UNIVERSIDAD LOYOLA ANDALUCÍA

TESIS DOCTORAL

"Dynamic Stability with Artificial Intelligence in Smart Grids"

Vol. Nº 1

Doctorando/a:	Nicholas Gregory Baltas
Directores:	Pedro Rodriguez Cortez
Tutor/a de Doctorado:	María Luisa Rodero Casano

Sevilla, 2021

UNIVERSIDAD LOYOLA ANDALUCÍA

TESIS DOCTORAL

"Dynamic Stability with Artificial Intelligence in Smart Grids"

Vol. Nº 1

Tesis defendida en Sevilla en septiembre de 2021

DOCTORADO EN CIENCIA DE LOS DATOS

Doctorando/a:	Nicholas Gregory Baltas
Directores:	Pedro Rodriguez Cortez
Tutor/a de Doctorado:	Maria Luisa Rodero Casano

Sevilla, 2021

Universidad LOYOLA

This thesis has been partially supported by the European Commission under Project FLEXITRANSTORE—H2020-LCE-2016-2017-SGS-774407, and in part by the Spanish Ministry of Science under Project ENE2017-88889-C2-1-R

To my best friend and wife Danai

and my daughter Ariadna Ourania

Abstract

Environmental concerns are among the main drives of the energy transition in power systems. Smart grids are the natural evolution of power systems to become more efficient and sustainable. This modernization coincides with the vast and wide integration of energy generation and storage systems dependent on power electronics. At the same time, the low inertia power electronics, introduce new challenges in power system dynamics. In fact, the synchronisation capabilities of power systems are threatened by the emergence of new oscillations and the displacement of conventional solutions for ensuring the stability of power systems. This necessitates an equal modernization of the methods to maintain the rotor angle stability in the future smart grids. The applications of artificial intelligence in power systems are constantly increasing. The thesis reviews the most relevant works for monitoring, predicting, and controlling the rotor angle stability of power systems and presents a novel controller for power oscillation damping.

The intelligent Power Oscillation Damper (iPOD) proposed in this thesis integrates the modelling power of artificial intelligence (superficially the Random Forest ensemble) and the ability of power converters to emulate the behaviour of a synchronous generator. The iPOD can attenuate an underdamped oscillation while adapting in the changing operating conditions. The multi-band iPOD (MiPOD) extends the controller to target more than one oscillation by emulating three rotors. In both cases, however, the main idea is that the oscillation frequencies are treated as a known parameter, hence making this implementation possible. The effectiveness and simplicity of the proposed controllers is demonstrated through a series of simulations for different types of contingencies. To demonstrate the adaptability of the iPOD, in each simulation the operating conditions vary randomly.

Following the above rationale, this thesis presents the development of a deep neural network to monitor the electromechanical interactions of a gas-turbine power plant for a system in Europe for the FLEXITRANSTORE project. The trained neural network is embedded into a controller to adaptively update the parameters of the PSS device installed in the power plant. The neural network is trained using measurements within the area on influence of the power plant to predict the oscillation frequency to compute the parameters for the phase compensation. The applications of artificial intelligence are restricted to simulation platforms; therefore, this project is a unique opportunity to obtain invaluable results about the effectiveness, limitations, performance of the application of neural networks in real life conditions.

Acknowledgments

First and foremost, I would like to express my gratitude and sincere appreciation towards my supervisor Professor Dr. Pedro Rodriguez Cortez, Loyola University, for giving me the opportunity to work on the project and for all his guidance, and support during my PhD.

Very special thanks to the members of the team Mr. Ngoc Bao Lai, Intelligent Clean Energy Systems, Luxembourg Institute of Science and Technology, Mr. Andres Tarraso and Mr. Leonardo Marin, Renewable Electrical Energy Systems, Polytechnic University of Barcelona, and Ms. Isabel Morina, Loyola University, for their help and contributions.

I am grateful to my wife Danai and daughter Ariadna Ourania who gave me strength to overcome any obstacle that came my way. I would like also to say special thanks to my mother, who believed on me from the very first day of my academic career and supported me in many ways through all these years. Lastly, but no less importantly, I thank my family and friends for their encouragement and support.

Table of Contents

List of Acronyms

AI	Artificial Intelligence
AVR	Automatic Voltage Regulator
BESS	Battery Energy Storage Systems
CART	Classification and Regression Trees
CCT	Critical Clearing Time
DER	Distributed Energy Resources
DL	Deep Learning
DM	Direct Methods
DR	Demand Response
DT	Decision Tree
EV	Electric Vehicles
FACTS	Flexible Alternating Current Transmission System
FCT	Fault Clearing Time
FOT	Fault Occurrence Time
GB	Gradient Boosting
GHG	Greenhouse Gases
GSOM	Growing Self Organising Maps
GTPP	Gas-Turbine Power Plant
iPOD	Intelligent Power Oscillation Damper
LFO	Low Frequency Oscillations
LR	Linear Regression
MA	Modal Analysis
MAE	Mean Absolute Error
MiPOD	Mulit-band intelligent Power Oscillation Damper
ML	Machine Learning
MSE	Mean Squared Error
NN	Neural Networks
ODE	Ordinary Differential Equations
PLC	Power Loop Controller
PMU	Phasor Measurement Unit

RA	Ringdown Analysis
RES	Renewable Energy Sources
RF	Random Forests
RMSE	Root Mean Squared Error
RTU	Remote Terminal Unit
SG	Smart Grids
SMIB	Single Machine Infinite Bus
SOM	Self-Organising Maps
SPC	Synchronous Power Controller
ST	Structured Text
SVM	Support Vector Machines
TDS	Time Domain Simulations
VSM	Virtual Synchronous Machine

List of Figures

List of Tables

CHAPTER 1

Introduction

Energy Transition and Smart Grids

Over the course of time, energy has been an indispensable part of the empowerment and rise of human civilization. At a great extent our society depends on energy such that demand keeps increasing. However, environmental concerns are among the major drives of the energy transition for a sustainable future. The Paris Agreement formally recognises the impact of human activity with regards to Greenhouse Gas[1] (GHG) emissions and sets a target of limiting the temperature increase at 2 degrees Celsius maximum with respect to pre-industrial levels [1]. Naturally, the electricity sector, which accounts for a high share of GHG emissions [2], is at the core of this transition.

The drastic measures needed to reduce GHG emissions leads towards the higher participation of carbon-neutral resources, such as Renewable Energy Systems (RES) and Battery Energy Storage Systems (BESS) [3]. At a greater extent, this decarbonisation is particularly appealing to countries with limited and/or costly access to fossil fuels as the exploitation of abundant sources can lead to energy independence and financial gains [4]–[8]. Yet, under the centralized operational paradigm the integration of new energy assets is challenging due the majority of RES and BESS located at the medium- and low-voltage (MV/LV) distribution networks. In fact, residential, commercial, and industrial consumers can actively contribute in markets and support mechanisms as producers or flexibility resources leading to higher efficiency and lower costs [9]–[12]. For instance, industrial consumers, which are used to deal with automatized and optimized processes, have been providing flexibility through demand response schemes typically under bilateral agreements [13]. To cope with the large number of actors and complexity, digital technologies like communications, cloud/fog/edge computing and Internet-of-Things (IoT) are vital for (i) connectivity, (ii) data collection, and (iii) advance analytics [14]. Therefore, power systems are faced with an unprecedented transformation characterized by three strongly couple pillars: Decarbonisation, Decentralization and Digitalization [15].

[1] Greenhouse gases include Carbon Dioxide ($C0_2$), methane (CH_4), nitrous oxide (N_2O) and Fluorinated gases.

Smart grids (SG) are the natural evolution of power systems to become more efficient and sustainable. Naturally, the development of SG coincides with the energy transition as the vast and diverse integration of carbon-free resources increases the complexity and uncertainty in the system. Even though there are several definitions about what constitutes a SG the underlying structure, goals and objectives are similar [16]. Specifically, SG incorporate state-of-the-art Information and Communications Technologies (ICT) on top of the electrical infrastructure to permit the dynamic interaction between the various elements/actors/entities in the grid. This cyber-physical system is based on a larger network of sensors and metering devices, automatized control, two-way communications and big data analytics [17], [18].

The operation of power systems is monitored using Supervisory Control and Data Acquisition Systems (SCADA) where the slow update rate (up to 15 seconds) add difficulties regarding real time dynamic analysis. In contrast, SG incorporate an information and data exchange network in to increase connectivity among the various elements in the system and integrate ubiquitous sensing and artificial intelligence (AI) [19]. Specifically, the development of Phasor Measurement Unit (PMU) [20], [21] and Wide Area Monitoring Systems (WAMS) [22] has led to new approaches for monitoring and controlling the power systems. Commercial PMU's have a sampling rate of 60 samples per second, ideal for dynamic analysis [23]. To illustrate the high resolution that PMU measurements offer, [24] states that for a single day and for a single parameter, a PMU can generate over 2.5 million samples. Furthermore, [25] shows that a network, located in central China with 155 PMUs installed in different parts of the grid, can process almost 300 GB of data per day. Besides PMUs, other measurement devices such as advanced meter reading (AMR) and intelligent electronic devices (IED) allow the use of artificial intelligence to improve system reliability, efficiency and sustainability [26].

Dynamic stability in future smart grids

Generally, *Power system stability* is defined as the capability of a power system (a) to operate in an equilibrium state and (b) to reach a stable equilibrium point when subjected to a disturbance [27]. However, because of the complexity of the problem, stability has been traditionally categorised into *Rotor Angle Stability*, *Voltage Stability* and *Frequency Stability*. Since the first classification of stability [28], power systems have experienced a wide transformation. The decarbonisation of generation along with the use of energy storage introduces new challenges. For example, the stochastic nature of atmospheric conditions affects the power output of RES-based

power plants. The varying patterns of wind speed and nebulosity cause fluctuations in the power output of wind and solar power plants creating power imbalances, which result in frequency deviations [29] [30]. Coupling RES and ESS can mitigate that variability [31]. However, as power electronics are required for the connection between RES/ESS and the grid, in HVDC lines, loads (such as EV) and FACTS devices, their presence is expected to rise resulting in complex intra- and inter-connection synergies [32]. For instance, due to the lack of mechanical parts, the overall system inertia will be significantly implying that during a disturbance the system will respond much faster [33].

Evidently, the dynamics and stability of power systems is affected dramatically by the increased presence of power electronics. The significance of this impact is depicted by the recently revised version of the stability problem in power systems, shown in Figure 1 where the *Resonance Stability* and *Converter-driven Stability* have been added [34]. The former refers to the torsional and electrical resonances while the latter to the fast and slow interactions between the control systems of power electronics.

In this thesis the focus of the work is on *Rotor Angle Stability*, which will become even more relevant with the integration of power electronics. In fact, ENTSO-E has identified the angular stability of power systems as one of the key challenges related to High Penetration of Power Electronic Interfaced Power Sources [35]. It is expected that in future SG (a) new power oscillations will emerge, (b) damping of existing modes will decrease and (c) transient stability margins will be reduced [34], [36].

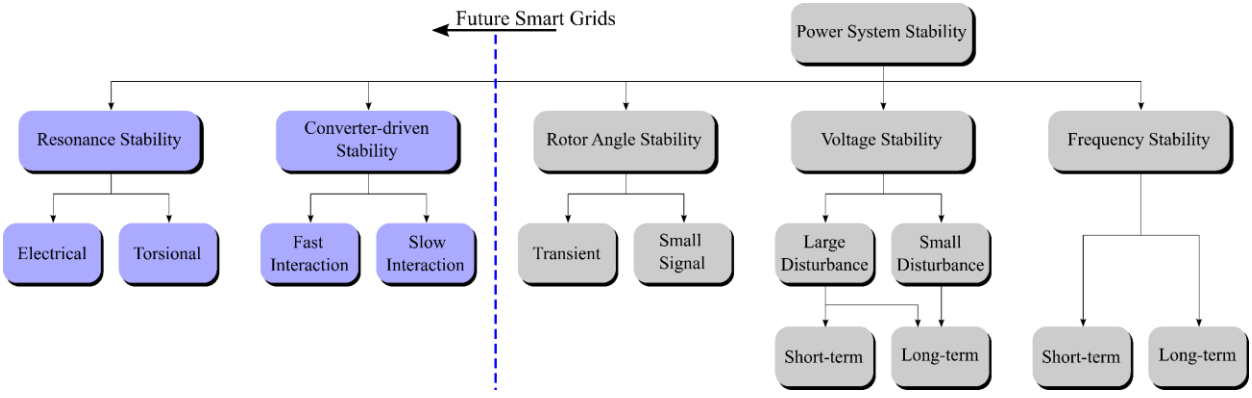


Figure 1: Power System Stability Classification

Overview of Conventional Methods in Rotor Angle Stability

Rotor Angle Stability is defined as the ability of a power system to remain in synchronism after a disturbance [27]. *Small Signal Stability* analyses the synchronism of the system under mild disturbances, such as load variations. Predominately, instability of this form is expressed as oscillations with higher amplitudes due to insufficient damping torque. In contrast, *Transient Stability* refers to the synchronization capabilities under severe disturbances, such as short circuits. Lack of synchronizing torque results in aperiodic divergence from equilibrium. After the disturbance, the time frame of interest is typically lower for *Transient Stability* than for *Small Signal Stability*. De-synchronization of the system may lead to unintentional separation of parts of the grid (i.e. islanding), cascading events and even widespread blackouts [37]. Such an event is the blackout in the USA and Canada, during the summer of 2003, affecting 50 million people and incurring financial losses estimated at 7 billion dollars [38]. Besides the particular sequence of events prior to the blackout, insufficient situational awareness lessen the operators' ability to detect the contingency and prevent its grave effects [39]. Similar cases have been recorded in Europe [40], [41] and Asia [42].

Until recently, there were two dominant approaches to determine the *Transient Stability* of the system: Time Domain Simulations (TDS) and Direct (or analytical) Methods (DM). In *Small Signal Stability*, Modal Analysis (MA) has been the most reliable and accurate method used by system operators, while Ringdown Analysis (RA) is based on signal processing techniques developed for the same purpose. In this section, the above conventional methods are briefly described for the readers' convenience.

Time Domain Simulations

In TDS, the non-linear ordinary differential equations (ODE) describing the power system, are solved using numerical integration techniques with respect to time: (1.1) and (1.2) where $\boldsymbol{x} \in \mathbb{R}^{n_x}$, $\boldsymbol{y} \in \mathbb{R}^{n_y}$ are the state and algebraic variables respectively, whereas $f(\cdot)$ and $g(\cdot)$ are continuous functions. Conclusions can be made about the stability of the system by evaluating the evolving trajectories. This method is very accurate and flexible in terms of model design and complexity. TDS is widely adopted in off-line applications, where a predefined list of credible operating and contingency scenarios is studied thoroughly through simulations to create the operating guidelines [43].

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{y}) \tag{1.1}$$

$$\mathbf{0} = g(\mathbf{x}, \mathbf{y}) \tag{1.2}$$

Direct Methods

The foundation of DM is based on Lyapunov's stability theorem [27], which can estimate the transient stability of a system without solving the nonlinear equations. DM does not replace TDS; instead both methods complement each other to provide stability estimates faster [44]. Generally, DM approaches determine stability by comparing the energy of the system at fault clearance with the critical energy of the system. The system is stable if the energy of the system is less than the critical energy [44], [45]. Formally, the energy is defined as the sum of the kinetic and potential energy stored in the system and is calculated using Lyapunov's transient energy function, as in (1.3) where V_{KE} is the kinetic energy defined in (1.4) while V_{PE}^{mech}, V_{PE}^{elec} and V_{PE}^{damp} are the potential energy components defined in equations (1.5) - (1.7) respectively, where M is generator inertia, ω is generator speed, $P_{m,i}$ is generator mechanical power, $P_{e,i}$ is generator electrical power, D_i is generator damping coefficient, δ_i is the generator rotor angle with $i = 1, \dots, n$ is the generator index.

$$\mathrm{V} = \mathrm{V_{KE}} - \mathrm{V_{PE}^{mech}} + \mathrm{V_{PE}^{elec}} + \mathrm{V_{PE}^{damp}} \tag{1.3}$$

$$\mathrm{V_{KE}} = \frac{1}{2}\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{M_i}\,\omega_\mathrm{i}^2 \tag{1.4}$$

$$\mathrm{V_{PE}^{mech}} = \frac{1}{2}\sum_{\mathrm{i=1}}^{\mathrm{n}} \int \mathrm{P_{m,i}}\,\mathrm{d}\delta_\mathrm{i} \tag{1.5}$$

$$\mathrm{V_{PE}^{elec}} = \frac{1}{2}\sum_{\mathrm{i=1}}^{\mathrm{n}} \int \mathrm{P_{e,i}}\,\mathrm{d}\delta_\mathrm{i} \tag{1.6}$$

$$\mathrm{V_{PE}^{damp}} = \frac{1}{2}\sum_{\mathrm{i=1}}^{\mathrm{n}} \int \mathrm{D_i}\,\omega_\mathrm{i}\,\mathrm{d}\delta_\mathrm{i} \tag{1.7}$$

DM based stability analysis procedure can be broadly summarized into two fundamental steps. First, TDS is implemented up to the point of fault clearance to retrieve the system state at that point. Second, the critical energy (V_{cr}) and current system energy (V_{en}) is calculated. Several methods have been developed for calculating V_{cr}. One such method, the boundary controlling unstable equilibrium (BCU) has been widely adopted by system operators [46].

Modal Analysis

The assumption that a system is linear within a region near an equilibrium point is the basis of modal analysis. Conversely, the oscillatory response of a system after an event can be represented by a number of superimposed response signals [43], as in (1.8) where (σ, ω) are the modal parameters representing damping and frequency respectively, A_0 denotes amplitude at $t = 0$ while θ is the phase of the signal $s(t)$.

$$s(t) = A_0 e^{(-at)} \cos(bt + \theta) \tag{1.8}$$

Further, the ODE describing a system can be transformed into their state space representation. Specifically, (1.1) and (1.2) can be written as in (1.9) and (1.10) respectively, where $\mathbf{A} \in \mathbb{R}^{n_x \times n_x}$ is the state matrix, $\mathbf{B} \in \mathbb{R}^{n_x \times n_u}$ is the input matrix, $\mathbf{C} \in \mathbb{R}^{n_y \times n_x}$ is the output matrix, $\mathbf{D} \in \mathbb{R}^{n_y \times n_u}$ is the feedthrough matrix and $\mathbf{u} \in \mathbb{R}^{n_u}$ is the input vector. An eigen-decomposition of matrix $\mathbf{A}$ returns the full list of eigenvalues $\lambda = a + jb$ along with their corresponding left and right eigenvectors. Modal analysis is highly accurate, able to provide all the details of the possible oscillatory response of the system.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{1.9}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{1.10}$$

Ringdown Analysis

In reality, the model of the system can be either too complex or not fully known, making MA infeasible [47]. As an alternative, Ringdown Analysis (RA) focus on the multi-modal decomposition of a signal because of a contingency e.g. a 3-phase fault. Prony analysis a typical method for determining the modes that constitute the measured signal. For instance, for a signal $\hat{s}$ discretized with a constant time step Δt as in (1.11), where $z_i = e^{\lambda_i \Delta t}$, B_i are the residues and n is the number of modes to be extracted. By finding the roots z_i of the characteristic polynomial described by (1.12), it is possible to calculate the frequencies and amplitudes comprising the

measured signal. The coefficients $a_1, \ldots, a_{n-1}$ can be computed by a least square fit using the vector of the discretized signal. Other RA methods are the Matrix-Pencil and Eigensystem Realization Algorithm [48].

$$\hat{s}(\mathrm{k}) = \sum_{i=1}^{n} B_i \, z_i^k \tag{1.11}$$

$$z^n - (a_1 z^{n-1} + \cdots + a_{n-1} z^0) = 0 \tag{1.12}$$

Limitations of Conventional Methods

Despite its accuracy, TDS requires high processing power for on-line applications, which can become impractical and expensive. In addition, current ODE solvers are time consuming rendering TDS unsuitable for on-line applications [45]. As an alternative, the AC emulation approach developed in [49] showed that simulation of power systems can be faster than real time although model limitations still exist. A trade-off between the model complexity and necessary processing resources exists, indicating that it is vital to determine the optimal balance among them [50].

Overall, after removing a large part of numerical integrations, the use of DM for stability assessment can be faster than and provide insights into the degree of stability. Nevertheless, DM depends on the existence of an UEP, whose calculation is considered to be a NP-hard problem. Moreover, deriving a transient energy function is not always possible; hence model simplifications are necessary [51]. Signal processing methods such as Prony Analysis, although able to process multiple signals, are prone to noise and non-linearities, possibly hindering their ability to identify modal properties [43].

Regarding the effects of power electronics in power systems, most research efforts are concerned with *Small Signal Stability* analysis by linearizing the system around an operation point and evaluating its eigenvalue trajectories [27]. Although this technique, as well as others suitable for linear systems such as Nyquist or Routh-Hurwitz analyses, has been used in small power systems (microgrids) with high penetration of grid-connected power converters [52], they cannot examine the system response under large disturbances. Specifically, during a severe contingency, the operating point of the non-linear system changes and the state-space model used for the *Small Signal Stability* is no longer valid. Therefore, large signal analysis i.e. *Transient Stability*, which is mainly focused on analysing the energy state of the system during the disturbance and after its clearance, should be used.

Furthermore, recent studies that evaluate the transient behaviour of power electronic devices use the equal area criterion approach [53], [54], [55], which assumes a single machine infinite bus system (SMIB) structure. Assuming a complex system such as a SMIB, the results of the transient stability analysis are limited to a single interaction i.e. between a power electronic interface device and an infinite bus. This is sub-optimal considering the increasing number of power electronics in the future smart grids, where interactions between different devices affect stability. A summary of the mature methods used in *Rotor Angle Stability* along with their advantages and limitations can be found in Figure 2.

