have further validated Section 5.6.1, that the feature extraction technique shortens the training completion time and finetuning technique improves validation accuracy for pretrained ImageNet dataset. It also shows that the finetuning technique can reach high validation accuracy compared to scratch model training or feature extraction technique after a few training epochs. The curves shown on Figures 5- 4 to 5- 6 also match the observation from Figures 5- 2 and 5- 3.

In the scenario of bus #3, no transmission line is tripped as protection measures. In the scenarios of bus #19 and #22, transmission lines are tripped, respectively. On a large complex power system, it is a general practice to trip the fault-related transmission lines to isolate the fault location and stabilise the power system under severe fault condition. The scenarios of bus #19 and #22 which achieve high validation accuracy are more applicable to industrial practice.

5.7 Conclusion

The simulation results of transient stability prediction supported by feature extraction and finetuning techniques for pretrained ImageNet dataset and target dataset obtained from the WECC 9 bus system with dimensionality reduction by ReliefF and KPCA are shown in Figures 5- 2 and 5- 3 and Tables 5- 5 to 5- 7. The results illustrate that the feature extraction technique reduces training completion time and finetuning technique improves validation accuracy if transfer learning approach is applied, respectively. It also demonstrates that the finetuning technique maintains stronger convergence ability with higher validation accuracy than the feature extraction technique or scratch DenseNet model. The simulation results using the target dataset obtained from the New England 39 bus system with pretrained ImageNet dataset, as shown in Tables 5- 8 to 5- 10 and Figures 5- 4 to 5- 6 have further validated the above conclusion.

As the ReliefF and KPCA present the target dataset in a different feature space, the DenseNet model with finetuning technique is capable of achieving the high validation accuracy around 0.995098 to 0.995951 in Section 5.6.1. The feature extraction technique can reduce training completion time to 49.54% and 38.69%, respectively.

Besides, it demonstrates that KPCA is more effective than ReliefF to process features for dimensionality reduction. Also, the DenseNet successfully extracts features from the 2D histogram images of simulated OPs for further processing of transient stability assessment. The proposed approach works for pre-fault DSA, and the procedure is also appliable to post-fault DSA.

The methodologies can be applied to different power systems. A figure for different IEEE bus systems is included in the appendix section. Some typical benchmark power systems with

different buses, such as the IEEE 118 bus, IEEE 39 bus, and IEEE 14 bus systems, can be referred to [267-269]. Readers can download the figures from these references.

According to [270], the IEEE 118 bus system contains 19 generators, 91 loads, 9 transformers, 177 lines, and 35 synchronous condensers [271]. This simplified representation of the American Electric Power system in the Midwestern United States was published in 1962, which later became well known as the IEEE 118 bus system. The system incorporates multiple voltage levels, with some buses operating at 138 kV and 345 kV. It is a benchmark for planning. The IEEE 39 bus system, also known as the New England 39 bus system, is another commonly used benchmark in power system analysis and research. It represents a simplified network of the New England power system in the northeastern region of the United States. The IEEE 39 bus system consists of 39 buses connected by 46 transmission lines, and it includes 10 generators, loads, and transformers [242, 268]. The buses in the system operate at 345 kV. This system serves as a standard test case for stability. The IEEE 14 bus system is a simplified network of the American Electric Power system as of February 1962 [267]. It consists of 14 buses connected by transmission lines and includes 5 generators, loads, and transformers. The buses in the system operate at various voltage levels, such as 13.8 kV, 18 kV, and 69 kV. It was initially developed for state estimation purposes.

6 Conclusions and Future Work

This research concerns the relationship between load, load model, power system stability, and the application of artificial intelligence to dynamic security assessment. The concepts of the load model and stability are discussed in Chapters 1 and 2. Section 2.3 concludes that the load model affects power system stability. The applications of artificial intelligence to load model and power system stability are described in Section 2.4.

In Chapter 3, this chapter investigates the possibility of developing a dynamic load model to represent the power system load characteristic during system fault when the power system operates at a high percentage of the power generation from wind farms, solar power and V2G. A new hybrid load model, consisting of two different load models representing the active and reactive power response respectively, is tested. The simulation results conclude that the proposed hybrid load model has a reasonable low fitting error to represent the system load dynamics on both active and reactive power responses.

In Chapter 4, the simulation results conclude that GANs can effectively generate synthetic data to address the challenge of limited data in NILM. Additionally, the transfer learning technique can be applied to resolve the mode collapse problem that arises when GANs generate synthetic data. Moreover, this technique helps stabilise the GAN training process in a shorter number of iterations. Considering the increasing number of new types of electrical devices being connected to the electricity network, the proposed methodology is suitable for domestic houses, and the procedure is also applicable to NILM applications in the industrial sector.

In Chapter 5, the simulation results of transient stability prediction supported by feature extraction and finetuning techniques for pretrained ImageNet dataset and target dataset

obtained from the WECC 9 bus system with dimensionality reduction by ReliefF and KPCA conclude that the feature extraction technique reduces training completion time and finetuning technique improves validation accuracy if transfer learning approach is applied, respectively. It also demonstrates that the finetuning technique maintains stronger convergence with higher validation accuracy than the feature extraction technique on scratch DenseNet model. The simulation results using the target dataset obtained from the New England 39 bus system with pretrained ImageNet dataset have further supported the above conclusion. Besides, it demonstrates that KPCA is more effective than ReliefF in processing features for dimensionality reduction. Also, the DenseNet successfully extracts features from the 2D histogram images of simulated OPs for further processing of transient stability assessment. The proposed approach works for pre-fault DSA, and the procedure is also applicable to post-fault DSA.

The researches mentioned above are only a part of the artificial intelligence transformation in power systems. For instance, Section 2.4 describes an artificial neural network-based dynamic load model that can accurately represent the dynamics of active and reactive power responses for a given disturbance. This differs from the traditional load model, which consists of components with physical meaning or reliable formulas.

It is challenging to maintain a stable grid with high integration of a significant increase in intermittent renewable energy to fulfil the low-carbon transition. As artificial intelligence supports the seamless integration of an increasingly distributed and multi-directional power system to unlock the vast potential of renewables, it relies on communications and data to make the system work. However, sufficient training data may not be available in some situations. The research outcome from this thesis supports part of the solution to this issue. GANs can generate datasets with the same statistical properties as the original target data,

e.g. solar and wind data. GANs and their variants may be further developed for a wider application.

In reviewing the challenges in online DSA, the data from PMU may be corrupted with noise. Therefore, employing new application concepts such as transfer learning, GANs, and other techniques can enhance the following research.

- The power system is monitored by the online DSA with unlearned faults, incomplete data, or unlearned faults and incomplete data.
- The confidence awareness of the trained model, which is used for online DSA with unlearned faults and incomplete data.

Reference

[1] A. Kahan, "EIA projects nearly 50% increase in world energy usage by 2050, led by growth in Asia," US Energy Information Administration, 2019. Accessed: Dec. 2022. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=41433

[2] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel, "Disaggregated end-use energy sensing for the smart grid," *IEEE Pervasive Comput.,* vol. 10, no. 1, pp. 28-39, Jan. 2011, doi: 10.1109/MPRV.2010.74.

[3] K. Yamashita *et al.*, "Modelling and aggregation of loads in flexible power networks: Scope and status of the work of CIGRE WG C4.605," *IFAC Proceedings Volumes,* vol. 45, no. 21, pp. 405-410, Jan. 2012, doi: 10.3182/20120902-4-FR-2032.00072.

[4] L. M. Korunovi´c, M. Ra˘si´c, N. Floranovi´c, and V. Aleksi´c, "Load modelling at low voltage using continuous measurements," *Facta Universitatis, Series: Electronics and Energetics,* vol. 27, no. 3, pp. 455–465, 2014.

[5] J. Marchgraber, E. Xypolytou, I. Lupandina, W. Gawlik, and M. Stifter, "Measurement-based determination of static load models in a low voltage grid," in *IEEE PES ISGT Europe*, Ljubljana, Slovenia, Oct. 2016, pp. 1–6.

[6] L. M. Korunovi´c, A. S. Jovi´c, and S. Z. Djokic, "Measurement-based evaluation of static load characteristics of demands in administrative buildings," *Int. J. Electr. Power Energy Syst.,* vol. 118, pp. 1–8, 2020.

[7] M. Merkle, "Probability and statistics for engineers and students of technical sciences," (in Serbian), *Akademska misao,* 2006.

[8] J. C. Lam, H. L. Tang, and D. H. W. Li, "Seasonal variations in residential and commercial sector electricity consumption in Hong Kong," *Energy Build,* vol. 33, pp. 513–523, 2008.

[9] A. S. Jovi´c, L. M. Korunovi´c, and S. Z. Djokic, "Application of meteorological variables for the estimation of static load model parameters," *Energies (Basel),* vol. 14, no. 16, p. 4874, 2021.

[10] N. J. Radmila, Ž. Mileva, L. Miodrag, and P. Dubravka, *Basics of Statistical Analysis*. Beograd, Serbia: Savremena Administracija (in Serbian), 1989.

[11] Multicollinearity and Other Problems of Linear Regression [Online] Available: www.matf.bg.ac.rs/p/files/69-Multikolinearnost.html

[12] OriginLab Tutorials [Online] Available: https://www.originlab.com/doc/Tutorials

[13] G. S. Vassell, "Northeast blackout of 1965," *IEEE Power Engineering Review,* vol. 11, no. 1, pp. 4-8, Jan. 1991.

[14] V. Ajjarapu and C. Christy, "The continuation power flow: a tool for steady state voltage stability analysis," in *Proc. Conf. IEEE PICA*, May 1991, pp. 304–311.

[15] P. A. Lof, T. Smed, G. Andersson, and D. J. Hill, "Fast calculation of a voltage stability index," *IEEE Trans. Power Syst.,* vol. 7, pp. 54–64, Feb. 1992.

[16] P. Kundur *et al.*, "Definition and classification of power system stability - IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.,* vol. 19, no. 2, pp. 1387-1401, May 2004.

[17] R. S. Zhang, Q. G. Qi, C. B. Li, L. Li, Y. P. Bao, and Y. S. Zhu, "Forecasting of load model based on typical daily load profile and BP neural network," in *Proc. Conf. ICSEM*, Singapore, 2012, vol. 34, pp. 38-43.

[18] D. N. Kosterev, C. W. Taylor, and W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Trans. Power Syst.,* vol. 14, no. 3, pp. 967–979, Aug. 1999.

[19] C. Concordia and S. Ihara, "Load representation in power system stability studies," *IEEE Trans. Power Appar. Syst.,* vol. 101, pp. 969–977, 1982.

[20] A. Adrees and J. Milanović, "Effect of load models on angular and frequency stability of low inertia power networks," *IET Gener. Transm. Distrib.,* vol. 13, no. 9, pp. 1520-1526, Apr. 2019.

[21] J. Cho, K. Lee, E. Shin, G. Choy, and S. Do, "How much data Is needed to train a medical image deep learning system to achieve necessary high accuracy?," in *Proc. Conf. ICLR*, Jan. 2016.
[22] P. Warden, "How Many Images Do You Need to Train A Neural Network?," ed, 2018.
[23] J. Donahue *et al.*, "Decaf: A deep convolutional activation feature for generic visual recognition," in *Prof. Conf. ICML*, 2014, pp. 647–655.
[24] Australian Energy Market Operator. "About the National Electricity Market (NEM)." https://aemo.com.au/en/energy-systems/electricity/national-electricity-market-nem/about-the-national-electricity-market-nem (accessed Jun. 2023).
[25] "Electric power system." Wikipedia. https://en.wikipedia.org/wiki/Electric_power_system (accessed Jun. 2023).
[26] "High voltage dc breakthrough." National Geographic. https://www.nationalgeographic.com/science/article/121206-high-voltage-dc-breakthrough (accessed Jun. 2023).
[27] Transgrid. *What is the high voltage transmission network?* [Online]. Available: https://www.transgrid.com.au/media/b2hpaiso/high-voltage-transmission-line-factsheet.pdf. Accessed: Jun. 2023.
[28] "Network Standard: NS238 Supply Quality," Ausgrid, May 2015.
[29] "Managing frequency in the power system." Australian Energy Market Operator (AEMO). https://aemo.com.au/learn/energy-explained/energy-101/managing-frequency-in-the-power-system (accessed Jun. 2023).
[30] S. Chapman, "Electric Machinery and Power System Fundamentals." Boston: McGraw-Hill, 2002, ch. 11.
[31] F. K. Basha and M. Thompson, "Practical EHV Reactor Protection," in *Annual Conf. CFPR*, TX, USA, Apr. 2013: Schweitzer Engineering Laboratories, Inc.
[32] W. Machine. (Feb. 2010)How does an RCD work? *PowerBreaker*.
[33] A. Maitra and A. Gaikwad, "Measurement based load modeling," Electric Power Research Institute (EPRI), Sep. 2006.
[34] P. Kundur, *Power System Stability and Control*. New York, USA: McGraw-Hill, Inc., Jan. 1994.
[35] J. Hou, Z. Xu, and Z. Y. Dong, "Load modeling practice in a smart grid environment," in *Proc. 4th Int. conf. DRPT*, Weihai, China, Jul. 2011: IEEE, pp. pp. 7-13, doi: 10.1109/DRPT.2011.5993854.
[36] W. S. Kao, C. J. Lin, C. T. Huang, Y. T. Chen, and C. Y. Chiou, "Comparison of simulated power system dynamics applying various load models with actual recorded data," *IEEE Trans. Power Syst.,* vol. 9, no. 1, pp. 248-254, Feb. 1994.
[37] W. S. Kao, "The effect of load models on unstable low-frequency oscillation damping in Taipower system experience w/wo power system stabilizers," *IEEE Trans. Power Syst.,* vol. 16, no. 3, pp. 463-472, Aug. 2001.
[38] J. V. Milanović, K. Yamashita, S. M. Villanueva, S. Ž. Djokić, and L. M. Korunović, "International industry practice on power system load modelling," in *IEEE PES General Meeting | Conference & Exposition*, National Harbor, MD, USA, Jul. 2014: IEEE, pp. pp. 1-1, doi: 10.1109/PESGM.2014.6938972.
[39] J. V. Milanovic, "On unreliability of exponential load models," *Electric Power Systems Research,* vol. 49, pp. 1–9, 1999.
[40] "Load representation for dynamic performance analysis," *IEEE Trans. Power Syst.,* vol. 8, no. 2, pp. 472–482, May 1993.
[41] L. Rodríguez-García, S. Pérez-Londoño, and J. Mora-Flórez, "Dynamic load modeling for small disturbances using measurement-based parameter estimation method," in *Proc. VII Int. Symposium on Power Quality (SICEL)*, 2013.
[42] D. J. Hill, "Nonlinear dynamic load models with recovery for voltage stability studies," *IEEE Trans. Power Syst.,* vol. 8, no. 1, pp. 166-176, Feb. 1993.

[43] M. A. Salam, "Equal Area Criterion," in *Fundamentals of Electrical Power Systems Analysis*, 1 ed.: Springer Nature Singapore Pte Ltd., 2020, ch. 9.9 p. 436.
[44] F. Shi, H. Zhang, and G. Xue, "Instability prediction of the inter-connected power grids based on rotor angle measurement," *Electrical Power and Energy Systems,* vol. 88, pp. 21-32, 2017.
[45] "Power System Stability Guidelines," Australian Energy Market Operator Ltd, May 2012.
[46] *Frequency Operating Standard*, A. E. M. O. Ltd., Jan. 2020.
[47] M. Y. Huh, P. Agrawal, and A. A. Efros, "What makes ImageNet good for transfer learning," in *Proc. Conf. IEEE CVPR*, Las Vegas, NV, Aug. 2016, pp. 1–10.
[48] "Frequency stability evaluation criteria for the synchronous zone of continental europe – Requirements and impacting factors," European Network of Transmission System Operators For Electricity entsoe, Mar. 2016.
[49] M. X. Wang and J. V. Milanovic, "The effect of the type and composition of demand on DSM contribution to system frequency stability," in *Proc. Conf. IEEE ISGT-Europe*, The Hague, Netherlands, Oct. 2020, doi: 10.1109/ISGT-Europe47291.2020.9248926.
[50] N. Hatziargyriou *et al.*, "Stability definitions and characterization of dynamic behavior in systems with high penetration of power electronic interfaced technologies," IEEE PES, PES-TR77, Apr. 2020.
[51] R. Schulz and P. Pourbeik, "Interconnected power system response to generation governing: present practice and outstanding concerns," IEEE PES, PES -TR13, May 2007.
[52] A. Sajadi, R. Preece, and J. V. Milanovi´c, "Establishing multidimensional transient stability boundaries for power systems with uncertainties," in *Proc. IEEE Conf. PSCC*, Dublin, Ireland, Jun. 2018.
[53] L. B. Shi, S. Q. Dai, Y. X. Ni, L. Z. Yao, and M. Bazargan, "Transient stability of power systems with high penetration of DFIG based wind farms," in *IEEE PES*, 2009, pp. 1-6.
[54] E. Telegina, "Impact of rotational inertia changes on power system stability," 2015.
[55] C. S. Saunders, M. M. Alamuti, and G. A. Taylor, "Transient stability analysis using potential energy indices for determining critical generator sets," in *IEEE PES General Meeting | Conference & Exposition*, National Harbor, MD, USA, Jul. 2014, pp. 1-5.
[56] J. D. Morales, X. Ye, and J. V. Milanović, "Comparative analysis of integral-based indices for on-line assessment of power system transient stability," in *IEEE PES ISGT Europe*, Espoo, Finland, Oct. 2021.
[57] L. Meegahapola and D. Flynn, "Impact on transient and frequency stability for a power system at very high wind penetration," in *IEEE PES General Meeting*, 2010, pp. 1-8.
[58] "Transient Security Assessment Tool: User Manual, Tech. Rep.," Powertech Labs Inc., Surrey, British Columbia, Canada, 2011.
[59] M. Reza, P. H. Schavemaker, J. G. Slootweg, W. L. Kling, and L. v. d. Sluis, "Impacts of distributed generation penetration levels on power systems transient stability," in *IEEE PES General Meeting*, Jun. 2004.
[60] L. Meegahapola and T. Littler, "Characterisation of large disturbance rotor angle and voltage stability in interconnected power networks with distributed wind generation," *IET Renew. Power Gener.,* Sep. 2014, doi: 10.1049/iet-rpg.2013.0406.
[61] M. Pavella, D. Ernst, and D. Ruiz-Vega, *Transient Stability of Power Systems: A Unified Approach to Assessment and Control*, 1st ed. Kluwer Academic Publishers, 2000.
[62] F. P. Demello and C. Concordia, "Concepts of synchronous machine stability as affected by excitation control," *IEEE Trans. Power Appar. Syst.,* vol. 88, no. 4, pp. 316-328, 1969.
[63] Y. V. Makarov, Z. Y. Dong, and D. J. Hill, "A general method for small signal stability analysis," *IEEE Trans. Power Syst.,* vol. 13, no. 3, Aug. 1998.
[64] Z. Y. Dong, Y. V. Makarov, and D. J. Hill, "Genetic algorithms in power system small signal stability analysis," in *Proc. 4th Int. Conf. APSCOM*, HK, Nov. 1997.
[65] P. W. Sauer, B. C. Lesieutre, and M. A. Pai, "Maximum loadability and voltage stability in power systems," *JEPE,* vol. 15, pp. 145-154, Jun. 1993.