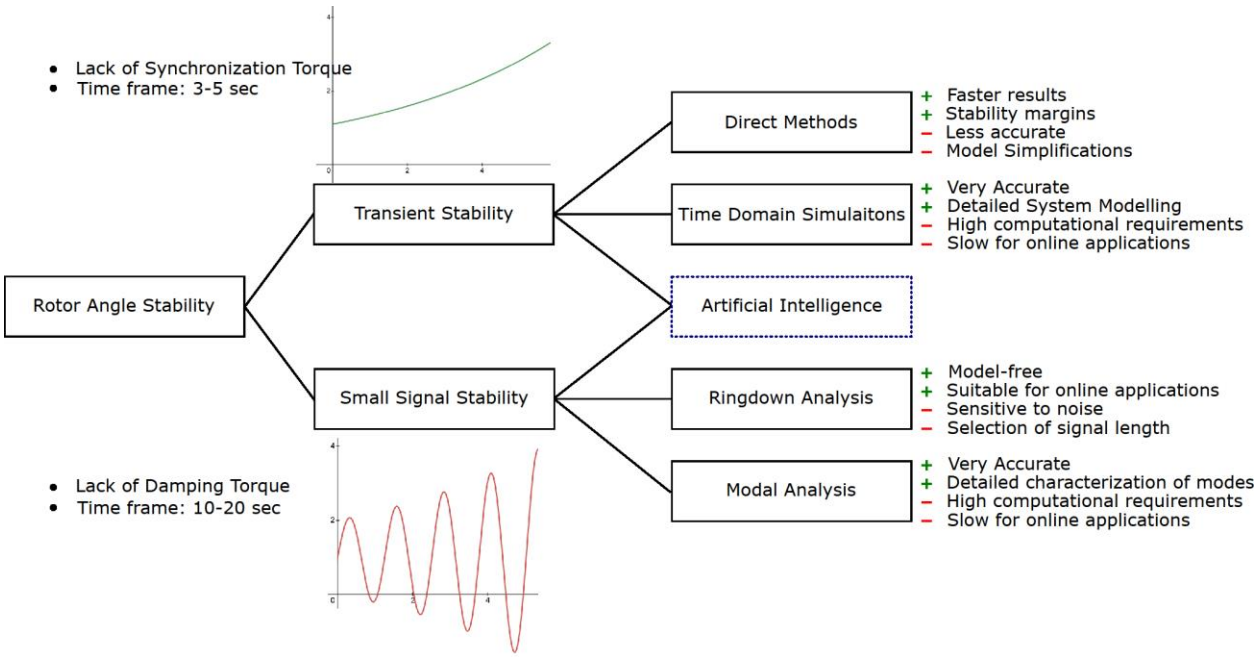


Figure 2: Summary of Rotor Angle Stability problem

Objectives

Evidently the conventional assessment methods have significant limitations with respect to the transition of power systems into SG, where the complexity and dynamics will be much harder to model. The AI is an alternative that can complement existing methods to improve the stability of the system. Driven by the high penetration of power electronics, RES, BESS, EV and in conjunction with the expectation that the *Rotor Angle Stability* of future SG will play an important role, this thesis proposes novel methods for ensuring the synchronism of the system under large or small disturbances. Specifically, this thesis examines the effectiveness of AI to model the oscillatory characteristics of the system, it integrates AI predictions in the control loop of power converters, which will dominated future power systems and describes the implementation of an AI

model in an actual power plant for providing adaptive PSS tuning to support with the rotor angle stability of the system. The latter is particularly significant as there a very few actual demonstrators of AI applications in power system stability.

Outline

The rest of the thesis is structured as follows. In Chapter 2 the review of relevant papers is presented. The focus is on the applications of AI in power systems and rotor angle stability. Chapter 2 is classified into monitoring, prediction, and control according to the theme of the review of the papers. In Chapter 3 the methods used in this thesis are thoroughly described. Specifically, form the methods of learning to the general model development framework and finally towards a description of ML type that is used throughout this thesis. Chapter 4 presents the development of the prediction models for the two-area system following the development framework presented in Chapter 3. The intelligent Power Oscillation Damper is presented in Chapter 5 for a single and multiple mode attenuation using the prediction models to track the characteristics of the oscillatory modes in the system. Chapter 6 presents the development of prediction models for an actual power system in Europe for the adaptive tuning of an actual PSS installed in a Gas-Turbine power plant. The work described therein is part of the work conducted for the FLEXITRANSTORE H2020 project. Finally, in Chapter 7 the conclusions of this thesis are summarized along with the indications for future work.

CHAPTER 2

Artificial intelligence Applications

The sources of energy data are growing and expanding, introducing the big data concept into power systems. As a result, research studies focus on the adoption of novel techniques based on AI to enhance efficiency and reliability. For instance, extensive surveys summarize the machine learning applications applied in SG [56]–[58] and in power electronics [59]. In this section relevant papers to monitoring, prediction and control are presented. Monitoring refers to algorithms that receive as inputs system variables at time t_k and provide outputs regarding the status of the system for the same time instance (see Figure 3), as in (2.1) and (2.2) where $\mathbf{x}(\mathrm{t_k})$ is the input vector at time $\mathrm{t_k}$, $f(\cdot)$ is the mapping function, such as a NN, and $\hat{y}(t_k)$ is the output of the AI model. In other words, system measurements drawn from time $\mathrm{t_k}$ are used to provide information regarding the system status for the same time instance $\mathrm{t_k}$. In contrast, prediction related applications receive similar inputs and provide outputs for time t_c. Finally, control applications represent algorithms that directly or indirectly aim to improve control actions of power systems.

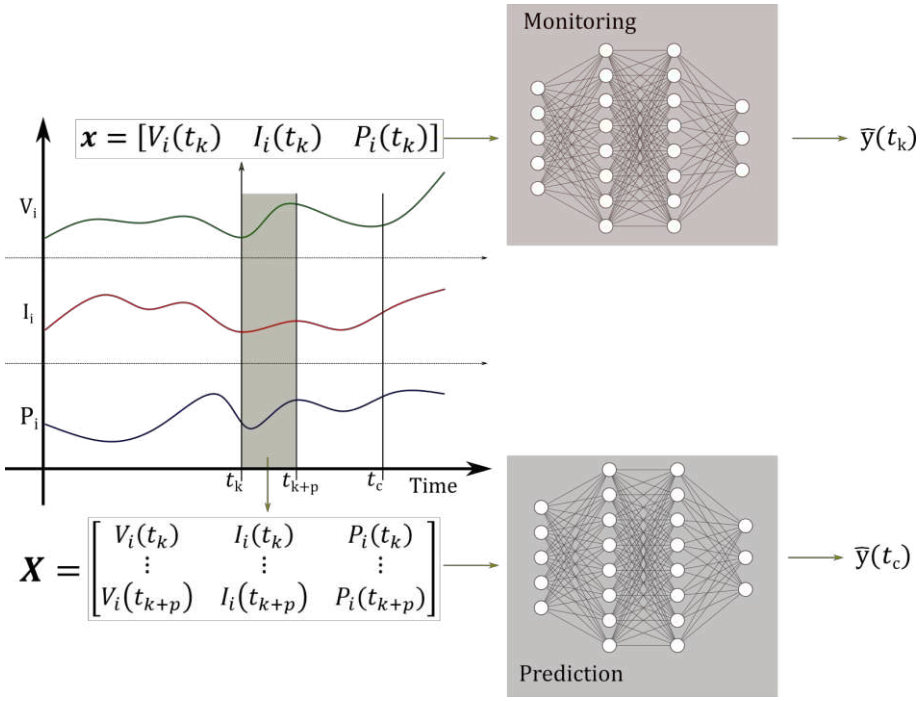


Figure 3: Example of monitoring and prediction AI models

$$\hat{y}(\mathrm{t_k}) = f\big(\mathbf{x}(\mathrm{t_k})\big), \qquad k \in \mathbb{Z}^+ \tag{2.1}$$

$$\hat{y}(\mathrm{t_c}) = f\big(\mathbf{x}(\mathrm{t_k})\big), \qquad c > k \in \mathbb{Z}^+ \tag{2.2}$$

Monitoring

CCT based Stability Assessment

The critical clearing time (CCT) is an indicator of the system strength against unplanned contingencies. It is defined as the maximum allowable fault duration for the given system conditions. Faults with a duration longer than the CCT will cause the system to lose synchronism due to the kinetic energy accumulation in the rotating masses of generators [27], [43] . Early studies of AI applications in power system stability train ML models to classify operating points as CCT stable/unstable. Specifically, in [60] a general framework is proposed, based on the CART algorithm, to develop binary DTs suitable for binomial and multivariate classification. The input-output pairs consist of simulation results of a single three phase fault for 201 operating points, which are labelled as secure/insecure according to their corresponding CCT. Following the above rationale, [61] implements a kernel ridge regression to derive the stability boundary of the system, as defined by the CCT of the current operating conditions. In other words, the CCT is directly estimated and compared to the fault clearing time (FCT), although only for a single contingency event. Similarly, in [62] a lasso regression is adopted to estimate the stability boundary. Yet in this study, the FCT is a tunable parameter representing a threshold value that can be modified according to the severity of the event under evaluation. This approach is further expanded in [63] to improve accuracy and reduce complexity; Lasso regression is used to eliminate redundant features and a univariate regression function is developed for each of the remaining features.

A type-2 fuzzy neural network (NN) is adopted in [64] to estimate CCT of four fault scenarios. Initially, an NN is used to construct the input feature vector. The fuzzy NN, first assigns each feature to a fuzzy set with an associated degree of membership and second the NN makes use of the fuzzified features to provide CCT estimations. Furthermore, a two-stage hybrid method is proposed by [65], consisting of an ensemble weighted Support Vector Machine (SVM) and an adaptive neuro fuzzy inference system (ANFIS). The hybrid method can achieve 98% classification accuracy under different contingencies and network topologies. However, the model

has limited capability in generalizing to new/unseen events. This is because the model is trained with CCT of a fixed number of fault scenarios.

Estimation of Stability Margins

Considering the gap on applications of AI in stability margin estimation, in [66] a Regression Tree (RT) is developed to assess the oscillatory and voltage stability margins of a system. The first two statistical moments i.e. mean and variance respectively, of the training samples of each RT terminal leaf were used to determine these stability margins. Authors in [67] argue that by estimating the value of the Transient Energy Margin (TEM) index, it is possible to determine the stability status and margin of the system. This approach is closely related to the CCT and critical energy of the system. Therefore, estimating the aforementioned index can provide additional information about the status of the system [27]. Subsequently, one NN is developed for each of the two events to estimate their TEM. From the results, it appears that accuracy depends on the type of the event, which the NN is trained for. In particular, the root mean square error (RMSE) between the two cases differs by a factor of 10. To mitigate this sensitivity, authors expand on their approach in [68] where an analysis is conducted with regards to feature selection and accuracy of estimation.

Similarly, [69] develops two RTs for the evaluation of the operating conditions of the Salt River Project system. An RT is developed to classify voltage violations from voltage magnitude deviations whereas another RT estimates the line thermal limits violations from deviations of current flows. Critical contingencies and operating conditions are identified, and for the less probable scenarios, N-2, N-3 and N-4, two additional CTs were trained to characterize them as secure or insecure in terms of transient and voltage stability.

The TEM index is used as the stability index of a system in order to develop four fuzzy sets in [70]**.** Several heuristic and direct system variables are ranked according to their correlation with the stability index. Concretely, the developed fuzzy model classifies the samples into the following four fuzzy sets: danger, high-risk, low-risk and safe. This approach provides additional details about the status of the system as opposed to the usual stable-unstable classification.

Likewise, [71] defines three classes of system stability: unstable, critical and stable. That distinction is derived from the relationship of FCT and stability of the system. Specifically, FCT between 5 and 7 cycles that result in instability are characterized as critical. The k-nearest neighbours (KNN) algorithm based on the Euclidean distance is implemented to classify the pre-

fault data into the classes. The model is evaluated for three contingency scenarios and the maximum and minimum misclassification errors are 2.54% and 0.42%, respectively.

Delivering or exchanging power between areas located geographically far apart requires long transmission lines. Yet, long lines are characterized by a reduced damping of oscillatory modes. As a result, operators enforce strict constraints on the maximum capacity of these lines. Several attempts have been made to correlate system variables with the damping of oscillatory modes using statistical and computational intelligence methods. Usually, regression approaches aim to directly estimate the modes of the system. For instance, in [72] a multiple regression model is proposed for estimating the damping ratio of the dominant mode in the Manitoba transmission system, where AC and DC long transmission lines interconnect two major regions. Due to the high length of those tie-lines, damping of oscillatory modes is of great importance. Likewise, in [73] operating points are classified according to the damping of a 0.35 Hz mode into low or well damped using a NN.

False Dismissals Mitigation

In the context of power systems stability analysis, False Dismissals (FD) are the type of error where inputs are wrongly classified as Stable. Such an outcome can put power systems in a critical condition or even lead to instability. In this regard, [74] defines a "*High Risk*" region for encompassing samples close to the decision boundary of a polynomial SVM. Points within that region are left unclassified while a warning informs the system operator about the critical state of the system. Concretely, the possibility of unstable cases being classified as stable (i.e. false dismissals) is minimized. The error rate is close to 4.8%, albeit this result is directly related to the size of the "*High Risk*". It is apparent, therefore, that a trade-off between false dismissals and SVM accuracy is present.

Following the above rationale, [75] defines the area close to the stability boundary as unreliable region. However, the points that lie within that area are re-evaluated instead of being ignored. Specifically, first, a Gaussian SVM evaluates the stability of input samples and checks if they belong to the incredible region. A second group of Gaussian SVM re-evaluates the samples that are located within the incredible region. Equally, [76] proposes the use of two separate SVM specifically designed so that both hyperplanes will constitute the boundaries of a grey zone. The difference between the two SVMs lies in the constraints of each algorithm that ensures purity of samples in one side of the hyperplane. If the outputs of the two SVMs do not agree, the instance is neither stable nor unstable and a warning flag is raised.

Further, in [77] a bagging ensemble of 30 Gaussian-SVMs is developed. The final output is designed based on a unanimous scheme as follows:

- A new sample is stable when the output of all individual Gaussian-SVMs is stable.
- A new sample is unstable when the output of at least one Gaussian-SVM is unstable.

By designing the ensemble in such a way, the possibility of a false dismissal is reduced. Additionally, a confidence index is proposed that calculates the distance between the hyperplane and sample. Low confidence outputs are flagged as unreliable and therefore not considered in the unanimous scheme presented above. Two tunable parameters discriminate between the reliable and unreliable results for stable and unstable classes. However, the selection of the threshold values is significant as it can lead to many cases being labelled as unreliable.

Generally, by design systems demonstrate resiliency against some disturbances. Therefore, databases contain more information about stable cases than unstable cases [78]. This imbalance will cause ML models to be biased towards the majority class i.e. resulting in an increased number of FD.

According to [79], a single DT cannot achieve high accuracy and generalization when faced with imbalanced datasets; hence, an ensemble model is adopted to classify the transient security status of a system. To defend the above statement, the authors develop a single DT and an RF and compare them over the same modified dataset, which consisted of 25% unstable cases and 75% stable cases. To counter this imbalance, usually one of the following options are used: oversampling and under-sampling. An oversampling approach is followed here that extends the original dataset by replicating the minority (unstable) samples three times. Ultimately, the RF achieves 99% accuracy, almost 10% higher than the single DT, proving the initial hypothesis in [79].

The class imbalance in power systems can be reduced more efficiently by using the Synthetic Minority Oversampling Technique (SMOTE) [80]. SMOTE creates synthetic samples that are similar but not identical with the samples in the minority class. In contrast, [81] develops a weighted RF that focuses on the correct classification of the minority class. Weights of 3 and 1 are assigned to the unstable and stable samples respectively. A feature reduction method based on the out-of-bag[2] error is used to determine the optimal feature vector. Only 45 out of the total 263

[2] Out-Of-Bag are samples left out in the training process.

features are used for the RF model, resulting in an improved accuracy of approximately 99% and in a 4% reduction on the classification error for the unstable cases.

Following the above, [82] attempts to implement a robust DT using information from the day-ahead market in order to formulate a database that can incorporate the intra-day variations of demand. Historical data of operating conditions, contingencies and data retrieved from several TDS are added to enrich the dataset. However, even with a large dataset, the authors argue that periodical tests are needed to assess the performance of the DT to maintain acceptable results.

The DT developed in [83], has more than two outgoing branches at its root node. This approach classifies the input variables as secure (or insecure) by first assessing the type of the contingency at the root node. In other words, it separates the data according to a predefined list that describes the type of topology and contingency. The rest of the DT has the same structure as the binary trees discussed thus far. For a specific network topology, the accuracy of the proposed algorithm reaches 98.6%. Because the foundation of this DT is a predefined list, it is unlikely that we could know beforehand which contingency will occur; hence, the proposed model is more suitable to be used as an indicator for preventive actions with visual aids.

Alternatively, [84] suggests encoding the system variables into an RGB image. An ensemble of CNN processes the input images to classify them as stable, unstable, and uncertain. Once classified, the images of the encoded voltages are used to estimate the stability margin for a predefined list of contingencies. The uncertain cases are included with the stable cases and unstable cases to create the datasets $D_{stable}^{uncertain}$ and $D_{unstable}^{uncertain}$. For each dataset a CNN model is developed to estimate both margins. The outputs from the two networks provide a robust estimation for the uncertain cases. In [85] a variation of SVM called the Core Vector Machine[3] (CVM) algorithm is implemented for the classification of transient stability system status. To emphasize the specific attributes of CVM, besides accuracy, false alarm and false dismissal, the CPU time and number of support vectors are also considered as performance metrics. Overall, CVM appears to perform better than SVM both in FD mitigation but also CPU time. Training time is considered as an important factor for online stability assessment. A variation called the TWin-SVM[4] (TWSVM)

[3] CVM is based on computational geometry and the minimum enclosing ball problem. Generally, it requires less resources both in time and space when dealing with big data. Briefly, an CVM aims to discover a hyper-sphere with centre $\boldsymbol{c}$ and radius $\boldsymbol{r}$ that encircles a set of points [189].

[4] The TWSVM algorithm aims to construct as many, non-parallel, hyperplanes as classes using the patterns of each class only, resulting in faster training time [190].

and the CNN algorithm is adopted by [86]. Practically, the objective function of the TWSVM is integrated in each CNN to improve speed and generalization.

Coherency Identification

The effectiveness of a LS-SVM model is validated in [87] where a TSA system is developed to predict generator coherency groups and stability. Three coherent groups are defined according to the values of the Transient Stability Index (TSI). A sensitivity analysis showed that the largest span of the TSI can be achieved when the maximum stable angle deviation is 120 degrees. By choosing the maximum angular deviation for the highest range of TSI, the first group consists of highly critical generators i.e. TSI >1, the second group consists of advanced generators i.e. 1 >TSI >0.75 and the last group consists of non-critical generators i.e. TSI >0.75. For the first and last group the authors have associated preventive actions, which refer to the rescheduling of active power generation. A particle swarm optimization is used to optimally select the features among all the active and reactive power of generator and load buses.

Likewise, the author of the thesis presents in [88] a coherency analysis method based on SOM. Therein, rotor angle signals are fed onto the SOM algorithm to identify the similarity between the oscillating response of the generators. Due to the 2-D graphical representation of clustering by SOM, this approach can be used as a situational awareness tool for real time coherency identification. The same authors present an extension of this approach in [89] using the Growing-SOM algorithm, which does not require the specification of the size of the SOM grid but expands according to a tunable parameter and the complexity of the data.

Detection of Islanding

Recently, the integration of DG into many parts of the grid has led to the development of intelligent methods for islanding detection. For instance, in [90], [91] a DT based approach is adopted for fast detection of islands through transient signal pattern recognition. These signals consist of voltage and current that are pre-processed by using discrete wavelet transformation to formulate the inputs of the DT. A database of 800 islanding and non-islanding cases are generated for the training of the DT. Overall accuracy of islanding detection is over 99% with less than 100ms of relay response after event detection.

Furthermore, in [92], the authors argue that conventional methods possess a high non-detection zone and that DT based relays are able to reduce it. In fact, the proposed DT-based relay was able to achieve higher detection rates while minimizing the cases where islanding is not detected.

Interestingly, the variables that affect the decision of the DT are RoCoF[5], frequency deviation, active and reactive power.

Further, in [93], features based on voltage, frequency and rotor angle over a sliding window of size $k\Delta t$ are used along with an SVM to predict islanding and trigger DG protection relay's. Approximately 3000 scenarios are simulated to generate the appropriate dataset size for training and testing the proposed algorithm. The SVM has been tested under different sizes of $k\Delta t$ and kernel functions. A 3rd order polynomial kernel yields the highest accuracy with an average detection speed for each scenario (less than 160 ms).

Prediction

Fixed Window Predictions

One of the first attempts to adopt NN for transient stability has been made by [94]. Specifically, a predictive tool is developed using rotor angle measurements from each generator after 3 and 6 cycles from fault occurrence time (FOT), as shown in Figure 4. These generators are divided into N subgroups where each subgroup consists of a unique pair of generators. Subsequently, N number of NNs are trained to map the relation between generators and their angular deviation after a contingency. The proposed scheme can predict the stability of the system and provide details about which pair of generators will cause the de-synchronization of the system. Furthermore, an improvement of the above approach is presented in [95], where generators are evaluated w.r.t. to Center of Inertia (COI) instead of other generators. This allows to detect which of the generator(s) will become unstable.

Alternatively, [96] proposes a multivariate boosting C5.0 ensemble to predict the out-of-step generators (rather than stability status). Using TDS, it is possible to identify 14 patterns (classes) of the generators' behaviour. In a comparison with single DT and SVM, it is shown that the boosting C5.0 ensemble achieves the highest accuracy, close to 91% within a fixed 30 cycle observation window as opposed to the other models achieving accuracy of 90% within a 60 cycle observation window.

[5] Rate of Change of Frequency

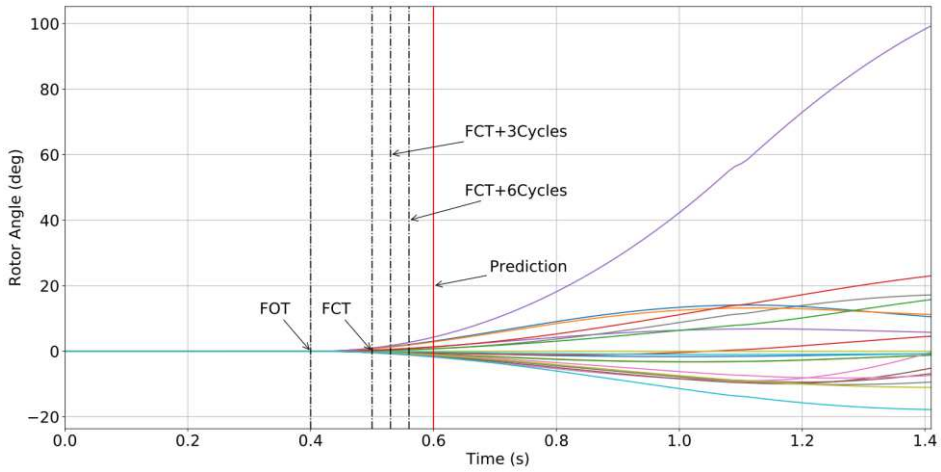


Figure 4: Timeframe of AI based Transient Stability Prediction

Faster predictions are attempted by [97] using rotor angle measurements of the 1st and 3rd cycle after FCT. Equally, in [98], stability predictions are attempted within one cycle after FC. The voltage magnitudes recorded from each bus at five different cycles - immediately, before and after fault occurrence - are used as inputs. By capturing the snapshots of the system at the above cycles, the severity of the fault can be incorporated within the input data. This approach yields an overall accuracy of approximately 99% for their test case. Speed of estimation is further improved in [99] by using a stability index vector consisting of the potential and kinetic energy computed at the fault occurrence and fault clearance. The stability index vector is used as input to a boosting DT, which returns accurate estimations even with high penetration of RES. Likewise, in [100] the on-fault period is used to predict the stability status and stability level. These are converted into a set of binary variables, which ensure the validity of the predictions.

Rolling Window Predictions

A two-step prediction model is designed in [101]. A pre-processing module is placed in series with a DT that consists of 19 Linear SVM (one for each feature) to transform the n^{th} cycle time series vector into a scalar. Conversely, the DT inputs are reduced to just 19 rotor angle measurements instead of 171. The proposed approach yields a similar performance to a single SVM but the main advantage of this particular design is the 97% acceleration of the training process. Nonetheless, this model is vulnerable to critical cases (near the decision boundary).