[66] M. A. Pai, "Structural Stability in Power Systems - Effect of Load Models," in *Systems and Control Theory for Power Systems*, J. H.Chow, P. V. Kokotovic, and R. J. Thomas Eds.: Springer-Verlag, 1995, pp. 259-281.

[67] P. W. Sauer and B. C. Lesieutre, "Power system load modeling," in *Systems and Control Theory for Power Systems*: Springer-Verlag, 1995.

[68] P. W. Sauer and M. A. Pai, "Power system steady-state stability and the load-flow Jacobian," *IEEE Trans. Power Syst.,* vol. 5, no. 4, pp. 1374-1383, Nov. 1990.

[69] M. A. Pai, A. Kulkarni, and P. W. Sauer, "Parallel dynamic simulation of power systems," in *Proc. Conf. IEEE ISCAS*, New Orleans, LA, US, May 1990, pp. 1264-1267, doi: 10.1109/ISCAS.1990.112360.

[70] C. Rajagopalan, P. W. Sauer, and M. A. Pai, "Analysis of voltage control systems exhibiting Hopf bifurcation," in *Proc. IEEE Conf. CDC*, Tampa, Florida, Dec. 1989, pp. 332-335.

[71] C. Rajagopalan, P. W. Sauer, and M. A. Pai, "An integrated approach to dynamic and static voltage stability," in *A CC'89*, Pittsburgh, Jun. 1989.

[72] C. Rajagopalan, B. Leisieutre, P. W. Sauer, and M. A. Pai, "Dynamic aspects of voltage power characteristics," *IEEE Trans. Power Syst.,* vol. 7, no. 3, pp. 990-1000, Aug. 1992.

[73] V. Venkatasubramanian, H. Schattler, and J. Zaborszky, "Dynamics of large constrained nonlinear systems - a taxonomy theory," *Proc. IEEE Inst. Electr. Electron Eng.,* vol. 83, no. 11, pp. 1530-1561, Nov. 1995.

[74] V. A. Venikov, V. A. Stroev, V. I. Idelchik, and V. I. Tarasov, "Estimation of electric power system steady-state stability in load flow calculation," *IEEE Trans. Power Appar. Syst.,* pp. 1034 -1041, May 1975.

[75] I. Dobson, "Computing a closest bifurcation instability in multidimensional parameter space," *J Nonlinear Sci,* vol. 3, no. 1, pp. 307-327, Dec. 1993, doi: 10.1007/BF02429868.

[76] Y. V. Makarov and I. A. Hiskens, "A continuation method approach to finding the closest saddle node bifurcation point," in *Proc. NSF/ECG Workshop on Bulk Power System Voltage Phenomena* Davos, Switzerland, Aug. 1994.

[77] I. A. Hiskens and R. J. Davy, "Exploring the power flow solution space boundary," *IEEE Trans. Power Syst.,* vol. 16, no. 3, pp. 389-395, Aug. 2001, doi: 10.1109/59.932273.

[78] D. J. Hill *et al.*, "Advanced small disturbance stability analysis techniques and MATLAB algorithms," in "Collaborative Research Project Advanced System Analysis Techniques," New South Wales Electricity Transmission Authority, The Department of Electrical Engineering, The University of Sydney, Apr. 1996.

[79] I. A. Gruzdev, V. A. Maslennikov, and S. M. Ustinov, "Development of methods and software for analysis of steady-state stability and damping of bulk power systems," *Methods and Software for Power System Oscillatory Stability Computations,* pp. 66-88, 1992.

[80] P. W. Sauer, M. A. Pai, and J. H. Chow, *Power System Dynamics and Stability - With Synchrophasor Measurement and Power System Toolbox*. JohnWiley & Sons Ltd, 2018.

[81] N. Hatziargyriou *et al.*, "Definition and Classification of Power System Stability – Revisited & Extended," *IEEE Trans. Power Syst.,* vol. 36, no. 4, Jul. 2021.

[82] IEEE Subsynchronous resonance working group, "Reader's guide to subsynchronous resonance," *IEEE Trans. Power Syst.,* vol. 7, no. 1, pp. 150–157, Feb. 1992.

[83] IEEE subsynchronous resonance working group of the system dynamic performance subcommitte power system engineering committe, "Terms, definitions and symbols for subsynchronous oscillations," *IEEE Trans. Power Appar. Syst.,* vol. PAS-104, no. 6, pp. 1326–1334, Jun. 1985.

[84] P. Pourbeik, D. G. Ramey, N. Abi-Samra, D. Brooks, and A. Gaikwad, "Vulnerability of large steam turbine generators to torsional interactions during electrical grid disturbances," *IEEE Trans. Power Syst.,* vol. 22, no. 3, pp. 1250–1258, Aug. 2007.

[85] J. Shair, X. Xie, L. Wang, W. Liu, J. He, and H. Liu, "Overviewof emerging subsynchronous oscillations in practical wind power systems," *Renew. Sustain. Energy Rev.,* vol. 99, pp. 159–168, Jan. 2019.

[86] J. V. Milanovic and A. Adrees, "Identifying generators at risk of SSR in meshed compensated AC/DC power networks," *IEEE Trans. Power Syst.,* vol. 28, no. 4, pp. 4438–4447, Nov. 2013.

[87] A. Adrees and J. V. Milanovic, "Methodology for evaluation of risk of subsynchronous resonance in meshed compensated networks," *IEEE Trans. Power Syst.,* vol. 29, no. 2, pp. 815–823, Mar. 2014.

[88] P. Pourbeik, A. Boström, and B. Ray, "Modeling and application studies for a modern static VAr system installation," *IEEE Trans. Power Del.,* vol. 21, no. 1, pp. 368–377, Jan. 2006.

[89] P. Pourbeik, R. J. Koessler, D. L. Dickmander, and W.Wong, "Integration of large wind farms into utility grids (part 2 - performance issues)," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Toronto, Onttario, Canada, 2003, pp. 1–6.

[90] A. E. Leon and J. A. Solsona, "Sub-synchronous interaction damping control for DFIG wind turbines," *IEEE Trans. Power Syst.,* vol. 30, no. 1, pp. 419–428, Jan. 2015.

[91] T. Ackermann and R. Kuwahata, "Lessons learned from international wind integration studies," AEMO, Tech Rep. AEMO Wind Integration, Nov. 2011, vol. WP4(A).

[92] V. B. Virulkar and G. V. Gotmare, "Sub-synchronous resonance in series compensated wind farm: A review," *Renew. Sustain. Energy Rev.,* vol. 55, pp. 1010–1029, Mar. 2016.

[93] A. Ostadi, A. Yazdani, and R. K. Varma, "Modeling and stability analysis of a DFIG-based wind-power generator interfaced with a seriescompensated line," *IEEE Trans. Power Del.,* vol. 24, no. 3, pp. 1504–1514, Jul. 2009.

[94] L. Fan, R. Kavasseri, Z. L. Miao, and C. Zhu, "Modeling of DFIG-based wind farms for SSR analysis," *IEEE Trans. Power Del.,* vol. 25, no. 4, pp. 2073–2082, Oct. 2010.