In a like manner, [102] implements a ML-based relay comprised of three DTs. The relay uses a 6-cycle rolling window for all DTs, trained for different tasks. These tasks include the detection of faults by DT_1, detection of fault clearance by DT_2 and instability prediction within 6 cycles from FCT by DT_3. To protect the generator, the relay trips according to the stability prediction or when

a fault clearance is not detected. By following this approach, the instability of the generator can be detected 140ms faster than the actual occurrence time. In [103] a NN predicts stability according to the fault type, which is provided by a fault detection algorithm. The proposed scheme returns additional details if an unstable case arises, such as the time frame until instability.

The exploitation of time dependencies within time series signals is discussed by [104]. Concretely, a LSTM-RNN[6] is adopted for fast transient instability prediction. Notably, performance appears to remain high even with noisy data. The gated recurrent unit RNN (GRU-RNN) capture the long-term dependencies of sequential inputs by implementing an information flow regulation without independent memory cells [105], meaning a lower number of parameters per unit. The authors in [106] study the effectiveness of an GRU-RNN ensemble to achieve higher accuracy and feature extraction. Prediction accuracy of the system status is compared with a GRU-RNN and a LSTM-RNN. Although accuracy with the other models is comparable, the average response time is 220ms to 240ms faster; hence, subsequent control actions can be initiated sooner.

High Penetration of RES

As integration of RES increases, stability is affected due to (a) the intermittent nature of RES prime movers and (b) the grid connection based on low-inertia power electronic converters. To analyse RES integration with respect to transient stability, [107] examines the effects of different wind power penetration levels to the overall generation mix. Modified systems with a ratio of 4: 1 between wind and conventional power plants indicate there are more unstable cases with a higher penetration of wind generation. An Extreme Learning Machine (ELM)[7] predicts the stability status with an accuracy of 97.5%, and the value of a frequency stability index (as in (2.3) where, $f_{min,i}$ is minimum frequency in bus i and $f_{cr,i} + k * t_{cr,i}$ is the acceptable frequency dip) with mean average percentage error (MAPE) of 0.64%,

$$FSI = \left(f_{min,i} - f_{cr,i} + k * t_{cr,i}\right) * 100 \tag{2.3}$$

A decision forest (DF) is suggested by [108] for stability status prediction of the modified NETS-NYPS with 20% RES penetration. The DF consists of several trees, each trained and tested under a specific topology so that the model will be robust against different network configurations. The

[6] The Long Short Term Memory Recurrent Neural Network has the distinct characteristic to store data of preceding cycles and process them together with the current data

[7] ELM is a special type of NN that comprises an input, a hidden and an output layer. The key aspect of the ELM is that it does not tune the parameters of the hidden layer through back-propagation or similar methods. Yet, the ELM guarantees convergence to the global minima at a much faster rate than conventional NN [191]

authors tried to incorporate into their study the uncertainty of high RES penetration, thus an optimal power flow is utilized, which considers the probability distribution function of those sources. Essentially, the prediction output answers the binary (stable/unstable) question through a multi-class (topology ID) approach. Furthermore, hierarchical clustering groups the generators that exceed 360^{o} angle separation, as well as the sequence of de-synchronization (patterns) to provide more details to system operators.

The sensitivity of DT models in regard to uncertainties is examined by [109]. These uncertainties are defined within four different test sets consisting of fault location, fault duration, operating conditions, and network topology. Each test set is used to examine the performance of a DT model according to the classification accuracy for different (i) inputs and (ii) previously unseen measurements (by the DT) of post-fault cycles. The results reveal that DT performance is not robust against uncertainties related to operating conditions and network configurations. As a general observation, the authors show that the accuracy of a DT improves when post fault measurements from longer cycles are used.

Control

Remedial Action Schemes

Generally, a remedial action scheme (RAS) detects potential cases of instability and performs control actions in order to maintain stability. Usually, RAS is either event-based, developed according to a particular combination of operating conditions and types of disturbances, or response-based, activated when measurement signals exceed a specific safety threshold.

Consequently, a DT algorithm triggering a RAS to ensure system security is examined by [110]. The DT evaluates the security of operating points against a set of credible contingencies. For the insecure cases, a transient stability constrained optimal power flow (TSC-OPF) reschedules generator power output to move the insecure case to secure regions at the lowest cost. The secure and insecure regions as defined by the DT in the parameter space provide an additional transparency in the operational limits of the system.

Likewise, [111] develops two SVMs for evaluating the transient security of a system coupled, also with control actions. Specifically, a pair of linear-SVM and Gaussian-SVM work in series to classify critical operating points as secure or insecure. The Relief-F algorithm selects the most important system variables reducing the size of feature vector while the Liner SVM screens out

critical samples that lie close to the hyperplane. Because of the linear Kernel, the trained SVM (in (3.6)) can be integrated into the TSC-OPF to ensure generator dispatch is transient stable. However, the linear Kernel might not provide the required performance in more complex cases. A reformulation of the ReLU converts the NN into mixed-integer linear program, which permits the modelling of more complex tasks with non-linear transformations while ensuring the problem tractable [112]. As shown in [113], both dynamic and static constraints can be integrated into the OPF, approximating the feasible space with higher accuracy.

As suggested by [114], a combination of the event and response based RAS can provide better results, both in speed of detection and adaptability to unseen scenarios. Specifically, a DT is trained to detect instability and activate RAS, considering the mode of disturbance (MOD) as well. The designed RAS aims to minimize the probability of islanding in the modified IEEE 39-Bus system. The proposed scheme, as the authors point out, needs improvement in detecting unstable cases caused by a wider range of MODs.

A hybrid model consisting of an unsupervised NN-based model called self organizing maps (SOM) and DT is proposed by [115]. This hybrid model initially clusters the set of probable operating points of the Greek National Grid into different nodes according to a similarity index. Subsequently, and for each node, the ratio of secure operating points to the total operating points is calculated. For a ratio close or equal to one, the node is characterized as secure. Conversely, nodes with a ratio close or equal to zero are characterized as insecure. However, nodes with an indeterminate ratio that cannot be recognized as belonging to either category (secure/insecure) are passed on to a DT for further examination. A load shedding scheme is also designed to curtail load so that a node's insecure operating points can move to a secure one.

Integration of Power Electronics

Microgrids and Distributed Generation (DG), in general, impose certain challenges to the conventional operation of power systems, such as frequency fluctuations. Concretely, [116] propose a ML pipeline comprised of three modules. The first two modules work independently to generate (a) the estimates of the load angle using an ANFIS and (b) the speed and electromagnetic torque using a state estimator. The last module predicts the torque coefficients of the generator using a NN to minimize frequency variations.

The response of RES with power electronics can be significantly improved with ML. For instance, using NN instead of PI modules in the grid connected converters (GCC) of a WPP result in faster

convergence to steady state and lower oscillation amplitudes [117]. Similar results are reported for an offshore power plant [118], for WPPs [119], [120] and for an ultra-capacitor [121].

LCL-filtered GCC coupled with conventional controllers are prone to instability. To mitigate this issue, [122] propose a NN to estimate the voltage control references from the current deviations, all in dq frame. The loss function modifies the NN parameters to minimize the errors of the d and q currents for the provision of suitable voltage references. This NN based control for a single phase GCC with an LCL filter demonstrates robustness to parameter changes, operated at low sampling rates with no special damping strategies. Regarding inverters, transient stability at Point of Common Coupling (PCC) can be predicted with an accuracy higher than 99% within 100mu s using a DNN [123], while its transients can be reduced through the derivation of a linear regression model and its integration with the phase-locked loop [124].

Power oscillations can limit the capacity of transmission lines, or even cause widespread blackouts. Damping of such oscillating can be improved by regulating properly the parameters of static series synchronous control as shown in [125]. The optimal regulation of the voltage magnitude and polarity is achieved through a Fuzzy-NN with self-constructing attributes, which allow for online training. The results show a 10-fold increase in the inter area oscillation damping with respect to conventional Power System Stabilizers (PSS).

Noisy and Missing Data

Communication delays or corrupted and missing information packages are highly likely to occur. For instance, communication delays can impose an upper bound of 6 cycles on the available predicting time [114]. In addition, during transients, PMU measurements can differ from true values. Driven by the above, [126] argues that global system variables incorporate higher discriminant power and computational efficiency, regardless of scale and robustness against missing data. Two groups of global variables are proposed based on rotor angles and voltage magnitudes. The Gaussian-SVMs, developed for the stability assessment, achieve 99% accuracy although voltage based SVM has a faster detection speed ($\approx$10 cycles less). Given the selection of the specified features, missing PMU data do not seem to affect performance when signals from critical generators are still active.

Considering data limitations imposed by PMU availability, in [127], only 8 buses out of the 39 buses of the New England Benchmark system have PMU installed. A boosting ensemble is developed to evaluate the security of operating conditions against a scenario list. The proposed

model could easily incorporate new cases by recalculating the weights of the elements in the training subset and by modifying only the subtree associated with the new case.

Also, a data denoising algorithm is proposed by [128]. The algorithm consists of stacked autoencoders[8] (AE) and a convolutional[9] NN. This approach increases robustness and suitability to highly non-linear problems with high dimensionality. Compared to other NN and SVM based models, the stacked AE-CNN achieves the highest accuracy of over 97% when evaluated against noisy inputs. Instead of feeding directly the PMU measurements into the ML model, [129] propose using the last steady state snapshot to extract equivalent dynamic models of the system for running TDS, which will provide clear signals. On average, this process takes around half a second to complete after fault occurrence.

Remarks

Despite the advantages of the AI methods presented in the above applications, several limitations exist. For instance, in fixed window predictions the authors use specific time points of an event to construct the feature vector. Identifying these time points can be extremely difficult. The process that triggers the ML model to make the prediction is not described although this is a crucial aspect and especially relevant for actual implementations.

Additionally, most of the reviewed papers assume either full observability of the system, no communication delays or very fast sampling rates. However, these assumptions are often invalid. Different studies suggest the use of rotor angles and speeds [130], the potential and kinetic energy [131] or the frequency deviation [132], [133], all of which are not readily available. Feature engineering [134] and dimensionality reduction techniques [63], [66] improve the performance but increase the computation time.

The natural imbalance between stable and unstable cases can result in (a) overestimating performance and (b) biasing the ML model towards the majority class. Authors prefer designing their models so that when unsure, they classify cases as unstable or uncertain. This reveals a limited capability to define the stability boundary accurately. Furthermore, the accuracy of stability predictions is not enough [135], therefore the sensitivity of ML models to small changes of the

[8] The auto-encoder is a type of unsupervised learning model consisting by an encoder and a decoder. The encoding-decoding process makes those models robust to noisy data as well as providing an alternative approach to dimensionality reduction

[9] The Convolutional Neural Network is commonly used in image processing. The convolution layer of the network uses filters to isolate specific patterns in the data to extract the characteristics of an image.

system measurements needs to be studied. This is known as *Adversarial Examples* [136] and describes the impact of small perturbations of inputs that cause the ML model to give highly confident but false predictions.

To improve the accuracy and credibility of their methods, many of the reviewed papers introduce the uncertain class, which includes critical cases that cannot be classified as stable/unstable with high confidence. Research should focus on approximating the decision boundary with higher precision to minimize cases with inconclusive results. Potentially a more efficient formulation of the training database can help towards that end [137].

Nevertheless, a better understanding of the above issues can arise from demonstrations in real life conditions. Currently there are projects involving actual demonstrators such as the FNET/GridEye project [138] and FLEXITRANSTORE project [139]. Finally, recent developments in computer science have shown that NN can learn the mathematical expressions describing a system [140]. Indeed, the NN can solve the differential equations faster than numerical integrators, making them an ideal alternative for accelerating TDS for power system rotor angle stability. Currently, this approach is used for state estimation in power systems [141].

CHAPTER 3

Artificial intelligence Methodologies

Among the plethora of emerging measurement devices, the PMU with its high sampling rate allows for the adoption of data-driven methods in *Rotor Angle Stability*. Machine Learning (ML) and Deep Learning (DL) can fully exploit the vast amount of data to enhance the stability of the system.

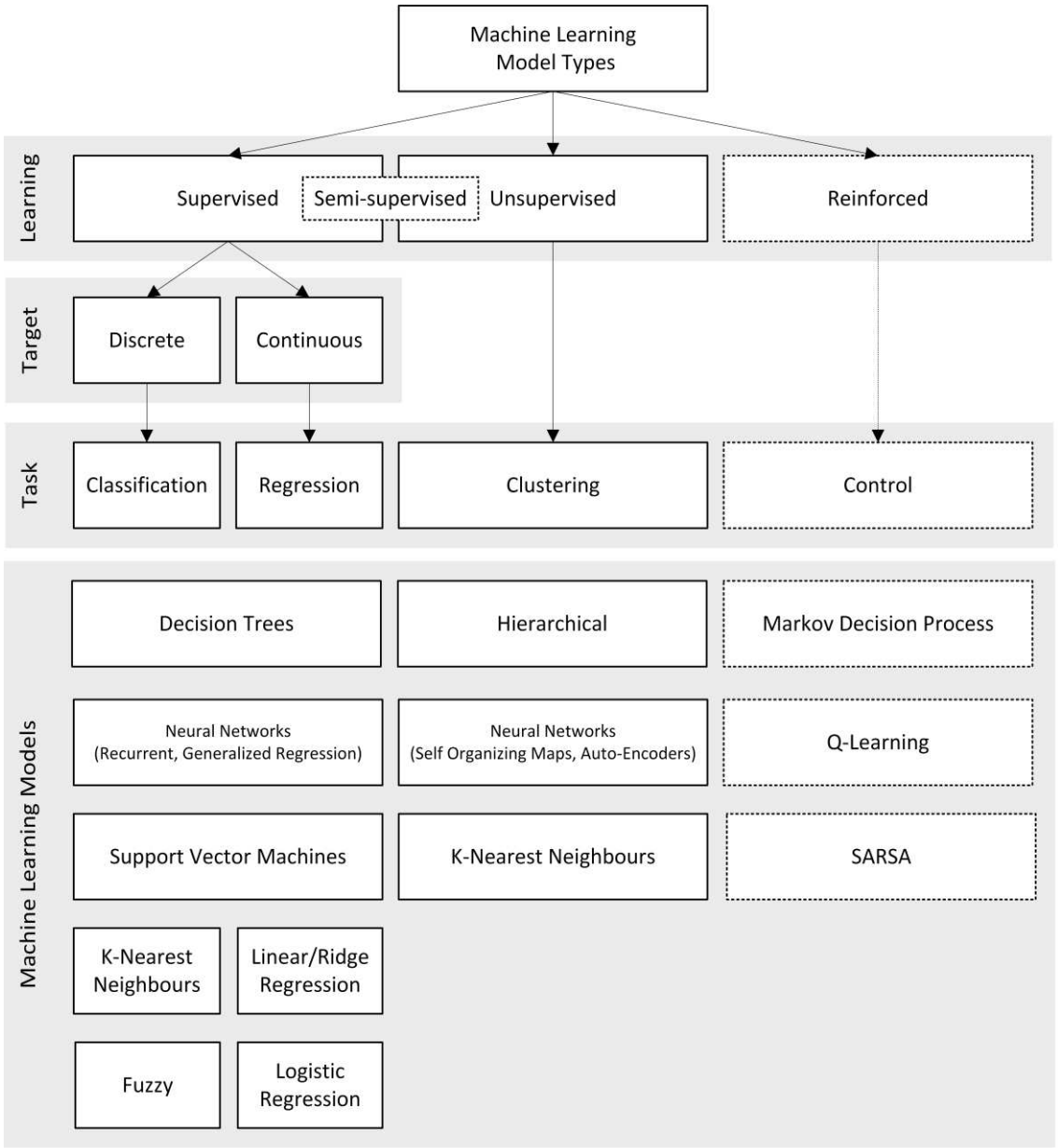
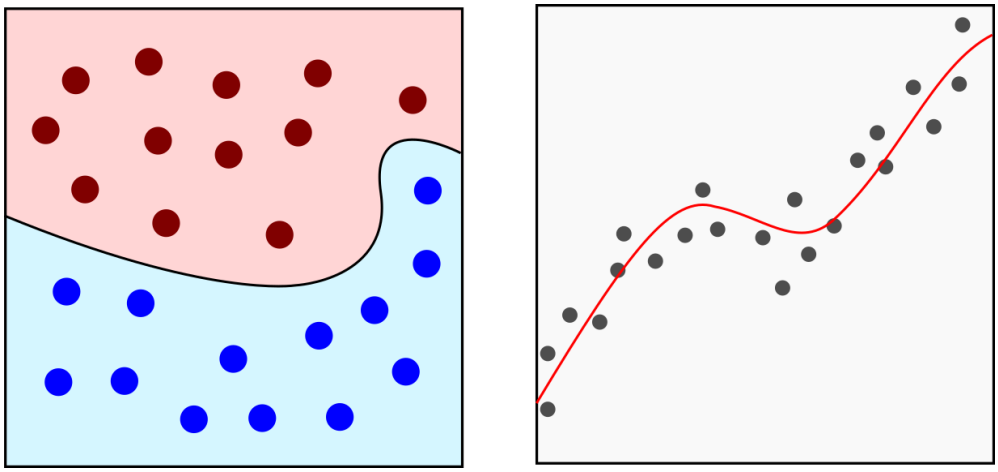


Figure 5: Overview of Machine Learning

In contrast to traditional approaches, such as TDS and MA, ML does not assume any knowledge of the physical system. Nevertheless, machines can learn from data to perform a specific task *"without being explicitly programmed"* [142]. ML encompasses a wide range of algorithms based on the type of learning and task as summarized in Figure 5.

Methods of Learning

ML consists of three distinct branches, each representing a different learning method; namely, supervised, unsupervised and reinforced learning. In supervised learning, machines use example input-output pairs to learn a mapping function that approximates the true relationship between them. Supervised learning tasks are regression for continuous output and classification for discrete. In regression tasks the algorithm tries to learn the best fit for the training data, while in classification the decision boundary that separates the classes, as shown in Figure 6. In contrast, unsupervised learning does not require the outputs to be known because it identifies hidden patterns within the input data to provide insights about the structure and characteristics of the system. Typical unsupervised learning tasks involve clustering and anomaly detection. Semi-supervised learning (Figure 5) combines the two previous learning types using both labelled and unlabelled samples to enhance generalization. Finally, in reinforced learning the model interacts with its environment in order to determine the optimal sequence of actions through a reward/penalty scheme [143].



a) Classification *b) Regression*

Figure 6: Supervised Learning Tasks

Naturally, to evaluate the performance of an algorithm for each task is different. Specifically, the metrics used in regression measure the error between the estimated and true output such as the mean absolute error (MAE) in (3.1), mean squared error (MSE) in (3.2) and root mean squared error (RMSE) in (3.3) being the most common. The suitability of the model w.r.t. the given task is measured by the R^2 score in (3.4). Accuracy in classification is defined as number of correct classifications over the total number of classifications. When classes are unevenly distributed accuracy can be a misleading measure. Instead, for uneven class distributions, a holistic performance evaluation of the model is necessary thus, the confusion matrix in Table 1, and the receiver operating characteristic curve are preferred over the simple accuracy ratio.

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \tag{3.1}$$

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{3.2}$$

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{3.3}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \bar{y}_i)^2} \tag{3.3}$$

Table 1: Confusion Matrix

	Actual Positive	Actual Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

Model Development Framework

The development of a machine learning model can be achieved by following the steps shown in Figure 7. The pre-process involves all these actions that prepare the dataset to be used for learning the given task. Some of these actions are cleaning, scaling, sampling, and transforming the inputs/outputs. Some algorithms, such as NN, are sensitive to inputs of different scales meaning that the ones with larger range will dominate over the other ones.

Normalization and standardization are two typical approaches for data scaling. The former refers to the rescaling of input features within the desired range (3.4), while the latter centres features to mean μ=0 and deviation to σ=1 (3.5).

$$\bar{x} = \frac{x - x_{min}}{x_{max} - x_{min}}(max^{new} - min^{new}) + min^{new} \tag{3.4}$$

$$\bar{x} = \frac{x - \mu_x}{\sigma_x} \tag{3.5}$$

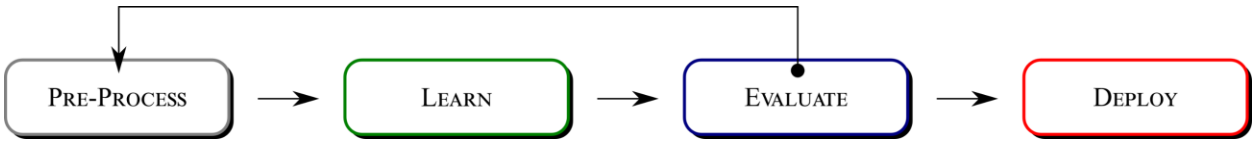


Figure 7: Model Development Workflow

Apart from general suggestions and guidelines there is now formal way to determine the best model a priori. Usually, a set of candidate models is trained using the same inputs and after an analysis of their errors and performance the least performing models are eliminated, and the process is repeated for the best models with hyper-parameter tuning. Concretely, the learn and evaluate steps are both closely related as the performance of the model in the unseen cases is the final selection tool. In many cases revisiting the pre-process step can result in improving the performance. The final step is to deploy the model to perform the required task using real data. As power systems are continuously change, it is necessary to monitor the performance of the model.

Machine Learning Models

Linear Regression

Simple yet effective, Linear Regression can solve a wide range of problems with good accuracy. It aims to find the parameters $\boldsymbol{w} \in \mathbb{R}^{n\times 1}$ that minimize the error between the set of independent variables $\boldsymbol{x} = [x_1, x_2, \dots, x_n]$ and dependent variable(s) y. This relationship is given by the linear combination between the inputs and parameters, as in (3.6). The optimal weights can be estimated by using iterative methods or the closed form expression in (3.7), where $\boldsymbol{X} \in \mathbb{R}^{n\times m}$ are the inputs, $\boldsymbol{y} \in \mathbb{R}^{n\times 1}$ are the outputs m and N are the number of patterns and features, respectively while the $\boldsymbol{X}^{+}$ is the Moore–Penrose inverse.

$$\hat{y}(\boldsymbol{x}, \boldsymbol{w}) = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + w_o \tag{3.6}$$

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{w}^{\mathrm{T}} \rightarrow \boldsymbol{w} = \underset{\boldsymbol{w}}{\operatorname{argmin}} \|\boldsymbol{X}\boldsymbol{w} - y\| \Rightarrow \boldsymbol{w} = \boldsymbol{X}^{+}\boldsymbol{y} \tag{3.7}$$

Decision Trees

Traditionally, Decision Trees (DT) were designed manually by analysts and researchers as a highly intuitive decision-making tool. Naturally, DT in machine learning follows the same structure i.e. nodes represent a decision and links the possible outcomes. A DT resembles an upended tree, meaning that the root lies at the top and the leaves at the bottom. Several algorithms have been developed for the construction of a DT. For instance, the ID3 algorithm, which uses the *information gain* to decide the splitting at each node, is one of the simpler methods [144]. The C4.5 algorithm, based on the ID3, includes a pruning stage, which reduces DT complexity by removing branches that add little discrimination power and can also increase robustness in relation to missing values [144]. Both algorithms, however, are suitable only for discrete classes. The classification and regression trees (CART) algorithm was developed to include continuous values as well. The main characteristics of the CART algorithm are (a) that it can generate both classification trees (CT) and regression trees (RT) and (b) all non-leaf nodes have exactly two outgoing branches, hence the term binary trees [145].

Support Vector Machines

Maximum margin classifiers or SVM separate samples into the predefined classes by constructing a decision boundary or hyperplane that maximizes the marginal space between them. A hyperplane can be mathematically expressed as in (3.8), where $\mathbf{w} \in \mathbb{R}^{N\times 1}$ is the weight vector $\boldsymbol{x} = [x_1, \ldots, x_N] \in \mathbb{R}^{N\times 1}$ is the feature vector and $\mathrm{b} \in \mathbb{R}$ is the bias. The perpendicular distance d between a point and a hyperplane is given by (3.9), where $\|\cdot\|_2$ is the Euclidean norm. Therefore, to maximize the margin, it is necessary to minimize the norm of the weight vector as shown in (3.10) subject to (3.11) and (3.12).

$$\mathbf{w}^{\mathrm{T}}\mathbf{x} + \mathrm{b} = 0 \tag{3.8}$$

$$d = \frac{\mathbf{w}^{\mathrm{T}}\mathbf{x} + \mathrm{b}}{\|\boldsymbol{w}\|_2} \tag{3.9}$$

$$\min_{w,b} \ \frac{1}{2}\mathbf{w}^{\mathsf{T}}\mathbf{w} \tag{3.10}$$

$$\text{s.t.} \quad t_j(\mathbf{w}^{\mathsf{T}}\mathbf{x} + \mathrm{b}) \geq 1 \tag{3.11}$$

$$t_j = \begin{cases} -1, y_j = 0 \\ 1, \quad y_j = 1 \end{cases} \forall j \in S_{train} \tag{3.12}$$

Formulating the optimization problem in this way does not allow misclassifications, which can cause overfitting and, in certain cases, convergence issues. To overcome these limitations, a slack variable ζ measures the margin violation of inputs, and parameter C defines the balance between maximizing the margin size and minimizing the violations. Hence, the soft margin SVM shown is defined in (3.13) subject to (3.14), (3.15) and (3.12).

$$\min_{w,b} \ \frac{1}{2}\mathbf{w}^{\mathsf{T}}\mathbf{w} + \mathrm{C}\sum_{j \in S} \zeta_j \tag{3.13}$$

$$\text{s.t.} \quad t_j(\mathbf{w}^{\mathsf{T}}\mathbf{x} + \mathrm{b}) \geq 1 - \zeta_{\mathrm{j}} \qquad \forall j \in S_{train} \tag{3.14}$$

$$\zeta_j \geq 0 \qquad \forall j \in S_{train} \tag{3.15}$$

The above optimization problem it is usually expressed in its dual form using the Lagrange multiplier $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_n] \in \mathbb{R}$, as shown in (3.16). By solving the dual problem, it is possible to incorporate the kernel function, a very important aspect of SVM for the non-linear cases [146]. In addition, the Lagrange multipliers identify the inputs, also known as support vectors (SV), which contribute to the construction of the hyperplane, i.e. with $\alpha_j > 0$. The number of SV indicate the complexity of the decision boundary [144].

$$\min_{\boldsymbol{\alpha}} \ \frac{1}{2}\sum_{j \in S}\sum_{i \in S} a_j a_i t_j t_i K(\mathbf{x}_j, \mathbf{x}_i) - \sum_{j \in S} a_j \tag{3.16}$$

$$\text{s.t.} \quad a_j \geq 0 \qquad \forall j \in S_{train} \tag{3.17}$$

From (3.18) the classification of new samples solely depends on the kernel function $K(\cdot,\cdot)$. For problems where the input space is non-linear, this function defines a feature space that can linearly separate the data [147]. In general, Kernel functions enable SVM to use dot products to separate the data in higher dimensions without explicitly defining a mapping function, which is difficult to find.

An example regarding the mapping of the input space into a linear separable feature space is shown in Figure 8. The XOR classification problem is non-linear separable at the input space as indicated by the four data points that correspond to the truth table of the XOR gate. By applying a mapping function ϕ the input space is raised into a higher dimension feature space where the problem becomes a linear separable one. The most used Kernel functions are the Gaussian, the polynomial and linear as listed in Table 2. Gaussian and polynomial can be slower than the linear kernel but they have greater power in separating highly non-linear datasets [148].

$$\hat{y}(\mathbf{x}) = \sum_{j \in S} a_j t_j K(\mathbf{x}, \mathbf{x}_j) + b \tag{3.18}$$

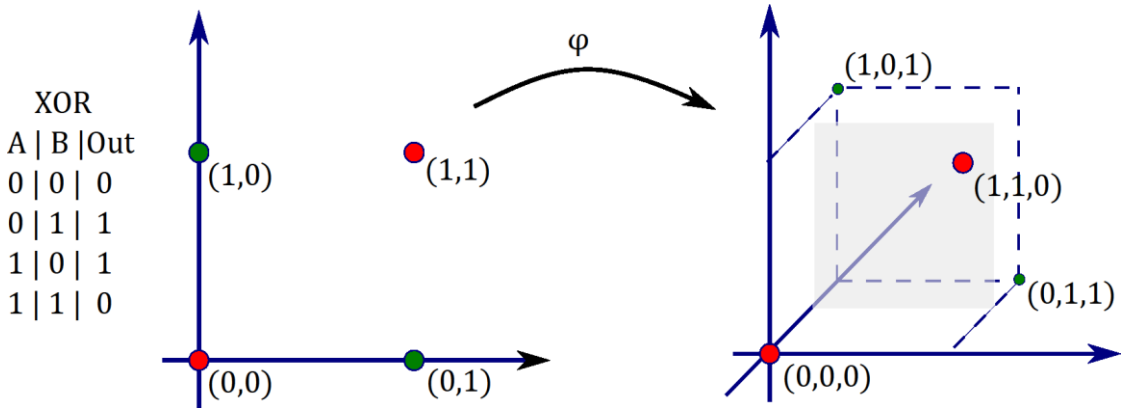


Figure 8: Input space mapping to feature space: the XOR classification problem.

Table 2: Common Kernel Functions [149]

	Kernels
Polynomial	$K(x_i, x_j) = (x_i x_j)^p$, p : degree
Gaussian	$K(x_i, x_j) = e^{-\gamma \lVert x_i - x_j \rVert_2^2}$, $\gamma = \frac{1}{2\sigma^2} > 0$
Linear	$K(x_i, x_j) = x_i x_j + c$, c : constant

Neural Networks

Neural Networks (NN) are nonlinear ML models based on the architecture and function of the human brain cells. In principle, NNs aim to solve highly complex problems through linear combinations of the outputs among the set of nodes (neurons) grouped together in different layers, which communicate with each other via weighted links (synapses). A NN is characterized by the number of layers L and number of neurons K_l at each layer, where $l \in \{0,1,\dots,L\}$. In supervised learning the NN is a function $g(x; \boldsymbol{W}, \boldsymbol{b}): \mathbb{R}^N \rightarrow \mathbb{R}^M$, which is parameterized by weights $\boldsymbol{W}$ and biases $\boldsymbol{b}$.

Graphically this can be represented as in Figure 9, where inputs and outputs are denoted by the vector $\boldsymbol{x} = [x_1, \dots, x_N] \in \mathbb{R}^N$ and $\hat{\boldsymbol{y}} = [\hat{y}_1, \dots, \hat{y}_M] \in \mathbb{R}^M$, respectively. Note that $K_0 \coloneqq N$ and $K_L \coloneqq M$ are the dimensions of the input and output vector respectively. Connections between the neurons are represented by weights $\boldsymbol{W}_l \in \mathbb{R}^{K_l \times K_{l-1}} \forall\ l \in \{1, \dots, L\}$ and biases $\boldsymbol{b}_l \in \mathbb{R}^{K_l}\ \forall \in \{1, \dots, L\}$. For convenience, it is customary to augment the weight matrices with the biases, introducing an auxiliary neuron at each layer with a fixed value equal to unity.

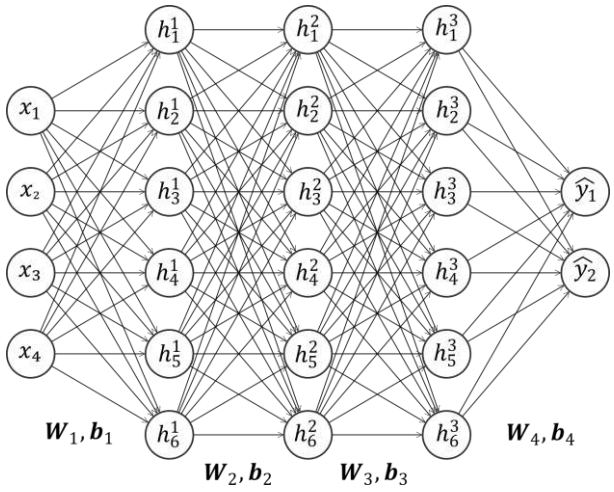


Figure 9: Typical NN structure

The principal operations conducted by the NN (see (3.19)-(3.23)) are the linear combination of the inputs at each neuron, as in (3.21) and a transformation $\sigma(\cdot)$, usually through a nonlinear function e.g. hyperbolic tangent (i.e. tanh), sigmoid and Rectified Linear Unit (ReLU). Some of the most common activation functions along with their first derivatives are plotted in Figure 10. Input-output pairs, $\{(\boldsymbol{x}_j, \boldsymbol{y}_j)\}_{j \in S}$ of train set S are presented to the model, where parameters are modified to minimize a loss $\mathcal{E}(y, \hat{y})$ according to the error metric $\mathcal{E}$. Generally, all NN are known for their ability to approximate any function [150], [151].

$$h_i^{(0)} = x_i\ \forall\ i \in \{1, \dots, N\} \tag{3.19}$$

$$\hat{y}_i^{(L)} = h_i^L\ \forall \in \{1, \dots, K_L\} \tag{3.20}$$

$$\overline{\boldsymbol{h}}_i^{(l)} = \boldsymbol{W}^{(l)} \boldsymbol{h}^{(l-1)} + \boldsymbol{b}^{(l)}\ \forall\ l \in \{1, \dots, L-1\} \tag{3.21}$$

$$\boldsymbol{h}^{(l)} = \sigma\left(\overline{\boldsymbol{h}}_i^{(l)}\right) \forall\ l \in \{1, \dots, L-1\} \tag{3.22}$$

$$\hat{\boldsymbol{y}}^{(L)} = \boldsymbol{W}^{(L)}\boldsymbol{h}^{(L-1)} + \boldsymbol{b}^{(L)} \tag{3.23}$$

Since the late 80s, *backpropagation* [152] has been the standard training procedure of NN. This procedure consists of two stages: the forward pass for computing the error and the backward pass for the gradient of that error w.r.t. each parameter. Once the gradients are computed, the parameters are modified towards the opposite direction. The procedure is repeated for a given number of iterations. Typical optimization algorithms used in *backpropagation* are Stochastic Gradient Descent, RMSprop and Adam [150], [153]. In power systems the use of neural networks has been studied since the early 1990s in many areas like planning, operation and analysis [154].

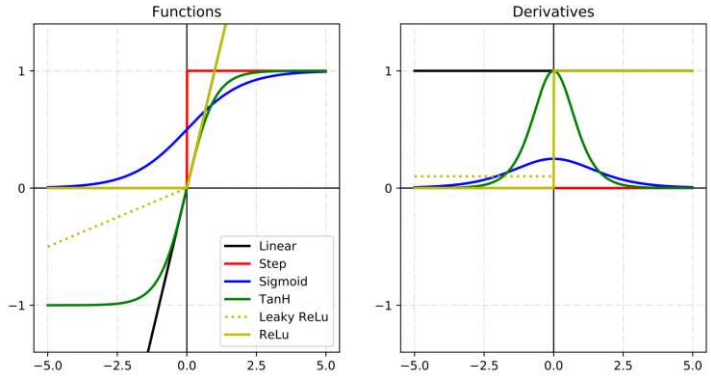


Figure 10: Common Activation Functions

Ensemble Models

Ensemble learning refers to the aggregation of a group of base learners that work collectively for the same objective to achieve more accurate results. By using different subsets of the training set it is possible to incorporate diversity (i.e. low correlation between the individual classifiers) to improve generalization, even in datasets with highly imbalanced class distributions [155]. There are two main techniques for developing ensemble models: *bagging* and *boosting*.

In Bagging, k classifiers are trained separately, each using the S_k subset of the original dataset S. The base classifiers operate independently each providing its own output. In classification tasks, the final output is usually determined through a voting scheme [155], [156] whereas in regression tasks, a weighted average. A representative application of this approach is Random Forests (RF) which, as the name indicates, consists of k uncorrelated classifiers which are trained using random subsets of dataset S in addition to random features. The output of RF is determined either through a majority voting scheme (for classification) or through soft voting (for regression) [79], [157].

In Boosting, classifiers are trained in series using a sample weighting scheme to improve the correct classification of misclassified samples. Initially each element of the dataset S is assigned an equal weight. Then k classifiers are developed through k iterations. After each iteration, the weight of each element in S is adjusted upwards if misclassified by the current learner and downwards if otherwise [155]. This approach creates a new train set after each iteration for the next classifier which helps favour the misclassified elements. Lastly, the output of the boosting ensemble is generated through the voting procedure as mentioned before, but this time each classifier vote is associated with a weight factor based on its classification accuracy [156].

CHAPTER 4

Monitoring Electromechanical Modes

The electromechanical interactions expressed through low frequency oscillations (LFO) are a phenomenon that exists naturally in power systems due to the operation of generators in parallel and the transfer of power over long distances [158]. Fast control devices such as the Automatic Voltage Regulator (AVR) improve the transient stability but have the opposite effect on the oscillatory response of the system. Besides damping, LFO are characterized by their frequency and damping ratio. In general, frequencies of LFO range between 0.1Hz - 2.0Hz and are further categorised into Local and Inter-area [27]. The latter have a deeper impact in the stability and operation of power systems. Underdamped inter-area modes of oscillation limit the maximum available power transfer capacity of transmission lines leading to significant financial costs [159]. In addition, power swings of with increasing amplitudes can trip protective devices and cause power outages [160]. Ensuring that sufficient damping resources with the capacity to adapt are vital, especially as power electronics impact the stability of power systems.

As discussed, conventional methods for identifying the characteristics of LFO are either unsuitable for online applications or too slow for control purposes. Specifically, MA is accurate but requires the state space model of the system and its slow. Signal processing methods, such as Fast Fourier Transform (FFT) or Prony require a probing signal and the selection of a time window, which is difficult to determine [159]. On the other hand, with AI it is possible to model the relationship between the operating points of the system with the corresponding characteristics of targeted LFO to provide faster the estimation results. To illustrate, Figure 11 shows both the steady state and ringdown response of a system where an AI model can provide faster the prediction of the mode characteristics that the signal processing methods like Prony and FFT.

Two-area system

The two-area system is a popular model for analysing oscillations and damping methods [161], originally presented in [27]. As shown in Figure 12, it is a symmetric system that consists of two areas each with two generators (G1-G2, G3-G4) one load (L7, L9) and one shunt capacitor (C7, C9) connected through a tie line with two parallel circuits.

Every generator has installed an IEEE Type AC4A Excitation System (AVR), a steam turbine governor (GOV) and a PSS of type 1. By design Area 1 exports to Area 2 approximately 400 MW. There are three LFO: an inter-area at 0.55 Hz and two local modes with frequencies of 1.05 Hz and 1.08 Hz in Area 1 and Area2, respectively. The system is modelled in Digsilent's Powerfactory 2019 SP3 simulation platform.

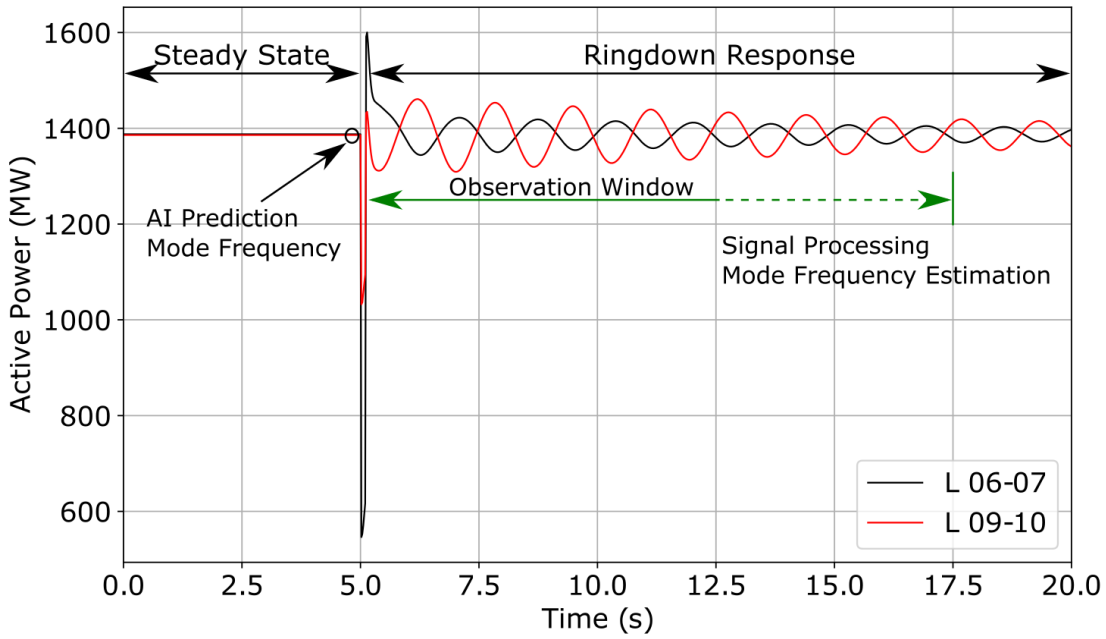


Figure 11: Differences between AI and signal processing methods

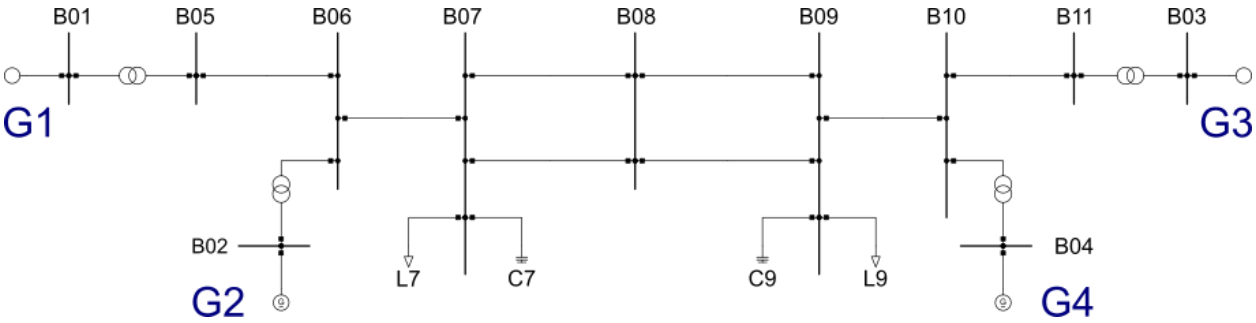


Figure 12: Two-area system

Table 3: Number of elements in Two area system

Elements	Number
Buses	11
Lines	8
Transformers	4
Generators	4
Loads	2

Database Generation

The performance of data-driven approaches, such as AI, depends on the quality and quantity of the data. In many cases these data are not readily available or not labelled for the desired supervised learning task. Using the system model and a simulation platform it is possible to manipulate the parameters and operating conditions of the system to generate a database by simulating each scenario. Naturally, the accuracy of the system model to the real system is critical for generating data that will reflect the actual conditions. However, at the case of benchmark systems, such as the two-area system, this is not an issue.

The characteristics of both local and inter-area modes are closely related with the operating conditions of the system, i.e. generation dispatch, loading levels, and network topology [27]. For instance, as demand stress the system towards its limits the damping of the modes moves closer to the y-axis and the stability boundary. Concretely, the database should consist of a wide range of operating scenarios to capture as accurately as possible the behaviour of the inter-area modes. Algorithm 1 summarizes the process followed to generate the required data by considering loads to be random variables drawn from a Gaussian distribution. Apart from the number of *Simulations,* the initial active and reactive power $P_{initial}$ and $Q_{initial}$ respectively, the mean μ and standard deviation σ of the Gaussian distribution are given as inputs to the algorithm.

ALGORITHM 1: **DATABASE GENERATION**

***input**: Simulations, μ, σ, $P_{initial}$, $Q_{initial}$*
$i \leftarrow 0$
while (**$i <$ Simulations**) do
  $r_1, r_2 \sim U(0,1)$
  $scaling \leftarrow \mu + \sigma\sqrt{(-2\log(r_1))}\cos(2\pi r_2)$
  $P \leftarrow P_{initial} \times scaling$
  $Q \leftarrow Q_{initial} \times scaling$
  failed ← Calculate Power Flow (.ComLdf)
  if** failed **then
    continue
  end if
  Calculate Modal Analysis (.ComMod)
  Export Results
  $I \leftarrow i+1$
end while

Here the Box-Muller transformation is used to convert a uniform to a Gaussian distribution [162]. The r1 and r2 coefficients are drawn randomly from the Uniform distribution of the open interval (0, 1) before calculating the scaling factor. In lines 6 and 7 of the algorithm the initial active and reactive power are scaled up or down. The power flow is computed and if converged the modal analysis returns the corresponding eigenvalues and eigenvalues of the system. Finally, both results are exported in .csv format and the process is repeated until the maximum number of simulations is reached. To automate the process the algorithm has been implemented in Python 3.6 to generate the random points and to interact with Powerfactory to execute the power flow (.ComLdf) and modal analysis (.ComMod) commands.

Data analysis and pre-processing.

For the two-area system, approximately 23000 scenarios have been simulated using Algorithm 1 with μ = 1 and σ = 0.1. The dataset, therefore, consists of equal input-output pairs represented by the system variables for each operating point and the corresponding eigenvalues. During these simulations it is assumed that the system is fully observable, meaning that all variables are available. The full dataset $S \in \mathbb{R}^{m\times(n+2g)}$ where $m = 23000$ is the total number of simulations $n = 122$ is the number of system variables and $g = 3$ are the number of modes. Before analysing the dataset, the three subsets are created S_{train}, S_{val}, S_{test} for training, validating, and testing the ML models with a ratio of 75%/15%/15%, respectively. The reason is that the final score needs to be recorded for the testing set that will remain unseen until the very end to prevent overfitting.

An overview of the characteristics of the train dataset $\boldsymbol{X}_{train} \in \mathbb{R}^{17250\times n}$, $\boldsymbol{Y}_{train} \in \mathbb{R}^{17250\times 2g}$, is depicted in Figure 13 where the variance of each variable is plotted as a function of its mean in log scale. This figure reveals the different scales of the data (i.e. per unit, MW, kW, degrees to name a few) and the low variance of some of the variables. Due to the nature of power systems and the co-dependence of the system variables there is a certain degree of collinearity in the dataset. In fact, most of the recorded variables have either positive or negative linear correlation above ±0.95. The correlation r is computed using the Pearson formula in (4.1) where x_i, y_i are two variables and $\bar{x}$, $\bar{y}$ are their respective means.

$$r = \frac{\sum_{i\in M}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i\in M}(x_i - \bar{x})^2}\sqrt{\sum_{i\in M}(y_i - \bar{y})^2}} \in [-1, 1] \tag{4.1}$$

The highly correlated features (i.e. blue dots in Figure 13) are discarded to remove redundant information. In addition, it has been shown that low variance features hold minimum predictive

power therefore, the remaining uncorrelated features that lie in the shaded area have a variance of less than 0.006 and can be eliminated from the dataset. The inter-area mode results are summarized in Figure 14.