[95] L. Fan and Z. Miao, "Nyquist-stability-criterion-based SSR explanation for type-3 wind generators," *IEEE Trans. Energy Convers.,* vol. 27, no. 3, pp. 807–809, Sep. 2012.

[96] Z. Miao, "Impedance-model-based SSR analysis for type 3wind generator and series-compensated network," *IEEE Trans. Energy Convers.,* vol. 27, no. 4, pp. 984–991, Dec. 2012.

[97] D. H. R. Suriyaarachchi, U. D. Annakkage, C. Karawita, and D. A. Jacobson, "A procedure to study sub-synchronous interactions in wind integrated power systems," *IEEE Trans. Power Syst.,* vol. 28, no. 1, pp. 377–384, Feb. 2013.

[98] H. Liu, X. Xie, C. Zhang, Y. Li, H. Liu, and Y. Hu, "Quantitative SSR analysis of series-compensated DFIG-based wind farms using aggregated RLC circuit model," *IEEE Trans. Power Syst.,* vol. 32, no. 1, pp. 474–483, Jan. 2017.

[99] X. Xie, X. Zhang, H. Liu, H. Liu, Y. Li, and C. Zhang, "Characteristic analysis of subsynchronous resonance in practical wind farms connected to series-compensated transmissions," *IEEE Trans. Energy Convers.,* vol. 32, no. 3, pp. 1117–1126, Sep. 2017.

[100] Y. Cheng, M. Sahni, D. Muthumuni, and B. Badrzadeh, "Reactance scan crossover-based approach for investigating SSCI concerns for DFIG-based wind turbines," *IEEE Trans. Power Del.,* vol. 28, no. 2, pp. 742–751, Apr. 2013.

[101] B. Benjamin, B. Johnson, P. Denholm, and B. Hodge, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power Energy Mag.,* vol. 15, no. 2, pp. 61–73, Mar./Apr. 2017.

[102] X. Wang, F. Blaabjerg, and W. Wu, "Modeling and analysis of harmonic stability in an ac power-electronic-based power system," *IEEE Trans. Power Electron.,* vol. 29, no. 12, pp. 6421–6432, Dec. 2014.

[103] J. He, Y. W. Li, D. Bosnjak, and B. Harris, "Investigation and active damping of multiple resonances in a parallel-inverter-based microgrid," *IEEE Trans. Power Electron.,* vol. 28, no. 1, pp. 234–246, Jan. 2013.

[104] X. Wang, F. Blaabjerg, M. Liserre, Z. Chen, J. He, and Y. Li, "An active damper for stabilizing power-electronic-based ac systems," *IEEE Trans. Power Electron.,* vol. 29, no. 7, pp. 3318–3329, Jul. 2014.

[105] C. Yoon, H. Bai, R. N. Beres, X. Wang, C. L. Bak, and F. Blaabjerg, "Harmonic stability assessment for multiparalleled, grid-connected inverters," *IEEE Trans. Sustain. Energy,* vol. 7, no. 4, pp. 1388–1397, Oct. 2016.

[106] E. Ebrahimzadeh, F. Blaabjerg, X. Wang, and C. L. Bak, "Modeling and identification of harmonic instability problems in wind farms," in *Proc. IEEE Energy Convers. Congr. Expo.*, Milwaukee, WI, USA, 2016, pp. 1–6.

[107] M. Liserre, F. Blaabjerg, and S. Hansen, "Design and control of an LCL -filter-based three-phase active rectifier," *IEEE Trans. Ind. Appl.,* vol. 41, no. 5, pp. 1281–1291, Sep.-Oct. 2005.

[108] E. Ebrahimzadeh, F. Blaabjerg, X. Wang, and C. L. Bak, "Harmonic stability and resonance analysis in large PMSG-based wind power plants," *IEEE Trans. Sustain. Energy,* vol. 9, no. 1, pp. 12–23, Jan. 2018.

[109] M. A. Laughton and D. J. Warne, *Electrical Engineer's Reference Book*, 16th ed. London, UK: Newnes, 2003.

[110] D. Wu, G. Li, M. Javadi, A. M. Malyscheff, M. Hong, and J. N. Jiang, "Assessing impact of renewable energy integration on system strength using site-dependent short circuit ratio," *IEEE Trans. Sustain. Energy,* vol. 9, no. 3, pp. 1072–1080, Oct. 2018.

[111] L. Fan and Z. Miao, "Wind in weak grids : 4 hz or 30 hz oscillations?," *IEEE Trans. Power Syst.,* vol. 33, no. 5, pp. 5803–5804, Sep. 2018.

[112] L. Fan, "Modeling type-4 wind in weak grids," *IEEE Trans. Sustain. Energy,* vol. 10, no. 2, pp. 853–863, Apr. 2019.

[113] D. Ramasubramanian, W. Wang, P. Pourbeik, E. Farantatos, S. Soni, and V. Chadliev, "Positive sequence voltage source converter mathematical model for use in low short circuit systems," *IET Gener. Transm. Distrib.,* vol. 14, no. 1, pp. 87–97, Jan. 2020.

[114] J. Z. Zhou, H. Ding, S. Fan, Y. Zhang, and A. M. Gole, "Impact of short circuit ratio and phase locked loop parameters on the small signal behavior of a VSC HVDC converter," *IEEE Trans. Power Del.,* vol. 29, no. 5, pp. 2287–2296, Oct. 2014.

[115] L. Papangelis, M. S. Debry, T. Prevost, P. Panciatici, and T. V. Cutsem, "Stability of a voltage source converter subject to decrease of short-circuit capacity: A case study," in *Proc. PSCC*, Dublin, Ireland, Jun. 2018.

[116] I. D. Pasiopoulou, E. O. Kontis, T. A. Papadopoulos, and G. K. Papagiannis, "Effect of load modelling on power systems stability studies," in *Proc. Conf. 12th MEDPOWER*, Nov. 2020, doi: 10.1049/ICP.2021.1266.

[117] P. Demetriou, M. Asprou, J. Quiros-Tortos, and E. Kyriakides, "Dynamic IEEE test systems for transient analysis," *IEEE Syst. J.,* vol. 11, no. 4, pp. 2108-2117, 2017.

[118] S. Wellhofer, S. Hohn, and M. Luther, "Impact of load modeling on small signal stability investigations," in *IEEE Int. Conf. POWERCON*, Wollongong, NSW, Australia, Sep. 2016, pp. 1-6, doi: 10.1109/POWERCON.2016.7753870.

[119] "Documentation on controller tests in test grid configurations," in "ENTSO-E SG SPD Report," European Network of Transmission System Operators for Electricity, 2013.

[120] H. B. Zhang, R. M. He, L. Li, and J. Zhange, "Application of different load models for the transient stability calculation," in *IEEE Int. Conf. POWERCON*, Kunming, China, Oct. 2002, vol. 4, pp. 2014-2018, doi: 10.1109/ICPST.2002.1047134.

[121] E. Z. Vaahedi, H. M. El-Din, and W. W. Price, "Dynamic load modeling in large scale stability studies," *IEEE Trans. Power Syst.,* vol. 3, no. 3, pp. 1039-1045, 1988 1988.

[122] B. Khodabakhchian and G. T. Vuong, "Modeling a mixed residential-commercial load for simulations involving large disturbances," *IEEE Trans. Power Syst.,* vol. 12, no. 2, pp. 791-796, 1997.