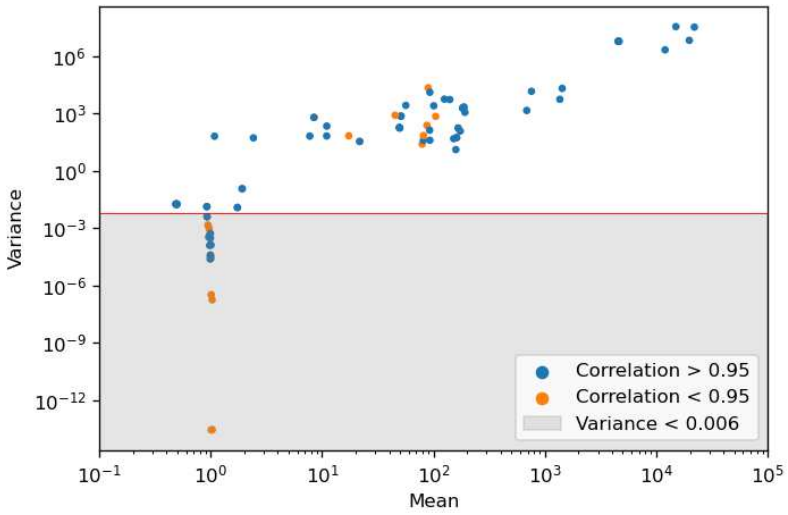


Figure 13 Mean vs Variance of System Variables

The majority of the simulated scenarios have shown that the damping of the inter-area mode is less than 5%. In many cases it is recorded a negative damping meaning that the system is in critical condition. Only a handful of cases has higher than 5% damping ratio. In contrast, the local mode in Area 2 (i.e. between generators G3 and G4) has several cases where the damping ratio is less than 5%, as depicted in Figure 15. The frequency of both modes varies significantly between the simulations.

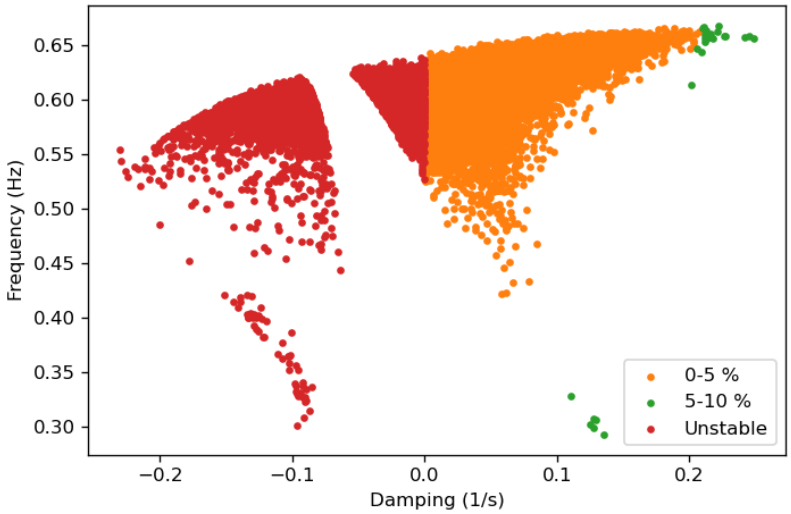


Figure 14: Inter-area mode results

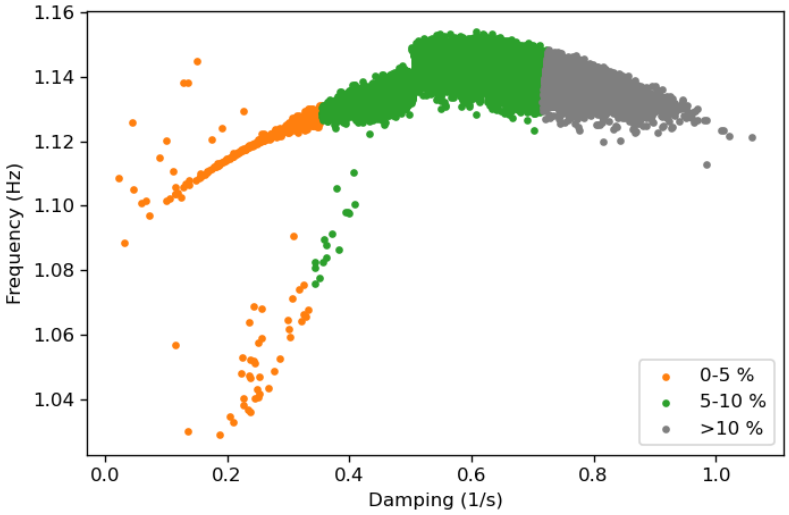


Figure 15: Local mode Area 2 results

Following the above the number of features are reduced from 122 to 22 hence the train and validations sets consist by 17250 and 3450 patterns respectively, i.e. $\boldsymbol{X}_{train} \in \mathbb{R}^{17250\times22}$, $\boldsymbol{Y}_{train} \in \mathbb{R}^{17250\times4}$, $\boldsymbol{X}_{val} \in \mathbb{R}^{3450\times22}$, $\boldsymbol{Y}_{val} \in \mathbb{R}^{3450\times4}$. It should be noted that because the algorithms cannot handle complex numbers directly, they are treated as separate target variables. As shown in Figure 13, the scale of the input features varies significantly due to the nature of the different variables. For instance, voltage magnitude is measured in per units while active power flow in MW. To alleviate any problems this may cause in the development of certain algorithms, all features are scaled to have mean $\mu = 0$ and standard deviation $\sigma = 1$ using the formula (4.2). However, cyclic features such as voltage angles should be treated differently because normalization does not reflect the actual relation between angles close to each other. For example, in Figure 16 reveals the difference between normalization and trigonometric representations for cyclic variables. It should be noted that the mean and standard deviation of the training set are used for normalizing the test and validation sets.

$$x_i' = \frac{x_i - \mu_{x_i}}{\sigma} \tag{4.2}$$

Model Learning and evaluation

The preliminary list of candidate models consists of Linear Regression, Gradient Boosting, Random Forests and Neural Networks. The latter has been developed using the Tensorflow library using the Keras Sequential API [163] while the rest using the Scikit-learn machine learning library [164]. The scikit models use the default hyper-parameters while the NN uses the ReLu activation

function for the two hidden layers of 50 neuros each. The adaptive momentum (Adam) optimize is used to train the NN using the default hyper-parameters. The goal is to learn a function $f: \boldsymbol{X} \rightarrow \boldsymbol{y}$ that maps the system operating conditions to the inter-area mode frequency. The candidate models are trained and validated using the corresponding sets described in the previous subsection.

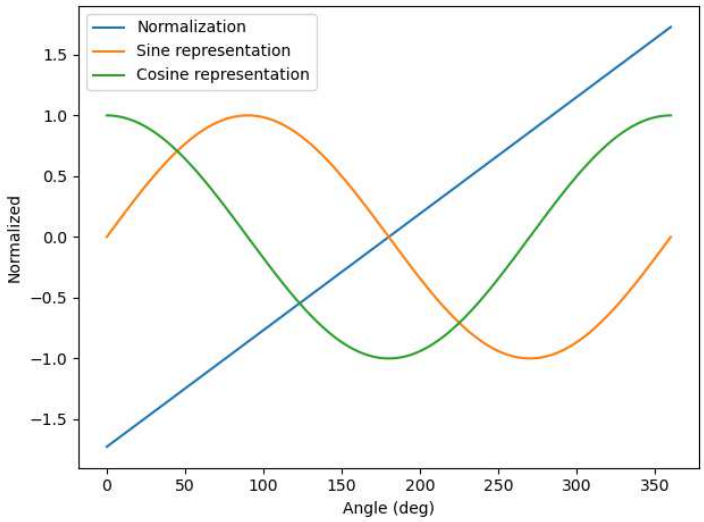


Figure 16: Illustration of normalizing cyclic variables using trigonometric representations

Using the training set and a 5-fold cross validation the learning curves of each model are plotted along with their fitting times in Figure 17. The dotted lines indicate the mean training error, the lines indicate the mean cross-validated error while the shade area around the line the standard deviation of the estimates.From the results the Gradient Boosting model has the lowest cross validated MAE with Random Forests being second to best. However, RF seem to have high variance, which can be reduced by identified the optimal hyper-parameters. The only issue is that the RF scales poorly w.r.t. the other models. In any case, the dataset size is small therefore, it permits an extensive parameter grid search. After performing grid search on the Random Forests, After identifying the optimal selection of parameters the new results are shown in Table 4. Figure 18 depicts a subset of the test set inter-area mode against the predict one from optimized RF.

There are some appealing properties about RF. They inherent their robustness to inputs of different types and scales, e.g. categorical features without encoding from decision trees. This eliminates the need of additional pre-processing steps that are necessary for the NN. However, as an ensemble the transparency of the decision-making process of a DT is diminished. In addition, RF have the capacity to provide a measure of importance with regards to the input features. This can help remove even further unnecessary features to reduce the complexity of the model but also it can be an indicator of where to install measurement devices.

Table 4: Optimized Random Forests Results.

RF	MAE	Parameters			
		DT	Min Pattern/Split	Max Features	Max Depth
Optimized	0.0866	1000	5	log25	80
Default	0.1501	100	2	-	-

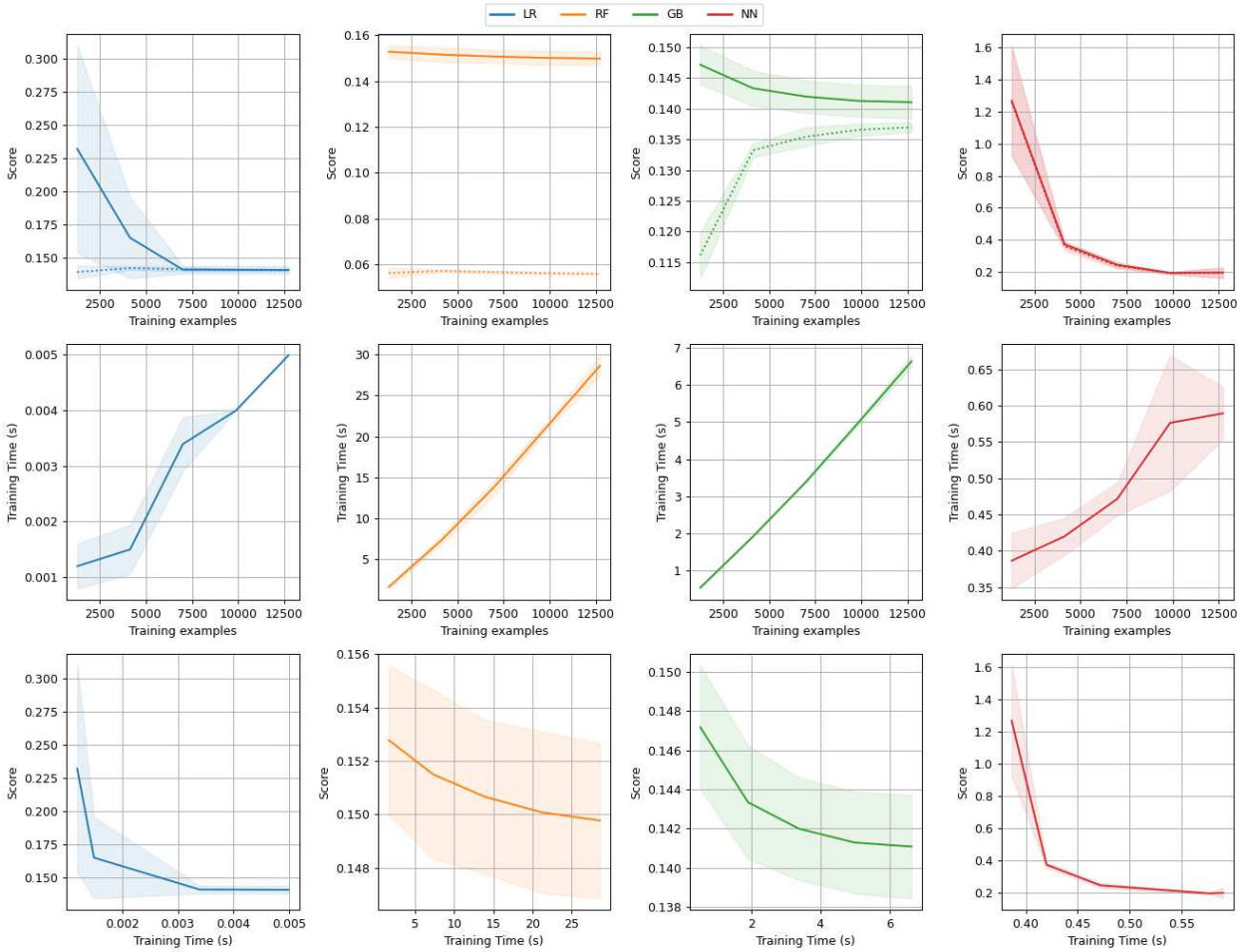


Figure 17: Learning curves of candidate models

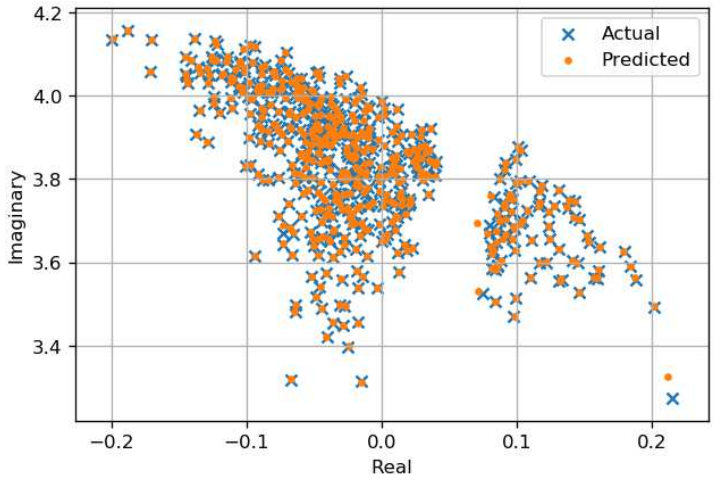
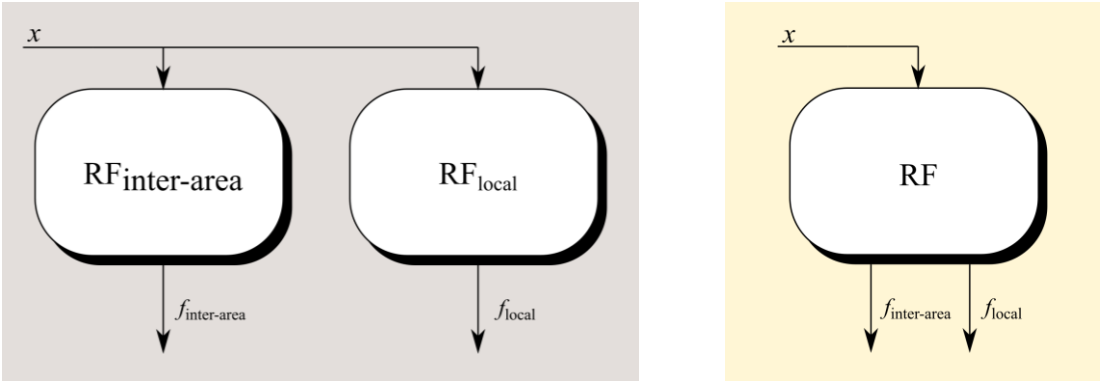


Figure 18: Inter-area mode actual vs RF-predicted.

Frequency Prediction of Multiple Modes

The above prediction RF model is providing estimates about the frequency of a single mode. However, in larger scale systems there are more than one inter-area and local modes. Therefore, in the following subsections two architectures of univariate and multivariate RF are compared with respect to their errors in estimating the frequencies of two modes. An illustration of the two architectures is shown in Figure 19. In the univariate approach an RF is trained for each of the targeted mode frequencies, for the two-area system these are the inter-area mode (see Figure 14) and the local mode between generators G3-G4 (see Figure 15). On the other hand, the multivariate RF aims to estimate the two frequencies simultaneously by optimizing the same parameters to achieve good performance in both targets.



a) Univariate *b) Multivariate*

Figure 19: Structures for predicting two frequencies in the two-area system

To obtain a more accurate approximation of the true error a 5-fold cross validation is used [165]. Apart from the MAE, the R^2 score is included to measure the performance of the models. Recall, that the R^2 score takes values from 0 to 1 depending on how well the model is performing on the given task. The two approaches are compared not only on the two metrics (i.e. MAE and R2 score) but also with respect to their CPU processing time.

The results are listed in Table 5 reveal that a trade-off exists among the prediction accuracy and processing speed. Logically, by training each RF for each frequency the error is lower, although the impact is more apparent in the predictions about the local mode. Regardless, by having a single model making two predictions at the same time has a faster processing time at the cost of slightly lower accuracy. Speed for control applications is more important given that the difference between the errors is very small. It should be noted that using parallel computing the processing time can be reduced.

Table 5: Performance on the test set for univariate and multivariate RF

	Inter-area		Local		Overall		CPU
	MAE	R^2	MAE	R^2	MAE	R^2	Time (s)
Univariate	1.4e-3	0.9883	5.0e-4	0.9942	9.5e-4	0.9913	3.9
Multivariate	1.5e-3	0.9882	8.0e-4	0.9886	11.5e-4	0.9882	2.7

IEEE 39-Bus System

The 39-bus system is a larger power system that is used for analysing the electromechanical interactions between different areas, depicted in Figure 20. In terms of developing ML models for larger scale systems it is necessary to evaluate their scalability. To consider a wide range of possible scenarios that will accurately capture the behaviour of the system the number of simulations will be higher than in small scale systems. In addition, in a realistic case partial observability is more likely than having measurements form every point in the system. For this reason, it is assumed that measuring devices are only available from bus 2 and bus 26 with red colour in Figure 20.

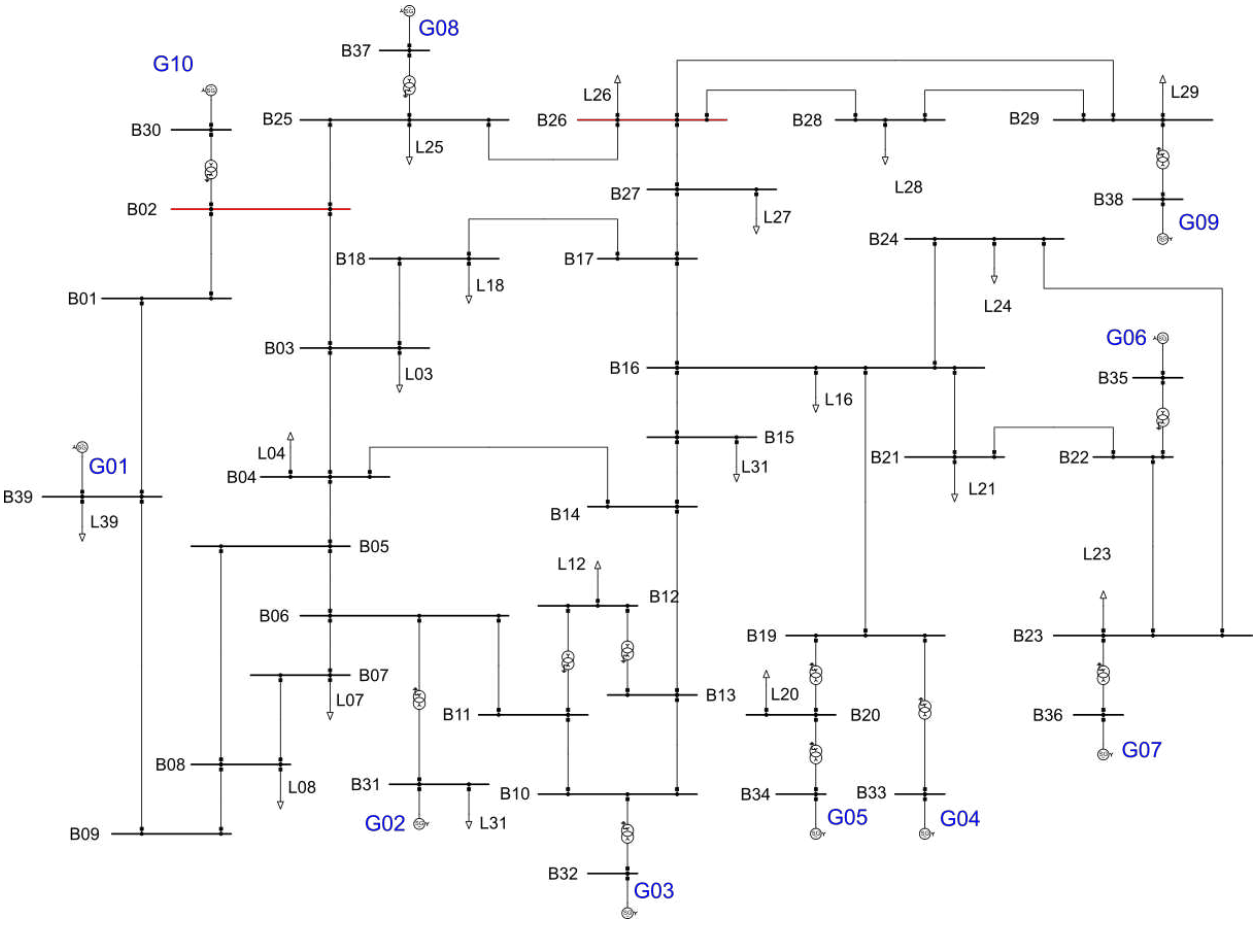


Figure 20: IEEE 39-bus test system

Table 6: Number of elements in 39-bus test system

Elements	Number
Buses	39
Lines	34
Transformers	11
Generators	10
Loads	20

Due to the larger number of elements in this system the database generation Algorithm 1 is revisited to improve the sampling of random operating points. Specifically, the process of computing the random demand of the loads now is computed using the Latin Hypercube Sampling (LHS) technique. The main reason is that with random sampling the required number of simulations to obtain sufficient and good data is much higher with larger power systems. In addition, the random points might not be unique, or they are very close to each other. On the other hand, the LHS generates the samples by dividing the parameter space a given number of evenly sized subsets [166] [167]. In two dimensions this can be thought of fitting a grid in the parameter space defined by two parameters, as shown in Figure 21. The 4 figures are generated by using random sampling (left) and LHS (right). Random sampling creates points that are not distributed in the full span of the parameter space, while the LHS creates a grid dividing the space into equal squares from which only one sample can be drawn per column and per row. In contrast to random sampling the size of the grid must be defined in advance hence the number of samples is fixed during simulations.

Algorithm 2 receives as inputs the matrix of the scaling factors $\mathbf{L}_s$ for each load and for each scenario. The $\mathbf{P}_{initial,}$and $\mathbf{Q}_{initial}$ are the default set points of loads active and reactive power demand. Dispatch scenarios $\mathbf{G}_s$ are also included in this algorithm that have been generated as well using the LHS method. The total number of converged simulations is close to 70000 and the final number of input features are 47 after removing the highly correlated variables and the ones that have low variance. The features include the voltage angles and the power flows at each feeder of these two buses.

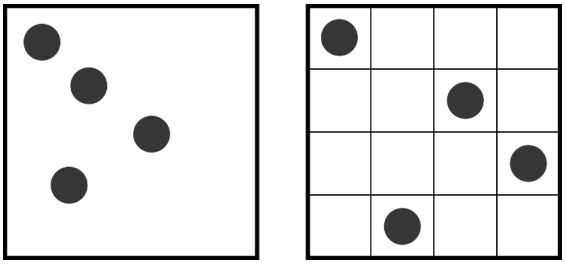


Figure 21: Random Sampling (left) and Latin Hypercube Sampling (right) example.

ALGORITHM 2: **DATABASE GENERATION WITH LHS**

```
input: Ls, Ds, Pinitial, Qinitial
for (scaling ∈ Ls) do
    PL ← Pinitial×scaling
    QL ← Qinitial×scaling
    for (dispatch ∈ Gs) do
        PG ← dispatch
        failed ← Calculate Power Flow ( .ComLdf )
        if failed then
            continue
        end if
        Calculate Modal Analysis ( .ComMod )
        Export Results
    end for
end for
```

Development of Prediction Models

Following the same procedure as in the previous section, the original dataset is split into train validation and test sets. The candidate models are trained to predict the inter-area mode between generator G1 with the rest of the system (i.e. G2-G10), where their learning curves are shown in Figure 22. The NN have the highest MAE, which might be due to the low number of epochs. The RF and GB perform the best although the former has a high variance[10]. The LR model has a comparable performance, which scales better as the size of the training set increases compare to RF and GB. Note that the size of the training set seems irrelevant to the error of the RF.

In contrast, to GB the cross-validation error does not reduce as more data are used in the training. Regardless, both models are able to approximate the true characteristics of the mode even though performance drops in cases when the frequency and damping deviates further from the mean, see Figure 23. These predictions are for the unseen test therefore their overall performance is promising, also considering that with fine-tuning the error can be increased further.

[10] Variance is the difference between the error in training and validation set

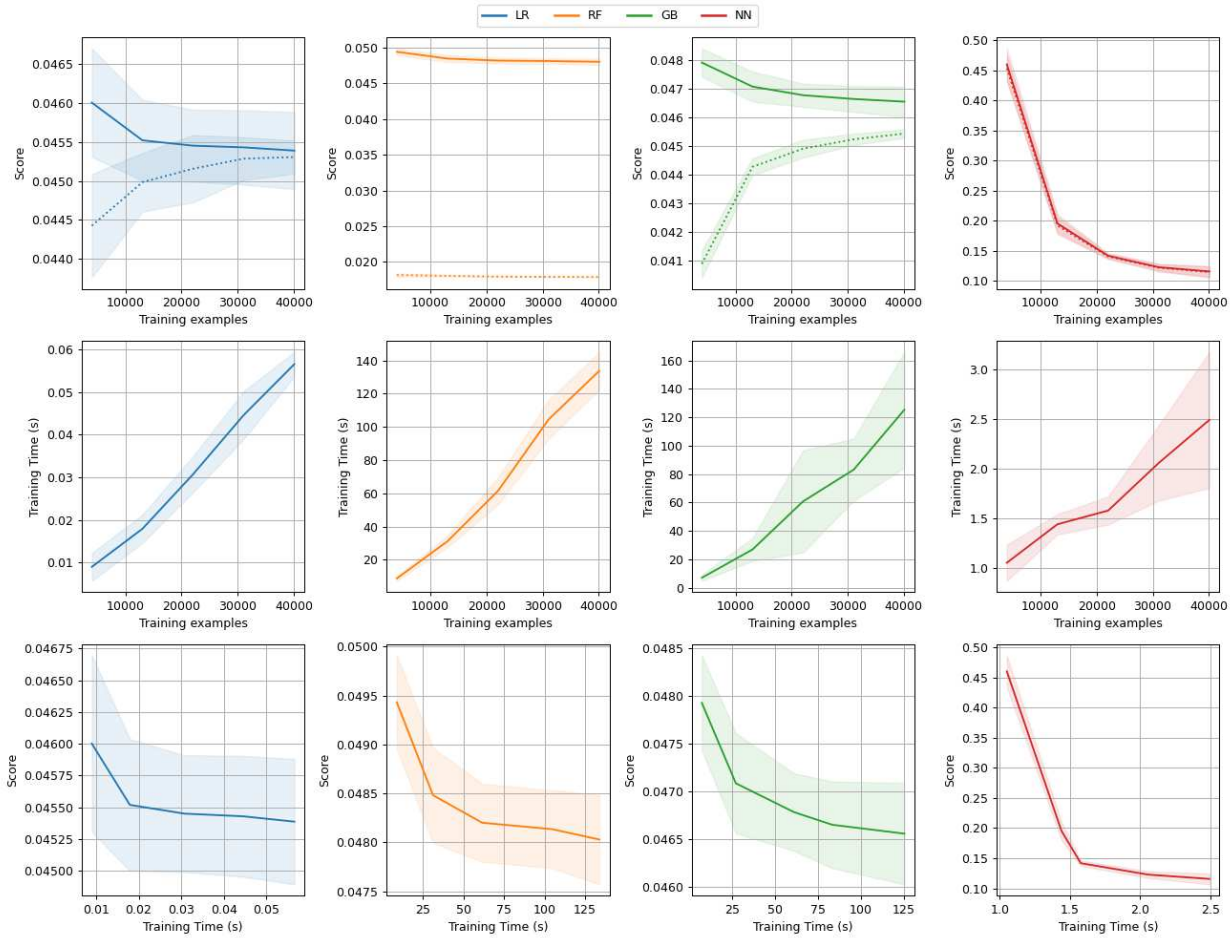


Figure 22 Learning curves for candidate models of 39 bus system

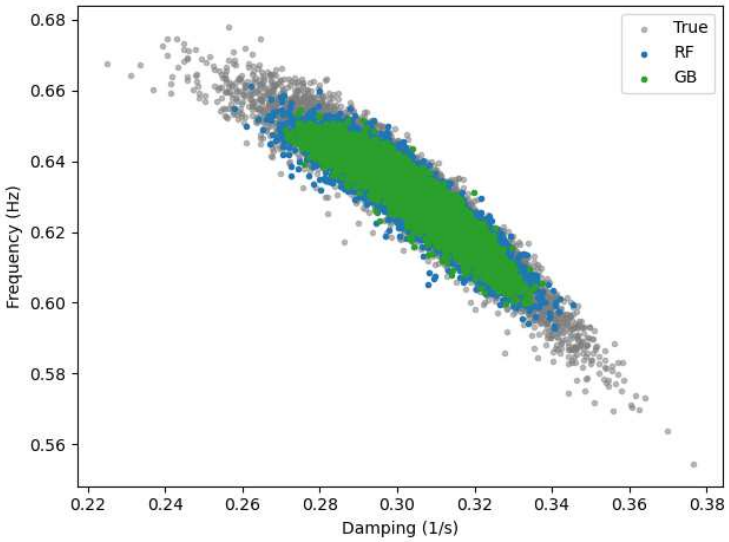


Figure 23: Inter-area mode frequency predictions

CHAPTER 5

Intelligent Power Oscillation Damper (iPOD).

The rotor angle stability of a power system can be modelled using machine learning given that the communication infrastructure and measurement devices exist. As discussed in Chapter 2 the applications of AI in rotor angle stability becomes more popular, especially with the advent of computational intelligence and resources as well as the transformation of power systems in SG. In the previous chapter it was demonstrated that it is possible to model the oscillatory characteristics of local and inter-area modes using a set of monitored system variables. Apart from monitoring, the information obtain from these models should be used to improve current control methods.

In the future, SG will experience faster and more complex dynamics as part of the wider integration of RES, BESS and power electronics interfaced resources in general. With regards to the rotor angle stability, damping of low frequency oscillatory modes will be reduce and new ones will emerge. To ensure the system synchronisation against different contingencies it is vital to plan for sufficient damping resources. In addition, these resources should always have the capacity to adapt to the ever-changing conditions in power systems to provide maximum damping.

Nevertheless, with the high penetration of power electronics and RES, many of the existing PSS will be displaced as a result of the decommissioning of old/carbon-based power plants [33]. Besides the main purpose, power electronic interfaced resources possess faster response (recall the absence of mechanical parts) than synchronous generators and therefore can support the system and its stability in case of a disturbance [168]. For instance, the control system of a grid-forming power converters can be based on either on droop [169] or on the virtual synchronous machine (VSM) [170]. With the former the converter can contribute in regulating voltage and frequency while with the latter it can emulate inertia or provide power oscillation damping [171]. In particular, the Synchronous Power Controller (SPC) is a popular approach for implementing a VSM [172].

Converters with SPC can provide virtual damping and synthetic inertia to the grid adding flexibility t to the system that can improve both local and inter-area stability [173], [174]. An SPC design with quasi-parallel power loop controllers (PLC) has been already presented in [175]. Specifically, the design based on the combination of a band-pass filter with a modified swing

equation can attenuate an oscillation at a specific frequency given that this frequency is known. In other words, the centre frequency of the bad-pass filter must match closely the frequency of the targeted mode centre. This information is not available by conventional means. Driven by the discussion in the previous section, here it is proposed to develop an intelligent Power Oscillation Damper (iPOD) with the ability to adaptively tune the centre frequency to provide maximum damping even when operating conditions change [176].

Control structure

Without focusing on the detailed control structure of a power electronic interfaced resource power plant, Figure 24 shows the iPOD receiving as input the active power deviation and sending to the voltage source converter the clean from the targeted frequency rotor speed deviation signal. The iPOD is comprised of two gains, a bandpass filter, the PLC and the ML model Figure 25, in this case this model is the RF from Chapter 4.

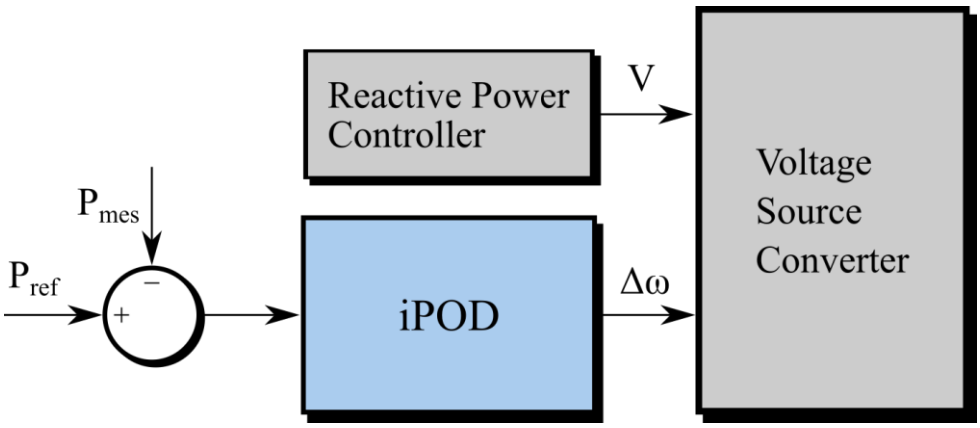


Figure 24: Overview of Power Electronics Interfaced Power Plant with iPOD

The PLC is essentially the swing equation with transfer function as in (5.1), where H is the inertia constant and D is the damping coefficient. However, this is not the only option; other alternatives can be easily adopted [168]. The use of a bandpass filter is supported by the argument that the centre frequency ω_c is known. Its transfer function is defined as in (5.2) where B is the passing band B. Furthermore, to remove the component of the input signal ΔP corresponding to the targeted frequency the gains k_1 is set to -1 before the signal goes into the PLC. The gain k_2 determines the amount of damping that the iPOD provides at the specified frequency. In contrast to the two gains the ω_c is updated by the predictions of the ML model.

$$G_{PLC}(s) = \frac{1}{Hs + D} \tag{5.1}$$

$$G_{BPF}(s) = \frac{Bs}{s^2 + Bs + \omega_c^2} \tag{5.2}$$

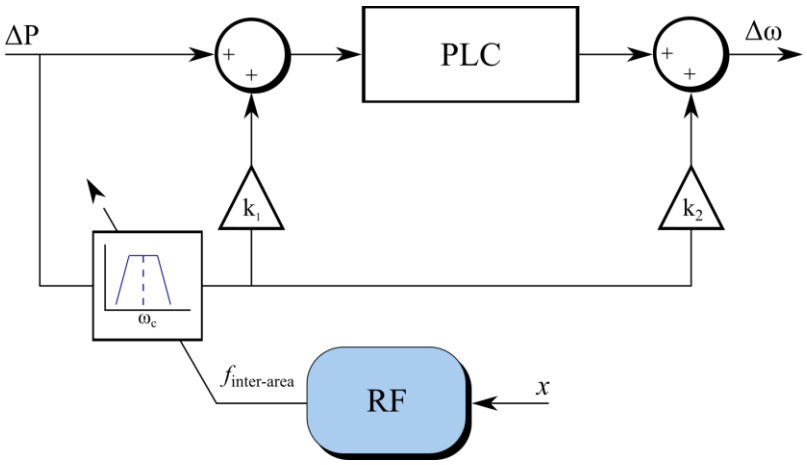


Figure 25: Control Structure of iPOD

The transfer function of the iPOD is therefore defined as in (5.3). As it can be seen the $G_{iPOD}(s)$ has two parts that correspond to the PLC and the integration of the bandpass filter with the gains. The second part can be reformulated as to resemble a lead-lag filter as in (5.4). Form this the significance of the two gains (k_1 and k_2) is revealed. Specifically, by tuning properly their values both the phase lag and the damping amount can be optimized.

$$G_{iPOD}(s) = G_{PLC}(s) + G_{BPF}(s)[k_1 G_{PLC}(s) + k_2] \Longrightarrow \tag{5.3.a}$$

$$G_{iPOD}(s) = G_{PLC}(s) + G_{BPF}(s) G_{ll}(s) \Longrightarrow \tag{5.3.b}$$

$$G_{iPOD}(s) = \frac{(2BHk_2 + 1)s^2 + (B + Bk_1 + BDk_2)s + \omega_c^2}{2Hs^3 + (D + 2BH)s^2 + (2H\omega_c^2 + BD)s + D\omega_c^2} \tag{5.3.c}$$

$$G_{ll}(s) = \frac{2k_2 HS + (k_1 + k_2 D)}{2Hs + D} \tag{5.4}$$

Single Mode Attenuation

Modified two area-system

A description of the two-area system has already been given in Chapter 4. However, to verify the proposed iPOD a few changes had to be made, see Figure 26. Specifically, at bus 7 a power electronics interfaced resource is connected through two step-up transformers. The new resource is equipped with an SPC and the iPOD presented above and has an apparent power equal to 100MVA. Note that the ML model used in the iPOD is trained using the dataset from the original system to track the frequency of the inter-area mode. Through trial and error, the gain k_2 is set to 0.018 while k_1 is equal to -1.

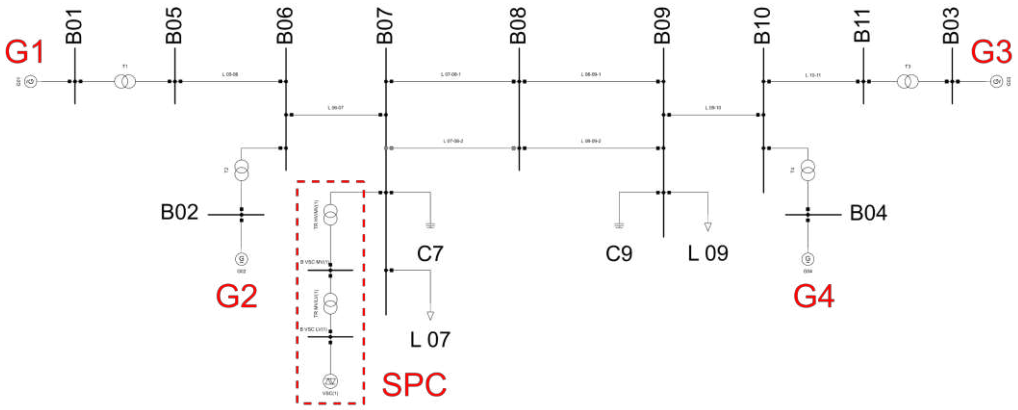


Figure 26: Modified two area system[11]

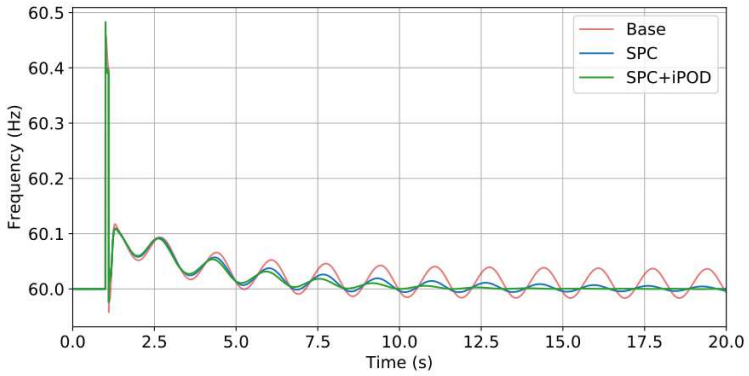
Verification

The proposed mode attenuation control scheme (i.e. the iPOD) is based on the premises that a ML model, like RF, can predict in real time the inter-area mode. Here the iPOD and the response of the system will be tested for a series of events. As a reference the same conditions and scenarios will be applied to the system for different case where the iPOD is out of service. It should be mentioned that communication delays are out of scope of this study and all inputs to the iPOD are in sync. This is mainly because in practical cases, a data processing module such as the Remote Terminal Unit (RTU) manage the exchange of information [177]. For instance, the IEC 61850 protocol support timestamping for ensuring that measurements can be grouped based on their time tags [178], [179].

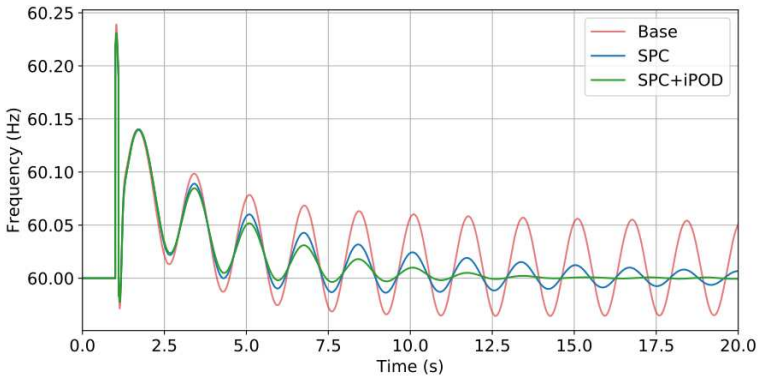
[11] Recall that the system is modelled in Powerfactory 2019 SP3.

System Response to Disturbance

A three-phase fault in the middle of the tie line will most likely excite the inter-area mode. For this reason, a fault is defined at the middle of the upper line connecting buses 7 and 8, denoted $L_{7\text{-}8\text{-}1}$ for a duration of 100 ms. In addition, to evaluate the contribution of the RF model to tune the centre frequency of the bandpass filter correctly, the loads at bus 7 and 9 are set randomly using the same procedure as in lines 4-5 of Algorithm 1. The system's frequency response is plotted in Figure 27 for buses 6 and 10 that correspond to areas 1 and 2, respectively. From these results it becomes apparent that the SPC has already the capability to provide damping to the system [180]. However, as the SPC is not targeting a specific frequency but rather a range of frequencies the amount of damping is much less that then one provided by the iPOD. The random variations of the system's operating conditions affect the frequency and the damping of the inter-area mode to vary in both frequency and damping. This is depicted in Figure 28 where the real and imaginary components of the inter-area mode vary along for random operating points. As expected, the SPC can improve the damping of inter-area but not as much as the iPOD. The accuracy of the RF estimations are reflected by the increased damping of iPOD.



a) Bus 6



b) Bus 10

Figure 27: Frequency response during a fault in line $L_{7\text{-}8\text{-}1}$ for a random operating point

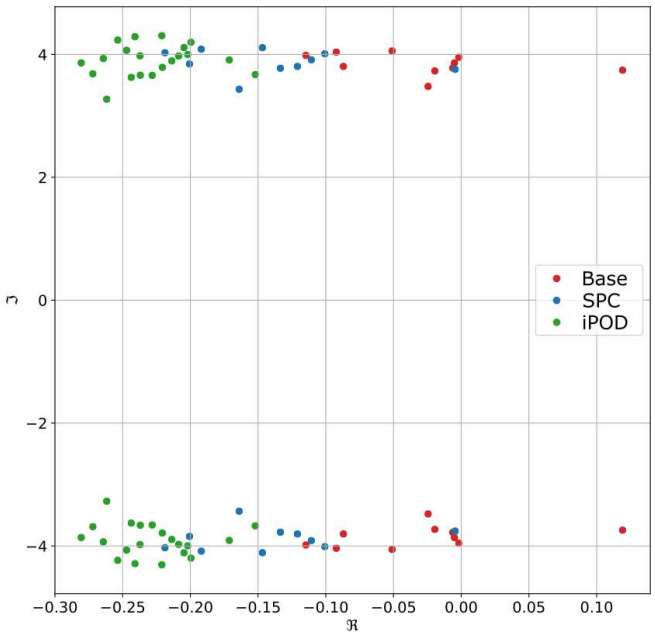


Figure 28: Modal characteristics for random loading levels

Oscillation Damping Impact of iPOD to Synchronous Generators

Area 2 is importing energy from area 1 due to the higher demand of load at bus 9. This dependence to area 1 become more apparent by examining the power output of each generator. To do so the same fault is applied but a new random point for the system loads is generated. As expected, the impact of the fault in generators of area 2, shown in Figure 29, is much more severe than in area 1. Nevertheless, with the iPOD the oscillation converges to a steady state faster and in addition the maximum power is lower compared to the base case. The performance of the system with the iPOD is interesting because the nominal capacity of the power plant is approximately 85% less than the synchronous generators.