[123] C. Grande-Moran, B. Fernandes, D. Feltes, J. Feltes, M. Wu, and R. Wells, "Case studies on dynamic load modeling," in *Conf. Clemson University PSC*, Charleston, SC, USA, 2018, pp. 1-8.
[124] R. M. He and A. J. Germond, "Comparison of dynamic load modeling using neural network and traditional method," *IEEE,* 1993.
[125] A. Keyhani, W. Lu, and G. T. Heydt, "Composite neural network load models for power system stability analysis," in *Proc. Conf. IEEE PES PSCE*, 2004, pp. 1159-1163, doi: 10.1109/PSCE.2004.1397702.
[126] T. Hiyama, M. Tokieda, W. Hubbi, and N. Andou, "Artificial neural network based dynamic load modeling," *IEEE Trans. Power Syst.,* vol. 12, no. 4, Nov. 1997.
[127] A. L. Bettiol, A. Souza, J. L. Todesco, and J. R. Tesch, "Estimation of critical clearing times using neural networks," in *IEEE Bologna PowerTech Conference*, Bologna, Italy, Jun. 2003.
[128] C. F. Küçüktezcan, V. Murat, and s. Genç, "A comparison between ANN based methods of critical clearing time estimation," in *Int. Conf. ELECO*, Bursa, Turkey, Nov. 2013.
[129] G. Rogers, *Power System Oscillations*. New York, NY, USA: Kluwer, 2000.
[130] *DSATools Dynamic Security Assessment Software*. Accessed: Dec. 2022. [Online]. Available: http://www.dsatools.com/.
[131] T. Aziz, S. Dahal, Z. Y. Dong, M. Nadarajah, and T. Saha, "Control methodologies of distributed generation for enhanced network stability and control," The University of Queensland, St. Lucia, Brisbane, Jul. 2009.
[132] R. M. He, J. Ma, and D. J. Hill, "Composite load modeling via measurement approach," *IEEE Trans. Power Syst.,* vol. 21, no. 2, pp. 663-672, May 2006.
[133] "Enhanced load modeling: survey results on industry load modeling practices," EPRI, Palo Alto, 2020.
[134] K. Morison, H. Hamadani, and L. Wang, "Practical issues in load modeling for voltage stability studies," in *IEEE PES General Meeting*, Jul. 2003, vol. 3, pp. 1392-1397, doi: 10.1109/PES.2003.1267355.
[135] D. Han, J. Ma, R. M. He, and Z. Y. Dong, "A real application of measurement based load modeling in large-scale power grids and its validation," *IEEE Trans. Power Syst.,* vol. 24, no. 4, pp. 1756-1764, Sep. 2009, doi: 10.1109/TPWRS.2009.2030298.
[136] J. Hou, Z. Xu, and Z. Y. Dong, "Measurement-based load modeling at distribution level with complete model structure," in *IEEE PES General Meeting*, San Diego, CA, USA, Jul. 2012, pp. 1-6, doi: 10.1109/PESGM.2012.6344939.
[137] D. Y. Wu, K. T. Chau, C. H. Liu, S. Gao, and F. H. Li, "Transient stability analysis of SMES for smart grid with vehicle-to-grid operation," *IEEE Trans. Appl. Supercond,* vol. 22, no. 3, pp. 5701105-5701105, Jun. 2012, doi: 10.1109/TASC.2011.2174572.
[138] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company, INC, Sep. 1988.
[139] K. Y. Lee and M. El-Sharkawi, *Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems*. Wiley-IEEE Press, 2008.
[140] K. Kristinsson and G. A. Dumont, "System identification and control using genetic algorithms," *IEEE Trans. Syst. Man Cybern.,* vol. 22, pp. 1033-1046, Sep. 1992, doi: 10.1109/21.179842.
[141] K. Iba, "Reactive power optimisation by genetic algorithm," *IEEE Trans. Power Syst.,* vol. 9, no. 2, pp. 685-692, May 1994.
[142] M. Tabassum and K. Mathew, "A genetic algorithm analysis towards optimization solutions," *IJDIWC,* vol. 4, no. 1, pp. 124-142, Jan. 2014, doi: 10.17781/P001091.
[143] S. M. Zali and J. V. Milanović, "Modelling of distribution network cell based on grey-box approach," in *Proc. Conf. 7th MedPower*, Agia Napa, Cyprus, Nov. 2010, pp. 1-6, doi: 10.1049/CP.2010.0852.
[144] B. K. Choi, H. D. Chiang, Y. H. Li, Y. T. Chen, D. H. Huang, and M. G. Lauby, "Development of composite load models of power systems using on-line measurement data," in *IEEE PES*

General Meeting, Montreal, QC, Canada, Jun. 2006, vol. 1, no. 2, pp. 161-169, doi: 10.1109/IEEECONF10470.2006.

[145] H. D. Chiang, J. C. Wang, C. T. Huang, Y. T. Chen, and C. H. Huang, "Development of a dynamic ZIP-motor load model from on-line field measurements," *Int. J. Electr. Power Energy Syst,* vol. 19, no. 7, pp. 459-468, Oct. 1997, doi: 10.1016/S0142-0615(97)00016-1.

[146] S. Perez-Londono, L. Rodriguez-Garcia, and J. Mora-Florez, "A comparative analysis of dynamic load models for voltage stability studies," in *IEEE PES Transmission & Distribution | Conference & Exposition (T&D)*, Nov. 2014, pp. 1-6, doi: 10.1109/TDC-LA.2014.6955260.

[147] X. Y. Zheng, R. M. He, and J. Ma, "A new load model suitable for transient stability analysis with large voltage disturbances," in *Proc. Int. Conf. ICEMS*, Incheon, Korea (South), Dec. 2010, pp. 1898-1902.

[148] D. Karlsson and D. J. Hill, "Modelling and identification of nonlinear dynamic loads in power systems," *IEEE Trans. Power Syst.,* vol. 9, no. 1, pp. 157-166, Feb. 1994.

[149] I. A. Hiskens and J. V. Milanovic, "Load modelling in studies of power system damping," *IEEE Trans. Power Syst.,* vol. 10, no. 4, pp. 1781-1788, Nov. 1995, doi: 10.1109/59.476041.

[150] J. Shi and R. M. He, "Parameter identification of dynamic load model using multi-curve fitting method," *Automation of Electric Power System,* vol. 27, no. 24, pp. 18-22, Dec. 2003.

[151] *MATLAB Getting Started Guide* (no. V.7.12.0.635). Natick, Massachusetts: The MathWorks Inc., 2011.

[152] *DIgSILENT PowerFactory User Manual* (no. 15.1). Gomaringen, Germany: DIgSILENT, 2013.

[153] C. Sreevidhya, K. Mukesh, and K. Ilango, "Design and implementation of non-intrusive load monitoring using machine learning algorithm for appliance monitoring," in *Proc. IEEE Int. Conf. Intell. Tech. Control Optim. Signal Process*, Tamilnadu, India, Apr. 2019, pp. 1-6, doi: 10.1109/INCOS45849.2019.8951312.

[154] "The future of heating: meeting the challenge," Department of Energy & Climate Change, London, UK, Mar. 2013. Accessed: Dec. 2022. [Online]. Available: https://www.gov.uk/government/publications/the-future-of-heating-meeting-the-challenge

[155] C. Laughman *et al.* (Dec. 2003) Power signature analysis. *IEEE Power Energy Mag.* 56-63.

[156] Y. C. Su, K. L. Lian, and H. H. Chang, "Feature selection of non-intrusive load monitoring system using STFT and wavelet transform," in *Proc. IEEE 8th Int. Conf. ICEBE*, Beijing, China, Oct. 2011, pp. 293-298, doi: 10.1109/ICEBE.2011.49.

[157] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE Inst. Electr. Electron Eng.,* vol. 80, no. 12, pp. 1870–1891, 1992.

[158] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," in *Proc. IEEE Int. Conf. ICCE*, Xianning, China, Feb. 2011, vol. 57, pp. 76-84, doi: 10.1109/TCE.2011.5735484.

[159] H. H. Chang, K. L. Chen, Y. P. Tsai, and W. J. Lee, "A new measurement method for power signatures of nonintrusive demand monitoring and load identification," *IEEE Trans. Ind. Appl.,* vol. 48, no. 2, pp. 764-771, Mar. 2012, doi: 10.1109/TIA.2011.2180497.

[160] Y. H. Lin and M. S. Tsai, "Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification," *IEEE Trans. Instrum. Meas.,* vol. 63, no. 6, pp. 1470-1483, Jun. 2014, doi: 10.1109/TIM.2013.2289700.

[161] S. M. Tabatabaei, S. Dick, and W. Xu, "Toward non-Intrusive load monitoring via multi-label classification," *IEEE Trans. Smart Grid,* vol. 8, no. 1, pp. 26-40, Jan. 2017, doi: 10.1109/TSG.2016.2584581.

[162] A. I. Cole and A. Albicki, "Data extraction for effective non-intrusive identification of residential power loads," in *Proc. IEEE Conf. IMTC*, St. Paul, MN, USA, May 1998, vol. 2, pp. 812-815, doi: 10.1109/IMTC.1998.676838.

[163] A. I. Cole and A. Albicki, "Algorithm for nonintrusive identification of residential appliances," in *Proc. IEEE Int. Symp. Circuits Syst.*, Florence, Italy, May 1998, vol. 3, pp. 338-341, doi: 10.1109/ISCAS.1998.704019.

[164] L. K. Norford and S. B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy Build.,* vol. 24, no. 1, pp. 51-64, Jan. 1996, doi: 10.1016/0378-7788(95)00958-2.

[165] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox, "Nonintrusive load monitoring and diagnostics in power systems," *IEEE Trans. Instrum. Meas.,* vol. 57, no. 7, pp. 1445-1454, May 2008, doi: 10.1109/TIM.2008.917179.

[166] Y. Jimenez, C. Duarte, J. Petit, and G. Carrillo, "Feature extraction for nonintrusive load monitoring based on s-transform," in *Proc. Clemson University PSC*, Clemson, SC, USA, Mar. 2014, pp. 1-5, doi: 10.1109/PSC.2014.6808109.

[167] T. Y. Wang and B. Yin, "A new method for the nonintrusive load monitoring based on BP neural network," in *Proc. Conf. ICMIP*, Wuhan, China, Mar 2017, pp. 93-97, doi: 10.1109/ICMIP.2017.55.

[168] Business Requirements Work Group, "Smart metering infrastructure minimum functionality specification," NSMP Business Requirements Work Stream, Nov. 2011. Accessed: Dec. 2022. [Online]. Available: https://link.aemo.com.au/sites/wcl/smartmetering/Pages/BRWG.aspx

[169] W. C. Kong, Z. Y. Dong, B. Wang, J. H. Zhao, and J. Huang, "A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing," *IEEE Trans. Smart Grid,* vol. 11, no. 1, pp. 148-160, Jan. 2020, doi: 10.1109/TSG.2019.2918330.

[170] A. Liddy, K. Bhogale, W. F. Lim, and L. Spencer, "Ministerial advice benefits of advanced digital metering," Queensland Competition Authority, Sep. 2019. Accessed: Dec. 2022. [Online]. Available: https://www.qca.org.au/wp-content/uploads/2019/08/ministerial-advice-benefits-of-advanced-digital-metering.pdf

[171] H. S. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. W. Han, "Unsupervised disaggregation of low frequency power measurements," in *Proc. Int. Conf. SIAM Data Mining SDM*, Phoenix, Arizona, USA, Dec. 2011, pp. 747-758.

[172] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *National Conf. 26th AAAI*, Toronto, Ontario, Canada, Jul. 2012, pp. 356-362.

[173] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, "Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring," *IEEE Trans. Smart Grid,* vol. 7, no. 6, pp. 2575-2585, Nov. 2016, doi: 10.1109/TSG.2015.2494592.

[174] K. A. Agyeman, S. Y. Han, and S. H. Han, "Real-time recognition non-intrusive electrical appliance monitoring algorithm for a residential building energy management system," *Energies,* vol. 8, no. 9, pp. 9029-9048, Aug. 2015, doi: 10.3390/EN8099029.

[175] K. Suzuki, S. Inagaki, T. Suzuki, H. Nakamura, and K. Ito, "Nonintrusive appliance load monitoring based on integer programming," in *Proc. Annual Conf. SICE*, Chofu, Japan, Oct. 2008, pp. 2742-2747, doi: 10.1109/SICE.2008.4655131.

[176] M. Z. A. Bhotto, S. Makonin, and I. V. Bajic, "Load disaggregation based on aided linear integer programming," *IEEE Trans. Circuits Syst. II Express Briefs,* vol. 64, no. 7, pp. 792-796, Jul. 2017, doi: 10.1109/TCSII.2016.2603479.

[177] Z. Chen, L. Wu, and Y. Fu, "Real-time price based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Trans. Smart Grid,* vol. 3, no. 4, pp. 1822-1831, Sep. 2012, doi: 10.1109/TSG.2012.2212729.

[178] C. Peng, G. Y. Lin, S. P. Zhai, Y. Ding, and G. Y. He, "Non-intrusive load monitoring via deep learning based user model and appliance group model," *Energies,* vol. 13, no. 21, pp. 5629-5629, Oct. 2020, doi: 10.3390/EN13215629.

[179] S. Welikala, C. Dinesh, M. P. B. Ekanayake, R. I. Godaliyadda, and J. Ekanayake, "Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting," *IEEE Trans. Smart Grid,* vol. 10, no. 1, pp. 448-461, Jan. 2019, doi: 10.1109/TSG.2017.2743760.

[180] A. L. L. Aquino, H. S. Ramos, A. C. Frery, L. P. Viana, T. S. G. Cavalcante, and O. A. Rosso, "Characterization of electric load with information theory quantifiers," *Physica A,* vol. 465, pp. 277-284, Jan. 2017, doi: 10.1016/J.PHYSA.2016.08.017.

[181] C. Dinesh, B. W. Nettasinghe, R. I. Godaliyadda, M. P. B. Ekanayake, J. Ekanayake, and J. V. Wijayakulasooriya, "Residential appliance identification based on spectral information of low frequency smart meter measurements," *IEEE Trans. Smart Grid,* vol. 7, no. 6, pp. 2781-2792, Nov. 2016, doi: 10.1109/TSG.2015.2484258.

[182] Y. D. Zhang and X. Chen, "Triad state space construction for chaotic signal classification with deep learning," *arxiv.org,* Mar. 2020, Art no. arXiv:2003.11931.

[183] K. Chen, Q. Wang, Z. Y. He, K. L. Chen, J. Hu, and J. L. He, "Convolutional sequence to sequence non-intrusive load monitoring," *IET JoE,* vol. 2018, no. 17, pp. 1860–1864, Oct. 2018.

[184] C. Y. Zhang, M. J. Zhong, Z. Z. Wang, N. Goddard, and C. Sutton, "Sequence-to-point learning with neural networks for non-intrusive load monitoring," in *Conf. AAAI*, New Orleans, Louisiana, USA, 2018.

[185] D. S. Kumar, K. L. Low, A. Sharma, and W. L. Woo, "Non-intrusive load monitoring using feed forward neural network," in *Proc. IEEE ISGT Asia*, Chengdu, China, May 2019, pp. 4065-4069, doi: 10.1109/ISGT-ASIA.2019.8880801.

[186] N. V. Linh and P. Arboleya, "Deep learning application to non-intrusive load monitoring," in *Proc. IEEE PowerTech*, Milan, Italy, Jun. 2019, pp. 1-5, doi: 10.1109/PTC.2019.8810435.

[187] X. Li, Y. Guo, M. Yan, and X. Wu, "Non-intrusive load monitoring for consistent shape loads based on convolutional neural network," in *Int. Conf. CCET*, Beijing, China, Aug. 2022.

[188] S. Yaemprayoon and J. Srinonchat, "Non-intrusive load monitoring using multi-layer perceptron for appliances classification," in *Int. Conf. ECTI-CON*, Prachuap Khiri Khan, Thailand, May 2022.

[189] M. Yu, B. Wang, L. Lu, Z. Bao, and D. Qi, "Non-intrusive adaptive load identification based on siamese network," *IEEE Access,* vol. 10, pp. 11564 - 11573, Jan. 2022.

[190] Q. Huang, H. Cheng, K. Fang, W. Yu, C. Fan, and Y. Li, "Non-intrusive load monitoring-based on deep convolutional generative adversarial network prediction," in *IEEE Int. Conf. ICET*, Chengdu, China, May 2022.

[191] D. Yang, X. Gao, L. Kong, Y. Pang, and B. Zhou, "An event driven convolutional neural architecture for non-intrusive load monitoring of residential appliance," *IEEE Trans. Consumer Electronics,* vol. 66, no. 2, pp. 173-182, May 2020.

[192] M.-S. Tsai, Y.-H. Li, and Y.-K. Lin, "A novel method combining grayscale images and convolutional neural networks for non-intrusive load identifications," in *Int. Conf. iFUZZY*, Kaohsiung, Taiwan, Nov. 2022.

[193] X. Wang and W. Li, "MTFed-NILM: multi-task federated learning for non-intrusive load monitoring," in *IEEE Int. Conf. DASC/PiCom/CBDCom/CyberSciTech*, Falerna, Italy, Sep. 2022.

[194] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Int. Conf. AISTATS*, Florida, USA, 2017, vol. 54: JMLR: W&CP, pp. 1273-1282.

[195] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *Proc. Conf. IIPhDW*, Świnoujście, Poland, 2018, pp. 117–122.

[196] I. Goodfellow *et al.*, "Generative adversarial nets," *Adv. Neural Inf. Process,* vol. 27, pp. 2672-2680, Dec. 2014, doi: 10.3156/JSOFT.29.5_177_2.

[197] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. CVPR*, Las Vegas, NV, USA, Jun. 2016, pp. 2536-2544, doi: 10.1109/CVPR.2016.278.

[198] Y. Chen, Y. Wang, D. Kirschen, and B. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Trans. Power Syst.,* vol. 33, no. 3, pp. 3265 - 3275, May 2018, doi: 10.1109/PESGM40551.2019.8974096.

[199] L. Song, Y. Li, and N. Lu, "ProfileSR-GAN: A GAN based super-resolution method for generating high-resolution load profiles," *IEEE Trans. Smart Grid,* vol. 13, no. 4, pp. 3278-3289, Jul. 2022.

[200] J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Int. Conf. Embedded Systems for Energy-Efficient Built Environments*, Seoul, South Korea, Sep. 2015, pp. 55–64.

[201] K. B. Bao, K. Ibrahimov, M. Wagner, and H. Schmeck, "Enhancing neural non-intrusive load monitoring with generative adversarial networks," in *Conf. Energy Informatics*, Oldenburg, Germany, Oct. 2018, vol. 1, no. 1: Springer Open, pp. 295–302.

[202] M. Kaselimi, A. Voulodimos, E. Protopapadakis, N. Doulamis, and A. Doulamis, "EnerGAN: A generative adversarial network for energy disaggregation," in *Proc. IEEE ICASSP*, Barcelona, Spain, May 2020, pp. 1578–1582.