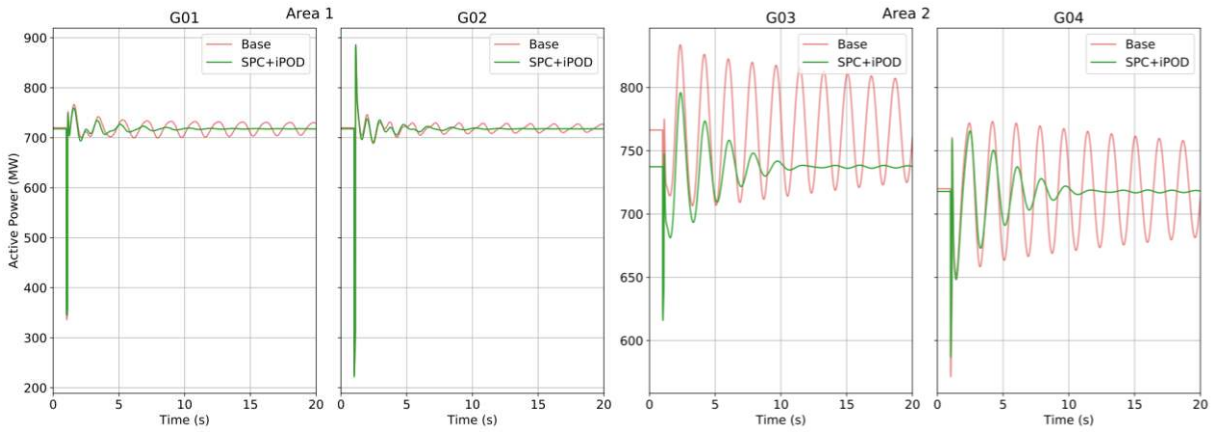


Figure 29: Active power output of generators after a fault for a random operating point

iPOD Comparison with PSS

Among the LFO damping solutions, the PSS is very popular. Therefore, it is logical to compare the performance of the iPOD individually with the PSS in the system. The tuning of the PSS parameters is done following the procedure in [181]. Like the previous scenarios, random operating points are generated, and the modal analysis of the system is calculated for each control case and random point. The average inter-area mode characteristics are listed in Table 7. The PSS of generators 1 and 3 provide (on average) the highest amount of damping

Clearly, PSS1 and PSS3, on average, possess the highest contribution towards inter-area power oscillation damping. On the other hand, the iPOD is in the middle meaning that it dominates the contributions of PSS2 and PSS4 but cannot surpass the damping of PSS1 and PSS3. Regardless, the average damping ratio is higher than 5%, which is acceptable. Perhaps increasing the capacity of the iPOD power plant this amount can be increased.

Table 7: Average Interarea mode characterises for each control case

Case	Frequency (Hz)	Damping (1/s)	Damping ratio (%)
Base	0.613	0.019	0.50
PSS1	0.604	0.289	7.60
PSS2	0.614	0.022	0.58
PSS3	0.582	0.359	9.79
PSS4	0.617	0.061	1.59
iPOD	0.659	0.229	5.55

Multiple Mode Attenuation: Multi-band iPOD

The iPOD presented above can provide additional damping for a single mode by adapting online a single parameter, the centre frequency of the bandpass filter. An extension is the Multi-band intelligent Power Oscillation Damper (MiPOD) that can increase the damping of two oscillatory modes instead of just one [182].

Control structure

The main variation of the control structure of the MiPOD is that the two gains have been replaced by a frequency selective POD, that has two band-pass filters and PLCs with the swing equation. Specifically, the swing equation is defined as in (5.5) while the transfer function of the bandpass filter is given by (5.6) where ζ_p is the damping ratio while subscript p indicates to which block the transfer function is referring to i.e. the zero refers to the main PLC block1 and 2 for the other two.

The MiPOD exploits the fact that regulating the virtual inertia the amount of damping to LFO can be controlled. The decoupling of the two frequency bands is achieved by feeding back the output of the bandpass filter of the first to the second and vice versa as defined by the transfer functions (5.7) and (5.8). This decoupling network is important not only for each damper to act on the specific frequency but also for the tuning procedure.

$$G_{PLC,q}(s) = \frac{1}{2H_q s + D_q}, \qquad q \in \{0,1,2\} \tag{5.5}$$

$$G_{BPF,q}(s) = \frac{2\zeta_q \omega_{c,q} s}{s^2 + 2\zeta_q \omega_{c,q} s + \omega_{c,q}^2}, \qquad q \in \{1,2\} \tag{5.6}$$

$$G_{p,1}(s) = \left(1 - G_{BPF,2}(s)\right) G_{BPF,1}(s) \tag{5.7}$$

$$G_{p,2}(s) = \left(1 - G_{BPF,1}(s)\right) G_{BPF,2}(s) \tag{5.8}$$

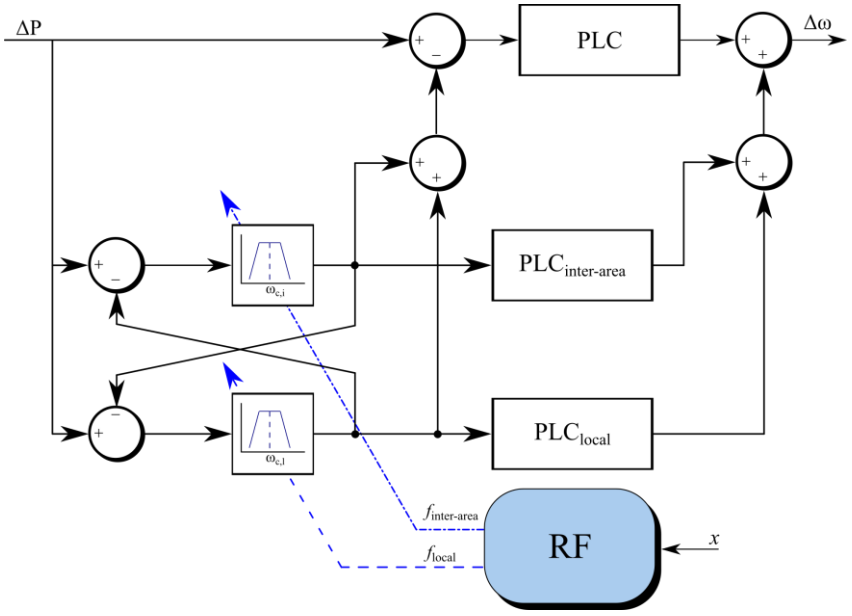


Figure 30: Control Structure of MiPOD

Modified two-area system

For implementing the MiPOD, two RES power plants are connected to bus 10 of the two-area system Figure 12, through a virtual synchronous power converter. The total apparent power of the plant is 200 MVA, around 6% of the total apparent power for the four synchronous generators. The reason of connecting the RES plants at bus 10 is for attenuation apart from the interarea mode, the local mode between generators G3 and G4.

The connection of the RES power plants increases the interarea and local mode frequency from 0.55 to 0.61 Hz and from 1.10 to 1.14 Hz respectively. The MiPOD is equipped with the multivariate RF model described in Chapter 4.

Verification

Like in the case of iPOD, to emulate the changing conditions of power systems, the active and reactive power of the two loads are scaled up and down using lines 4-5 from Algorithm 1. In the following sections the MiPOD performance is evaluated with respect to the base, the SPC and iPOD cases. The contingency events occur all at FOT=2s, with a sampling rate of 1kHz and a total simulation duration of 20 seconds.

Modal analysis for random operating points

Around 100 random active-reactive power parameters were generated for each load. Modal analysis computed the small signal characteristic of the system focusing on the inter-area and local mode. The Gaussian Kernel density estimation for each case and for the damping and A_1/A_2 ratios are illustrated in Figure 31. The area under the curve corresponds to the probability, which equals 1. By comparing the area defined by each curve and the vertical lines conclusions can be made about the different cases. Specifically, with the MiPOD the probability of new random operating points having a damping ratio above 5% (or A_1/A_2 ratio above 2) is much higher than the SPC or base case. Considering the performance of iPOD, it is still better than SPC however it cannot surpass the proposed MiPOD.

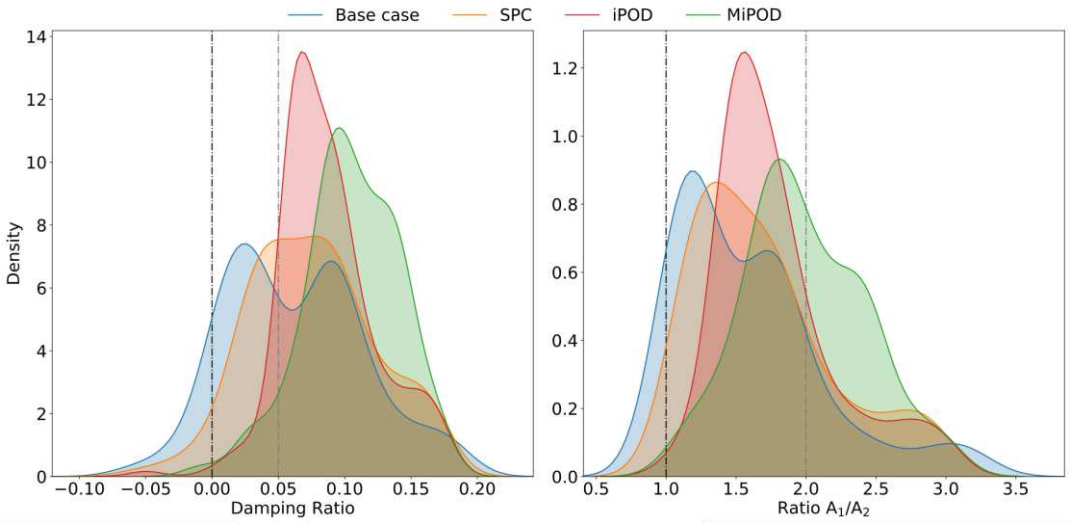


Figure 31: Probability density functions of damping and A_1/A_2 ratio for all random points

Symmetrical short circuit

A three-phase fault is defined at bus 7 with a duration of 100ms. The same event has been repeated for more than 30 random operating points. The frequency and active power response at bus 8 is recorded and plotted in Figure 32. For brevity only two operating points are shown in this (and the following figures). From these responses the contribution of the MiPOD to damp LFO is clear. The system reaches an equilibrium faster that in the other two cases. However, there is a significant frequency overshoot due to the severity of the fault that is present in all cases. Most likely such frequency increase would have caused the tripping of protective devices. Nevertheless, with MiPOD this overshoot follows the same trend, but it is smaller.

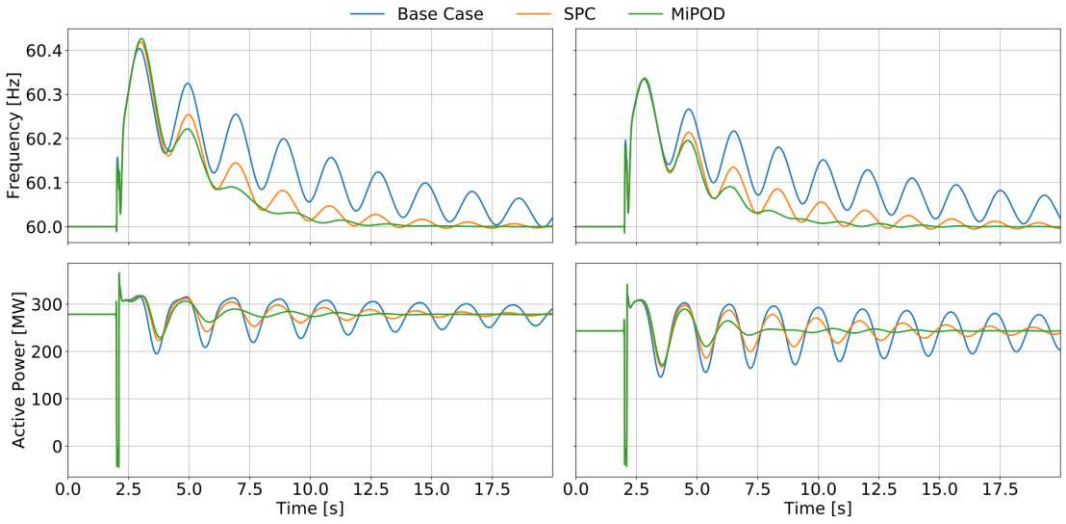


Figure 32: Frequency and Active power at B08 under a fault at B07

Step change of load active and reactive power

One of the most common disturbances in power systems are load variations. In power system analysis these variations are represented by either a ramp or step change. Both types affect the power balance albeit in a different way. For instance, a sudden load variation (i.e. step change) has a more severe impact in the stability of the system. Therefore, a 5% increase of the reactive power of load 9 is defined and repeated for 30 new random operating points. The system response is shown in Figure 33. Although the oscillations in the frequency have very small amplitude the impact of this event is present in the active power flowing through bus 8. Once again, the system oscillations for a smaller period and the new steady state point is closer to the synchronous speed.

The same responses are plotted in Figure 34 for a 5% increase of the active power demand of load 9. The response of the system when the MiPOD is connected is smoother, with lower amplitudes and a faster convergence to equilibrium.

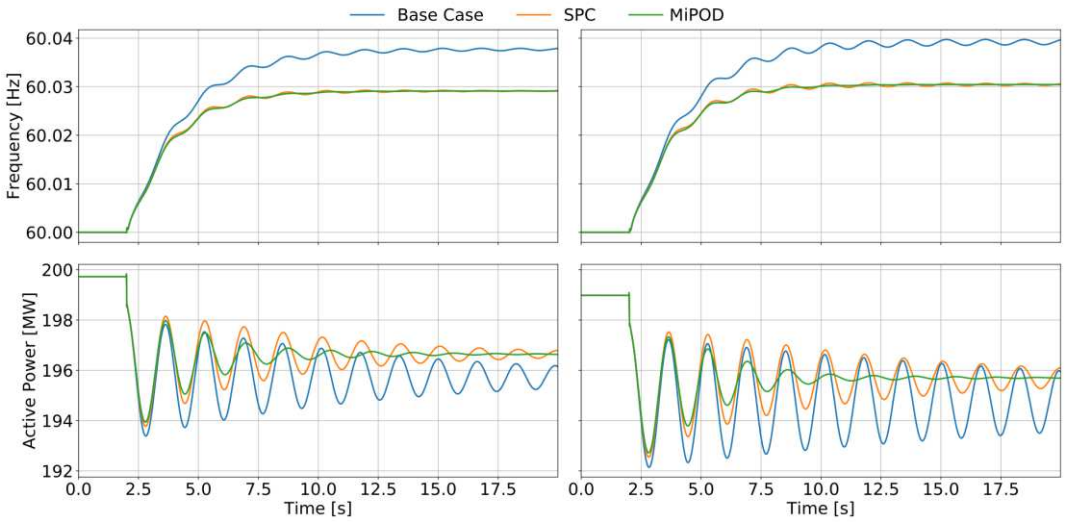


Figure 33: Frequency and Active power at B08 after an increase of L9 reactive power.

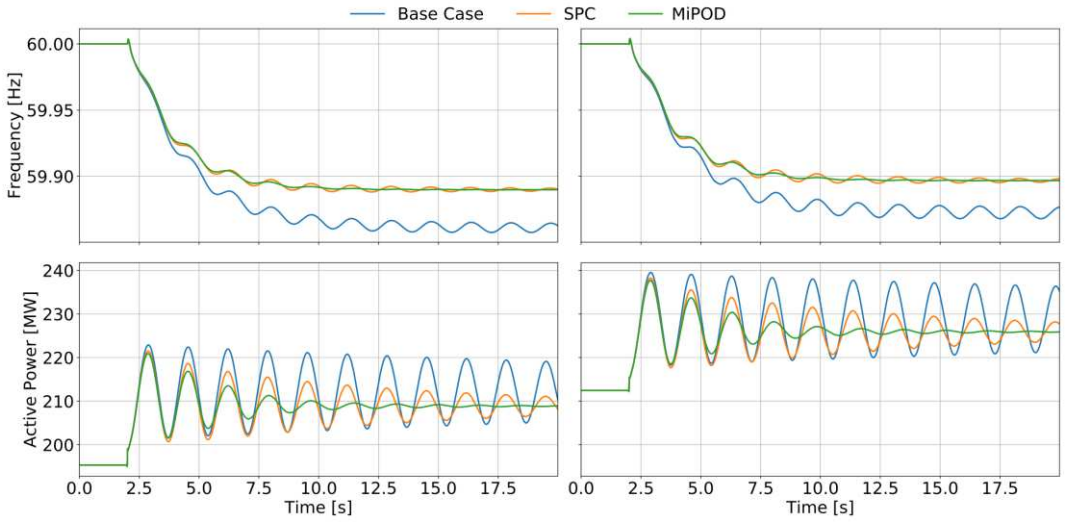


Figure 34: Frequency and Active power at B08 after an increase of L9 active power

Synchronous Generation Event

The local mode in the two-area system is already adequately damped. Attempting to excite the local mode G3 experiences a 0.1 per unit increase of its mechanical input torque while an equal is reduced from G4. The component of the local mode appears in the frequency of bus 8 but for a very brief period, Figure 35. Overall, the inter-area mode dominates the local mode for the rest of the simulation.

For the base case this event leads the system into instability as the active power oscillations with increasing amplitudes. Even for the SPC case the system seems to be in a critical state.

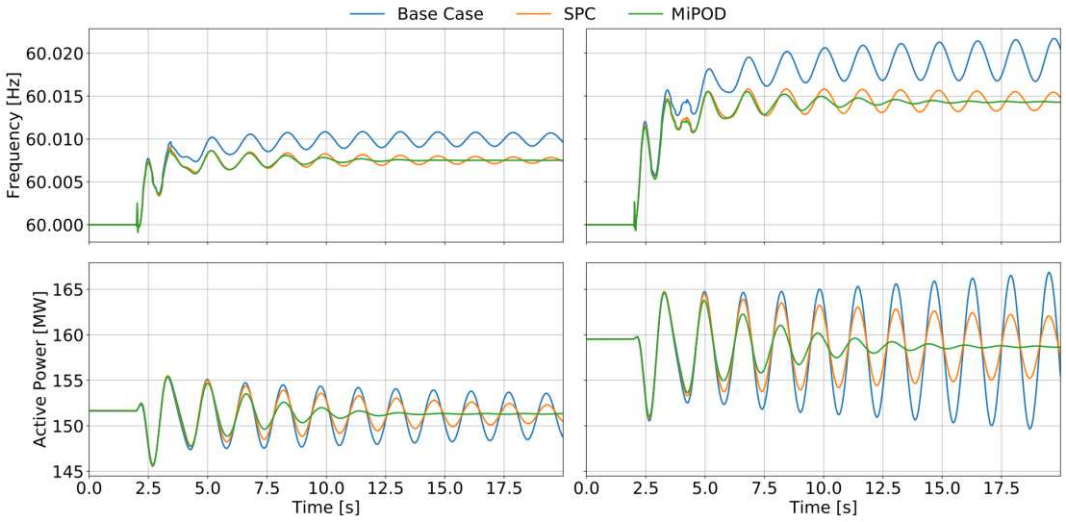


Figure 35: Frequency and active power at B08 after a synchronous machine event

Variation of Network Topology

The same three phase fault is applied however, the upper transmission line connecting buses 8 and 9 is out of service. As a result, the rest of the system is under high stress meaning that the modes of the system will move closer to the stability boundary. This contingency event aims to evaluate the MiPOD under the most severe case but also its ability to adapt in situations that the predictions model hadn't been trained before. From the results shown in the MiPOD is able to adapt to these new conditions and also to provide optimally additional damping Figure 36.

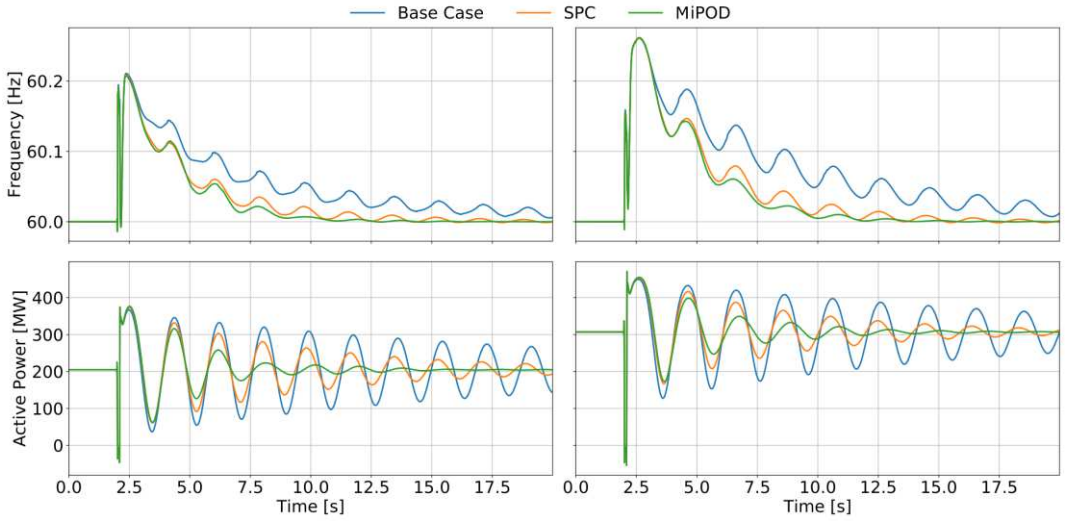


Figure 36: Frequency and active power at B08 under a fault at B7 with line $L_{8\text{-}9\text{-}1}$ out of service

CHAPTER 6

Gas Turbine Power Plant Adaptive PSS Tuning

To date, the PSS has been the main solution to LFO due to their simplicity and cost-effectiveness. Briefly, by controlling the excitation of a synchronous generator it is possible to create a damping torque that is in phase with the input signal deviation (usually the rotor speed) such that the overall damping of the system is improved. If the governor is fast enough this torque component can be provided by the governor-turbine system [183] although the majority are excitation-based PSS. There are many variations of the PSS ranging from the classical PSS1 controller to the more advanced designs of the PSS2B and PSS4B controllers [184]. In its simplest form the single input-output PSS consists of a gain (K_{pss}), a low-pass (T_l), a washout (T_w)and lead-lag filters as depicted in Figure 37. The low pass filter eliminates any high frequency components in the input signal while the washout filter removes the DC component [185]. The lead-lag filter creates the required phase compensation to attenuate the LFO in the signal. The tuning of the PSS parameters (e.g. K_{pss}, T_{lp}, T_w ,T_1 ,T_2) is typically done when during the installation of the PSS and afterwards only when it is absolutely necessary to be updated [186]. Improper parameter settings can lead to the PSS providing minimum damping and even amplify oscillations. In future SG, the oscillatory characteristics of the system will vary more frequently and as a result the possibility of updating the settings of the PSS adaptively is becoming a popular topic in the literature [185].

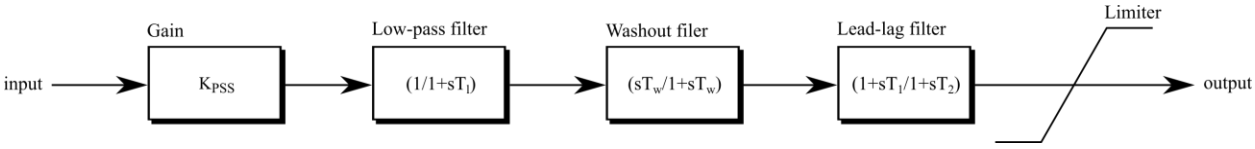


Figure 37: PSS1A structure

The iPOD presented in the previous section makes use of AI to monitor specific oscillatory modes' frequency to tune its parameter for maximum damping. This adaptive tuning however could be useful with the conventional damping devices, such as the PSS. For this reason, in the FLEXITRANSTORE project an AI model is developed to estimate the frequency of an interarea mode in an actual power system in Europe. This information will be used to compute the proper set of parameters for the lead-lag filter (s) of a PSS installed in a Gas-Turbine Power Plant (GTPP). Sensitive information with regards to the power system or the GTPP will be kept private.