[203] H. L. Chen and P. Cao, "Deep learning based data augmentation and classification for limited medical data learning," in *Proc. IEEE Int. Conf. ICPICS*, Shenyang, China, 2019, pp. 300–303.

[204] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," in *IEEE 15th Symp ISBI* Washington, DC, USA, Apr. 2018, pp. 289-293, doi: 10.1109/ISBI.2018.8363576.

[205] C. Bowles *et al.*, "GAN augmentation: Augmenting training data using generative adversarial networks," *Computer Vision and Pattern Recognition,* Oct. 2018, Art no. arXiv:1810.10863.

[206] S. Y. Zhao, Z. J. Liu, J. Lin, J. Y. Zhu, and S. Han, "Differentiable augmentation for data-efficient GAN training," *Adv. Neural Inf. Process,* vol. 33, pp. 7559-7570, Jun. 2020.

[207] N. T. Tran, V. H. Tran, N. B. Nguyen, T. K. Nguyen, and N. M. Cheung, "On data augmentation for GAN training," *IEEE Trans. Image Process,* vol. 30, pp. 1882-1897, Jan. 2021, doi: 10.1109/TIP.2021.3049346.

[208] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," *Adv. Neural Inf. Process,* vol. 33, pp. 12104-12114, Jan. 2020.

[209] H. Zhang, Z. Z. Zhang, A. Odena, and H. L. Lee, "Consistency regularization for generative adversarial networks," in *Proc. Int. Conf. ICLR*, Apr. 2020.

[210] A. Bora, E. Price, and A. G. Dimakis, "AmbientGAN: Generative models from lossy measurements," in *Proc. Int. Conf. ICLR*, Vancouver, Canada, Feb. 2018.

[211] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.,* vol. 22, no. 10, pp. 1345-1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[212] R. Hao, K. Namdar, L. Liu, and F. Khalvati, "A transfer learning–based active learning framework for brain tumor classification," *Frontiers in Artificial Intelligence,* vol. 4, pp. 635766-635766, May 2021, doi: 10.3389/FRAI.2021.635766.

[213] Y. X. Wang, C. S. Wu, L. Herranz, J. v. d. Weijer, A. Gonzalez-Garcia, and B. Raducanu, "Transferring GANs: Generating images from limited data," in *Proc. European Conf. ECCV*, Munich, Germany, Sep. 2018, pp. 220-236, doi: 10.1007/978-3-030-01231-1_14.

[214] A. Noguchi and T. Harada, "Image generation from small datasets via batch statistics adaptation," *Journal of the Imaging Society of Japan,* vol. 59, no. 6, pp. 607-616, Dec. 2020, doi: 10.11370/ISJ.59.607.

[215] Y. X. Wang, A. Gonzalez-Garcia, D. Berga, L. Herranz, F. S. Khan, and J. v. d. Weijer, "MineGAN: Effective knowledge transfer From GANs to target domains with few images," in *Proc. Conf. IEEE CVF CVPR*, Seattle, WA, USA, Jun. 2020, pp. 9332-9341, doi: 10.1109/CVPR42600.2020.00935.

[216] J. Li, Y. Liu, and Q. Li, "Generative adversarial network and transfer-learning-based fault detection for rotating machinery with imbalanced data condition," *Meas. Sci. Technol.,* vol. 33, no. 4, Jan. 2022.

[217] A. M. A. Ahmed, Y. Zhang, and F. Eliassen, "Generative adversarial networks and transfer learning for non-intrusive load monitoring in smart grids," in *IEEE Int. Conf. SmartGridComm*, Tempe, AZ, USA, 2020, pp. 1-7.
[218] Naveen, "What are the benefits of Transfer Learning in Machine Learning," vol. Jun. 2023, ed, Jan. 2023.
[219] SparkCognition, "Transfer Learning & Machine Learning: How It Works, What It's Used For, and Where it's Taking Us," ed, Dec. 2018.
[220] i-king-of-ml, "Unlocking the Power of Transfer Learning: How it Can Revolutionize AI," ed: MLearning.ai, Apr. 2023.
[221] C. Dilmegani, "Transfer Learning in 2023: What It Is & How It Works," ed, Jul. 2020.
[222] E. S. Olivas, J. D. M. Guerrero, and M. M. Sober, *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques* 1ed. Information Science Reference - Imprint of: IGI Publishing, Sep. 2009.
[223] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conf. CVPR*, Honolulu, Hl, Jul. 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.
[224] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proc. Int. Conf. ICLR*, San Juan, Puerto Rico, USA, Jan. 2016.
[225] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are GANs created equal? A large-scale study," *Adv. Neural Inf. Process,* vol. 31, pp. 700-709, Jan. 2018.
[226] A. Brodzicki, M. Piekarski, D. Kucharski, J. Jaworek-Korjakowska, and M. Gorgon, "Transfer learning methods as a new approach in computer vision tasks with small datasets," *Found. Comput. Decis. Sci.,* vol. 45, no. 3, pp. 179-193, Sep. 2020, doi: 10.2478/FCDS-2020-0010.
[227] Y. Xu, Z. Y. Dong, K. Meng, R. Zhang, and K. Wong, "Real-time transient stability assessment model using extreme learning machine," *IET Gener. Transm. Distrib.,* vol. 5, no. 3, pp. 314-322, Feb. 2011, doi: 10.1049/IET-GTD.2010.0355.
[228] K. Sun, S. Likhate, V. Vittal, V. S. Kolluri, and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and decision trees," *IEEE Trans. Power Syst.,* vol. 22, no. 4, pp. 1935-1943, Oct. 2007, doi: 10.1109/TPWRS.2007.908476.
[229] M. He, V. Vittal, and J. S. Zhang, "Online dynamic security assessment with missing pmu measurements: A data mining approach," *IEEE Trans. Power Syst.,* vol. 28, no. 2, pp. 1969-1977, Mar. 2013, doi: 10.1109/TPWRS.2013.2246822.
[230] P. Zhang, F. X. Li, and N. Bhatt, "Next-generation monitoring, analysis, and control for the future smart control center," *IEEE Trans. Smart Grid,* vol. 1, no. 2, pp. 186-192, Jul. 2010, doi: 10.1109/TSG.2010.2053855.
[231] I. Genc, R. Diao, V. Vittal, S. Kolluri, and S. Mandal, "Decision tree-based preventive and corrective control applications for dynamic security enhancement in power systems," *IEEE Trans. Power Syst.,* vol. 25, no. 3, pp. 1611-1619, Jan. 2010, doi: 10.1109/TPWRS.2009.2037006.
[232] L. Liu, F. Tang, and C. L. Bak, "An accurate online dynamic security assessment scheme based on random forest," *Electric Power Energy Syst. ,* vol. 11, no. 7, pp. 1-17, Jul. 2018, doi: 10.3390/EN11071914.
[233] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.,* vol. 26, no. 3, pp. 1474-1483, Aug. 2011, doi: 10.1109/TPWRS.2010.2082575.
[234] C. W. Liu, M. C. Su, S. S. Tsay, and Y. J. Wang, "Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements," *IEEE Trans. Power Syst.,* vol. 14, no. 2, pp. 685-692, May 1999, doi: 10.1109/59.761898.

[235] S. A. Siddiqui, K. Verma, K. R. Niazi, and M. Fozdar, "Real-time monitoring of post-fault scenario for determining generator coherency and transient stability through ANN," *IEEE Trans. Ind. Appl.,* vol. 54, no. 1, pp. 685-692, Jan. 2018, doi: 10.1109/TIA.2017.2753176.

[236] G. N. Baltas, P. Mazidi, F. Fernandez, and P. Rodriguez, "Support vector machine and neural network applications in transient stability," in *Proc. IEEE Int. Conf. ICRERA*, Paris, France, Oct. 2018, pp. 14–17 doi: 10.1109/ICRERA.2018.8567024.

[237] Y. Z. Zhou *et al.*, "Online identification of transient stability and unstable generators based on deep learning," in *IEEE Conf. EI2*, Beijing, China, Oct. 2018, pp. 1–6, doi: 10.1109/EI2.2018.8582287.

[238] C. Ren and Y. Xu, "Transfer learning-based power system online dynamic security assessment: Using one model to assess many unlearned faults," *IEEE Trans. Power Syst.,* vol. 35, no. 1, pp. 821-824, Jan. 2020, doi: 10.1109/TPWRS.2019.2947781.

[239] P. W. Sauer and M. A. Pai, "Simulation of multimachine power system dynamics," in *Control and Dynamic System*, vol. 43, Part 3, E. C. T. Leondes Ed.: Academic Press, 1991.

[240] B. Stott, "Power system dynamic response calculations," *Proc IEEE Inst Electr Electron Eng,* vol. 67, pp. 219-241, 2/1/1979 1979, doi: 10.1109/PROC.1979.11233.

[241] B. Dembart, A. M. Erisman, E. G. Cate, M. A. Epton, and H. Dommel, "Power system dynamic analysis: Phase I. final report," EPRI, Jul. 1977.

[242] A. Pai, *Energy Function Analysis for Power System Stability*. Kluwer Academic Publishers, 1989.

[243] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[244] A. Gupta, G. Gurrala, and P. S. Sastry, "An online power system stability monitoring system using convolutional neural networks," *IEEE Trans. Power Syst.,* vol. 34, no. 2, pp. 864-872, Mar. 2019, doi: 10.1109/TPWRS.2018.2872505.

[245] B. D. Tan, J. Yang, X. L. Pan, J. Li, P. Y. Xie, and C. L. Zeng, "Representational learning approach for power system transient stability assessment based on convolutional neural network," *J. Eng,* vol. 13, pp. 1847-1850, Jan. 2017, doi: 10.1049/JOE.2017.0651.

[246] J. Hou, C. Xie, T. Y. Wang, Z. H. Yu, Y. Lu, and H. Y. Dai, "Power system transient stability assessment based on voltage phasor and convolution neural network," in *Proc. IEEE Int. Conf. Energy Internet*, Beijing, China, May 2018, pp. 247-251, doi: 10.1109/ICEI.2018.00052.

[247] Y. J. Wang, H. Pulgar-Painemal, and K. Sun, "Online analysis of voltage security in a microgrid using convolutional neural networks," in *IEEE PES General Meeting*, Chicago, IL, USA, Jul. 2017, pp. 1-5, doi: 10.1109/PESGM.2017.8274200.

[248] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. CVPR*, Las Vegas, NV, Jun. 2016, pp. 1–12, doi: 10.1109/CVPR.2016.90.

[249] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Training very deep networks," in *Proc. Int. Conf. NIPS*, Montreal, Canada, Dec. 2015, vol. 28, pp. 2377-2385.

[250] G. Larsson, M. Maire, and G. Shakhnarovich, "FractalNet: Ultra-deep neural networks without residuals," in *Int. Conf. ICLR*, Palais des Congrès Neptune, Toulon, France, Feb. 2017, pp. 403-410.

[251] J. Jiang and C. X. Zhai, "A two-stage approach to domain adaptation for statistical classifiers," in *Proc. Conf. CIKM*, Lisbon, Portugal, Nov. 2007, pp. 401-410, doi: 10.1145/1321440.1321498.

[252] S. J. Pan, J. T. Kwok, and Q. Yang, "Transfer learning via dimensionality reduction," in *Proc. AAAI*, Chicago, IL, USA, Jul. 2008, vol. 2, pp. 677-682.

[253] J. Blitzer, R. McDonald, and F. Pereira, "Domain adaptation with structural correspondence learning," in *Proc. Conf. EMNLP*, NSW, Australia, Jul. 2006, pp. 120-128, doi: 10.3115/1610075.1610094.

[254] H. A. Lam and Z. Y. Dong, "Transfer learning based dynamic security assessment," *IET Gener. Transm. Distrib.,* Mar. 2021, doi: DOI: 10.1049/gtd2.12181.

[255] P. W. Sauer, J. H. Chow, and M. A. Pai, "Multimachine simulation," in *Power System Dynamics and Stability with Synchrophasor Measurement and Power System Toolbox*, 2nd ed ed.: IEEE Press, 2018.
[256] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Proc. Int. Workshop Mach. Learn*, Aberdeen, UK, Jul. 1992, pp. 249-256, doi: 10.1016/B978-1-55860-247-2.50037-1.
[257] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RRelief," *Mach. Learn,* vol. 53, no. 1, pp. 23-69, Oct. 2003, doi: 10.1023/A:1025667309714.
[258] K. Q. Weinberger, F. Sha, and L. K. Saul, "Learning a kernel matrix for nonlinear dimensionality reduction," in *Proc. Int. Conf. ICML*, Banff, AB, Canada, Jul. 2004, doi: 10.1145/1015330.1015345.
[259] N. I. Wahab, A. Mohamed, and A. Hussain, "Feature selection and extraction methods for power systems transient stability assessment employing computational intelligence techniques," *Neural Process. Lett,* vol. 35, no. 1, pp. 81-102, Feb. 2012, doi: 10.1007/S11063-011-9205-X.
[260] K. P. F.R.S. (1901) On lines and planes of closest fit to systems of points in space. *PhilosMag*. 559-572.
[261] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, 1st ed. John Wiley & Sons, 2001.
[262] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Trans. Instrum. Meas.,* vol. 53, no. 6, pp. 1517-1525, Nov. 2004, doi: 10.1109/TIM.2004.834070.
[263] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice-Hall, 1998.
[264] B. X. Zhong, "A method of simplified modeling based on kernel function principal component analysis," in *Proc. Conf. CC*, Yantai, China, Jan. 2011, pp. 1525-1528.
[265] B. Schölkopf, A. J. Smola, and K. R. Müller, "Kernel principal component analysis," in *Proc. Int. Conf. ANN*, Oct. 1997, pp. 583-588, doi: 10.1007/BFB0020217.
[266] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annual Workshop, COLT*, Pittsburge, PA, Jul. 1992, pp. 144-152, doi: 10.1145/130385.130401.
[267] "IEEE 14-Bus System." https://icseg.iti.illinois.edu/ieee-14-bus-system/ (accessed Jun. 2023).
[268] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Trans. Power Appar. Syst.,* vol. PAS-98, no. 2, pp. 573-584, Mar./Apr. 1979. [Online]. Available: https://icseg.iti.illinois.edu/ieee-39-bus-system/.
[269] "IEEE 118-Bus System." https://icseg.iti.illinois.edu/files/2013/10/IEEE118.png (accessed Jun. 2023).
[270] S. Peyghami, P. Davari, M. Fotuhi-Firuzabad, and F. Blaabjerg. (2019) Standard test systems for modern power system analysis: an overview. *IEEE industrial electronics magazine*. 86-105.
[271] K. Gogoi and S. Chatterjee, "L-index based weak area identification of IEEE 118 bus system using dynamic simulation in PSS ® E," *IJEAT,* vol. 9, no. 2, Dec. 2019.

Appendix

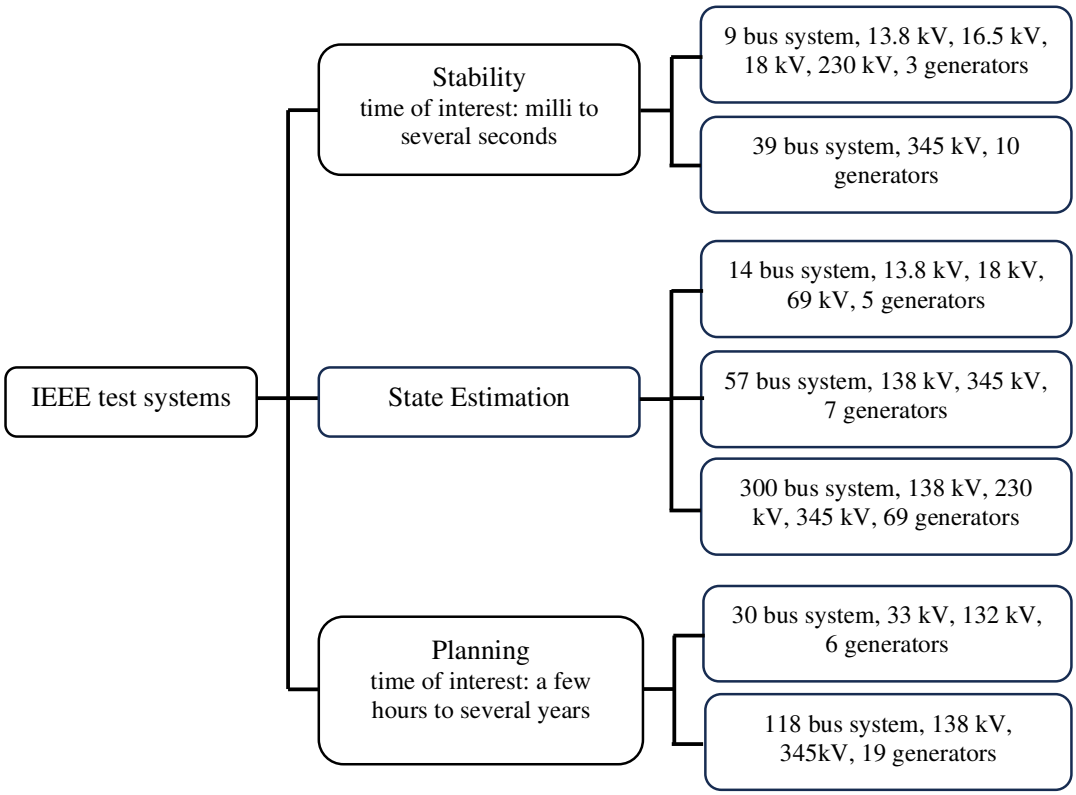

Fig. 6- 1 Most common applications for IEEE test systems [270]