Development of Prediction model

To create the required dataset Algorithm 2 is used. In addition, this algorithm has been repeated for a list of probable N-1 cases leading to a total number of simulations around 600000. After analysing the system states with respect to the system eigenvalues, the inter-area mode frequency that the GTPP has the highest participation is around 0.85Hz. The maximum and minimum values that have been recorded from the simulation results have indicated that the frequency of the mode does not deviate much from the mean, i.e. min = 0.81 Hz and max = 0.92Hz. The monitored system variables include voltage magnitudes, active and reactive power flows measured within the area of influence of GTPP and form a selected buses and lines. The linear correlation of these variables, shown in Figure 38, reveal that in general the correlation of the dataset is relatively low. The high correlation corresponds to variables from measurement points that are close to each other and mainly for the voltage magnitude. Nevertheless, all variables will be used as inputs to the AI model. The train, validation and test subsets are form by random selecting sample from the original dataset with a size ratio of 75%/15%/15% for each set respectively. A min max scaler is fitted on the train set to scale the inputs between the interval (-1,1). From a preliminary analysis the DNN can most accurately estimate the mode inter-areas mode frequency. Therefore, the DNN is selected to be finetuned and finally used in the adaptive tuning of the PSS.

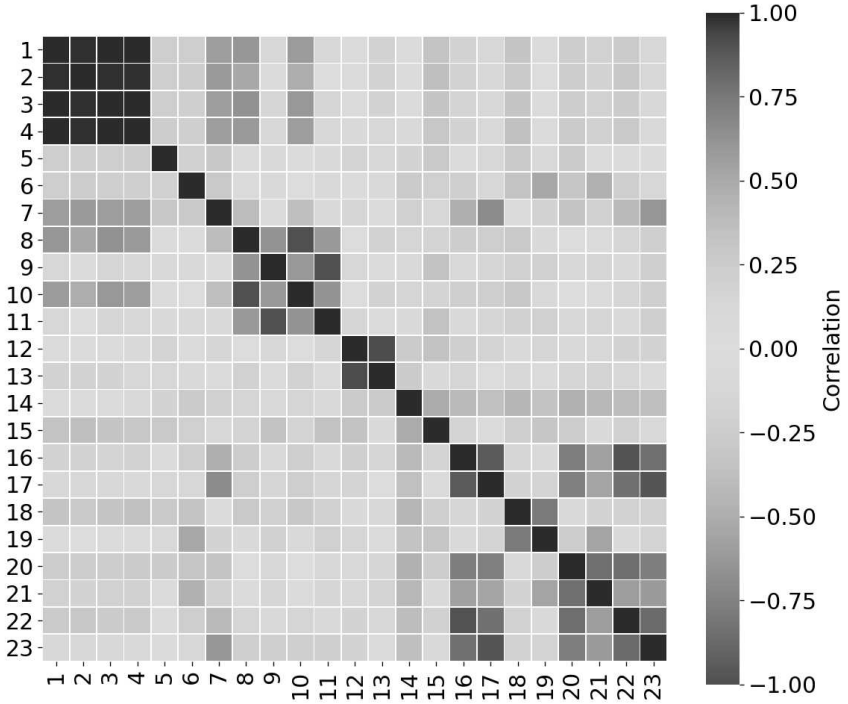


Figure 38: Demonstrator System Variables Correlation Plot

The DNN were trained using the TensorFlow Python library with the Sequential API from Keras. Several architectures fully connected architectures are tested with respect to the number of layers, number of neurons per layer and activation functions. In addition, different optimizers are used with different learning rates as well as different loss functions. All models are trained for 50 epochs however an *earlystopping callback* is included to stop the training when the models reach a plateau and do not improve their error for 3 consecutive epochs.

The parameters of the top 7 DNN models are listed in Table 8. The optimizers used are the Gradient Descent (GD) and the Nesterov Adaptive Momentum (NADAM) with different learning rates. The batch size is equal to 32 patterns while the remaining parameters (i.e. the momentum) are the default ones used by Tensorflow. The MAE on the validation set per each epoch is plotted in Figure 39. Models with the GD optimizer have higher error than the ones using the NADAM. This is common as GD often reach the local optima while NADAM is more robust in escaping them [153]. This is reflected both in the validation set MAE but also in the test set error. The model with the Huber loss function achieved the lowest loss value, however the MAE is not the lowest. In contrast, the best model uses the NADAM with a small learning rate and the MSE loss function.

Table 8: DNN models

	Parameters						Test Set
ID	$\sigma(\cdot)$	Layers[12]	Neurons	Optimizer	λ	LOSS	MAE
1	ELU[13]	2	50	GD	0.01	MSE	2.3e-2
2	ReLU	2	50	GD	0.01	MSE	2.6e-2
3	RELU	2+1	50+25	GD	0.01	MSE	2.1e-2
4	RELU	2+1	50+25	NADAM	0.001	MSE	1.6e-2
5	RELU	2+1	50+25	NADAM	0.001	HUBER	1.6e-2
6	ReLU	3	50	NADAM	0.0005	MSE	1.5e-2
7	ReLU	3	50	NADAM	0.0001	MSE	1.7e-2

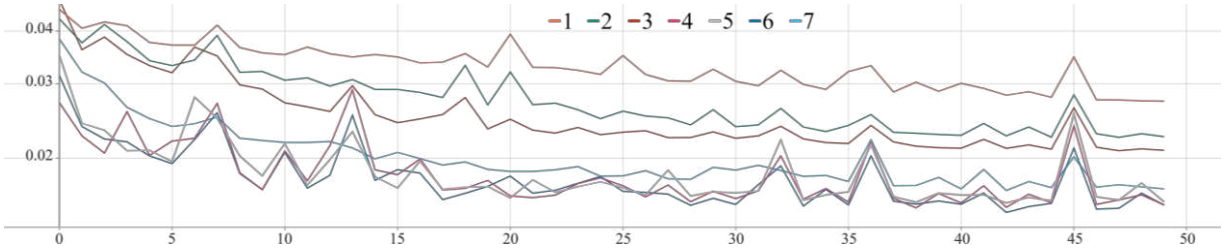


Figure 39: MAE on the validation set for each epoch for the models in Table 8.

[12] This column refers to the hidden layers of the network
[13] Exponential Linear Unit

Analysing the errors of the models on the test set it is possible to identify the limitations of the existing models. To illustrate models 3 and 6 are consider for their errors. Specifically, Figure 40 shows the predicted mode frequency (left) and the residuals (right) as a function of the true mode frequency. Using the zero error predictions as a reference, the lowest error of the model 6 is apparent. However, there is a pattern in these plots that reveal that both DNN are having difficulties on explaining properly the frequency. The red ellipsis emphasizes this issue by focusing on the clear error patterns. Revisiting some of the previous steps (i.e. scaling, sampling, feature engineering) might be able to solve this issue and reduce even further the performance.

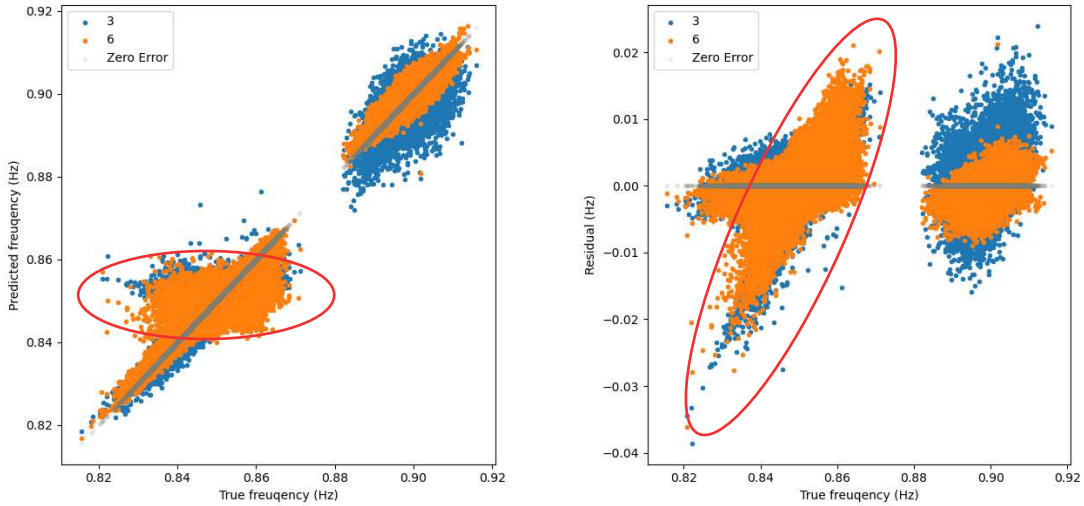


Figure 40: DNN models error analysis

Adaptive Tuner

The measurements are transmitted to the adaptive tuner so that the DNN can provide the estimation about the inter-area mode frequency. This estimation is then used by the tuner algorithm to compute the values of the time constants for the phase compensation block with the two-lead lag filters of the generators PSS, depicted in Figure 41. The adaptive tuner must operate in real-time so that the time constants are update regularly. To do so, the DNN and the tuning algorithm must be hard coded in a dedicated controller that will communicate with PSS using a suitable communication protocol. Recall that the DNN is developed using Python, which although is supported by the controller it is preferred to migrate from the high level programming language to the low level Structure Text (ST) language using TwinCAT because it is more robust and suitable for this application. Once trained, the DNN is represented by the set of optimized weights and biases of each layer. Importing these parameters is straightforward, as it is the implementation of the NN principal operation along with the ReLU activation function.

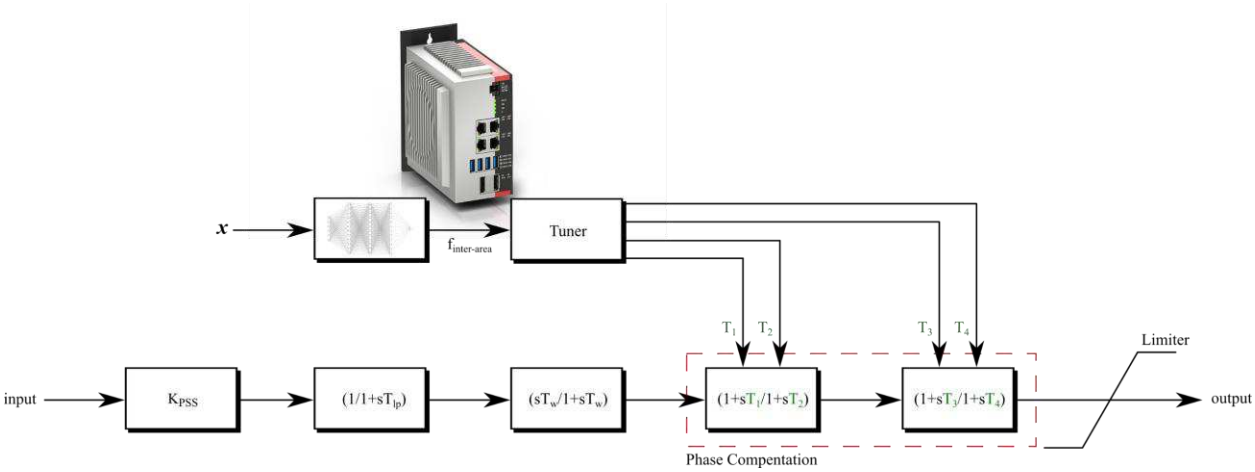


Figure 41: Overview of adaptive tuner

Hardware-in-the-loop simulations

The integration of the adaptive tuner controller in the real GTPP is tested using hardware in the loop simulations. Specifically, the power system and the power plant are modelled in real-time simulations and the communication systems and protocols are replicated fully. The controller receives the measurements from the area of influence, the adaptive tuner estimates the mode frequency and computes the new time constants that are transmitted to the GTPP PSS device. A graphical representation of the above architecture is presented in Figure 42. The different line colours indicate the different communication protocols, which have been omitted form this figure.

The FLEXITRANSTORE project is expected to end during 2022 and the next step for this demonstration is the deployment of the adaptive tuner in the GTPP site. The results from the field experiments will provide invaluable information about the effectiveness of the DNN to contribute in an actual system by improving the performance of the widely used PSS. Through FLEXITRANSTORE it will be possible to identify potential limitation of AI that cannot be seen in a fully controller environment that is in simulations.

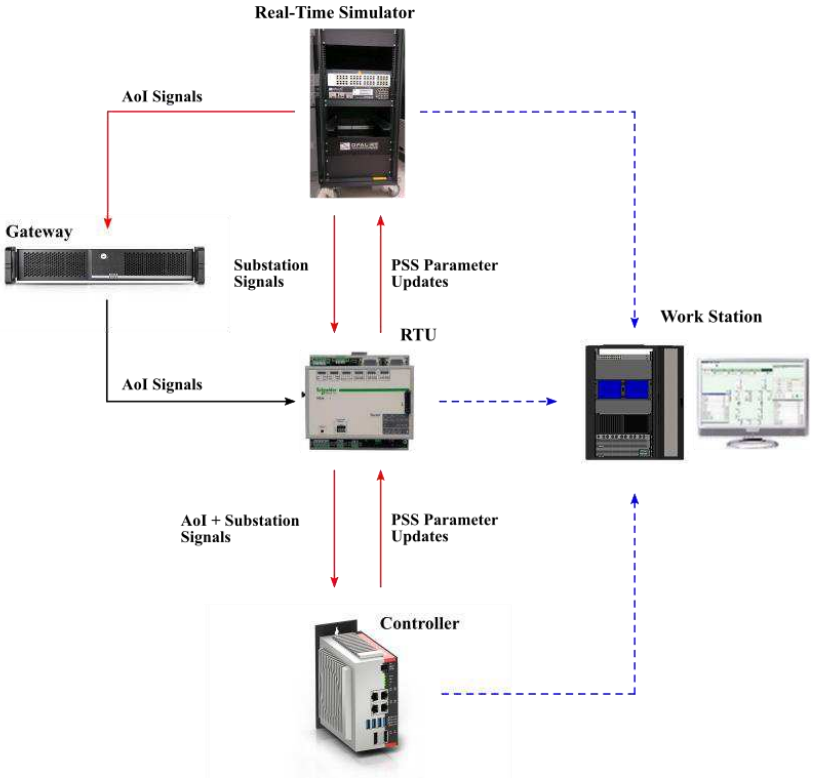


Figure 42: Hardware-in-the-loop architecture for simulations

CHAPTER 7

Conclusions

The energy transition necessitates the transformation of power systems into Smart Grids able to support decarbonisation, decentralization, and digitalization. The wide integration of carbon-neutral resources and the electrification of traditionally fossil-based sectors increase the complexity of the system at all levels. The newest classification of power system stability introduces additional branches for addressing specifically the issues related to the increasing presence of power electronics, which act as interfaces for the renewable energy resources. Apart from the uncertainty in both generation and demand, the absence of mechanical parts lowers the overall system inertia causing the dynamics to evolve much faster. The decarbonisation of power generation will lead to the displacement of conventional solutions for ensuring the stability of power systems. In fact, the higher penetration of power electronics has a deeper impact to the systems *Rotor Angle Stability.*

The digitalization of power systems with a state-of-the-art communication infrastructure and advanced metering, increases the volume, velocity, and variety of data in power systems. Artificial intelligence algorithms are becoming a popular alternative to complement conventional methods (or even surpass them) in monitoring, predicting, or controlling the stability of power systems. As shown, studies have developed solutions using a wide range of machine learning algorithms demonstrating high performance. Many of the related works in the literature, assume favourable conditions in terms of data availability (full observability) and communications (very high sampling rates) to name a few.

In this thesis, popular machine learning algorithms such as linear regression, neural networks, gradient boosting, and random forests are trained to estimate the modal characteristics of electromechanical interactions in a small and medium scale power system. The results show that these data-driven solutions can accurately approximate the changes of the modal characteristics as system conditions change. Apart from the monitoring purposes this information is proposed in this thesis to be used to increase the attenuation of underdamped oscillations. Once trained these models are fast enough to update near real-time the parameters of a controller to target the critical oscillations.

Power electronics reduce the damping of existing oscillations and introduce additional ones however, their fast response and flexibility have the potential to actively contribute in supporting the power system stability and alleviate certain challenges. The thesis presents the intelligent Power Oscillation Damper (iPOD) that combines the capability of power electronics to emulate the synchronous behaviour of conventional generators and the modelling capabilities of artificial intelligence. Specifically, the iPOD treats the frequency of the oscillations to be a known parameter. This is provided by the very accurate machine learning models developed to monitor the modal characteristics of the system. An extension of the iPOD to target two modes instead of one is the Multi-band intelligent Power Oscillation Damper (MiPOD). Driven by the assumption that the frequencies of the oscillations are known the MiPOD implements multiple rotors with a decoupling network to oscillate the attenuation at the specific frequencies. The integration of artificial intelligence permits the adaptation of the iPOD and MiPOD to the changing operating conditions that affect the modal characteristics.

The results presented in this thesis demonstrate the effectiveness of the proposed controllers. Simulations of various contingency scenarios for random variations of the operating conditions show that carbon-neutral power plant with the proposed controllers can provide enough damping to under-damped oscillations as to ensure that the system will remain stable. Note that the nominal capacity consider in the study is much lower than the conventional synchronous generators. In addition, the adaptive parameters for both iPOD and MiPOD are fewer for the Power System Stabilizers.

The applicability of machine learning in monitoring and control is investigated by the FLEXITRANSTORE project, which has supported this thesis. Specifically, following the above rationale, a deep neural network is developed to monitor the electromechanical interaction of a gas-turbine power plant for a system in Europe. The trained model is embedded into a controller along with a tuning algorithm that computes the optimal set of parameters for the phase compensation block according to model's predictions. The controller communicates with a wide area measurement system that provides with measurements within the area of influence of the plant and with the PSS device installed in the power plant to update its parameters near-real time and increase the damping of the system. The update rate is closely related to the sampling rate of the wide area measurement system. This project is still ongoing and by the year 2022 with field testing being the next step. The applications of artificial intelligence are restricted to simulation

platforms; therefore, this project is a unique opportunity to obtain invaluable results about the effectiveness, limitations, performance of the application of neural networks in real life conditions.

Future work

The iPOD and MiPOD do not require any knowledge about the type of the carbon-neutral resources behind the power converter. However, the prime mover can have an impact of the overall performance of the power plant. In addition, renewable energy source power plants can range from small scale to larger, connected at either the distribution or the transmission system. The impact of the capacity of the power plant to the performance of the iPOD will be further investigated.

The performance of the data-driven approaches, used in this thesis for monitoring and control purposes, depend on the quality and quantity of the data. In contrast to other fields, in power systems the required data are either unavailable, unlabelled or have low quality. Hence, equivalent system models are used to gather the necessary information. However, the design of the database generation process can create biased models, which can lead to an overestimation of its performance during evaluation. As a result, it is necessary to research further into their robustness and reliability. This is particularly important for black box models.

Towards that end, large scale, and actual applications, which currently are limited, can reveal valuable insights. For instance, the data generation process scales poorly with the size of power systems. Therefore, other alternatives will be investigated to improve the scalability of this models. Transfer learning appears to be a promising solution. Briefly, transfer learning is a human-inspired concept that encapsulates a wide spectrum of methods and techniques to make use of trained models for a given task and a given domain to aid the learning of another task that most likely exist in a different domain [187]. It has been shown that transfer learning can improve both scalability and adaptability [188].

Bibliography

[1] U. Nations, “Adoption of the Paris Agreement,” 2015.

[2] Interanational Energy Agency, “Data & Statistics.” https://www.iea.org/data-and-statistics/data-browser?country=WORLD&fuel=CO2 emissions&indicator=CO2BySector (accessed Jun. 15, 2021).

[3] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, “The role of renewable energy in the global energy transformation,” *Energy Strateg. Rev.*, vol. 24, 2019, doi: 10.1016/j.esr.2019.01.006.

[4] M. Höök and X. Tang, “Depletion of fossil fuels and anthropogenic climate change-A review,” *Energy Policy*, vol. 52, 2013, doi: 10.1016/j.enpol.2012.10.046.

[5] R. Hölsgens, “Resource dependence and energy risks in the Netherlands since the mid-nineteenth century,” *Energy Policy*, vol. 125, 2019, doi: 10.1016/j.enpol.2018.10.020.

[6] International Renewable Energy Agency (IRENA), “Renewables Readiness Assessment: The Republic of Tunisia,” 2021.

[7] International Renewable Energy Agency (IRENA), “Renewables Readiness Assessment: Republic of Moldova,” 2019.

[8] International Renewable Energy Agency (IRENA), “Renewables Readiness Assessment: Albania,” 2021.

[9] S. Howell, Y. Rezgui, J. L. Hippolyte, B. Jayan, and H. Li, “Towards the next generation of smart grids: Semantic and holonic multi-agent management of distributed energy resources,” *Renewable and Sustainable Energy Reviews*, vol. 77. 2017, doi: 10.1016/j.rser.2017.03.107.

[10] S. A. Aleem, S. M. Suhail Hussain, and T. S. Ustun, “A review of strategies to increase PV penetration level in smart grids,” *Energies*, vol. 13, no. 3. 2020, doi: 10.3390/en13030636.

[11] P. Mazidi, G. N. Baltas, M. Eliassi, and P. Rodriguez, “A Model for Flexibility Analysis of RESS with Electric Energy Storage and Reserve,” 2018, doi: 10.1109/ICRERA.2018.8566992.

[12] P. Mazidi *et al.*, “Zero Renewable Incentive Analysis for Flexibility Study of a Grid,” in *Lecture Notes in Electrical Engineering*, 2020, vol. 610 LNEE, doi: 10.1007/978-3-030-37818-9_5.

[13] M. H. Shoreh, P. Siano, M. Shafie-khah, V. Loia, and J. P. S. Catalão, “A survey of industrial applications of Demand Response,” *Electric Power Systems Research*, vol. 141. 2016, doi: 10.1016/j.epsr.2016.07.008.

[14] International Energy Agency (IEA), “Digitilization and Energy,” 2017.

[15] M. L. Di Silvestre, S. Favuzza, E. Riva Sanseverino, and G. Zizzo, “How Decarbonization, Digitalization and Decentralization are changing key power infrastructures,” *Renewable and Sustainable Energy Reviews*, vol. 93. 2018, doi: 10.1016/j.rser.2018.05.068.

[16] M. E. El-Hawary, “The smart grid - State-of-the-art and future trends,” *Electr. Power Components Syst.*, vol. 42, no. 3–4, 2014, doi: 10.1080/15325008.2013.868558.

[17] M. L. Tuballa and M. L. Abundo, “A review of the development of Smart Grid technologies,” *Renewable and Sustainable Energy Reviews*, vol. 59. 2016, doi: 10.1016/j.rser.2016.01.011.

[18] Y. Zhang, T. Huang, and E. F. Bompard, “Big data analytics in smart grids: a review,” *Energy Informatics*, vol. 1, no. 1, 2018, doi: 10.1186/s42162-018-0007-5.

[19] R. Elberg and M. Lawrence, “From Smart Grid to Neural Grid,” *Navig. Res.*, 2018.

[20] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, “Synchronized phasor measurement applications in power systems,” *IEEE Trans. Smart Grid*, vol. 1, no. 1, 2010, doi: 10.1109/TSG.2010.2044815.

[21] A. G. Phadke and J. S. Thorp, “Synchronized Phasor Measurements and their Applications,” *Springer*, 2008.

[22] V. Terzija *et al.*, “Wide-area monitoring, protection, and control of future electric power networks,” *Proc. IEEE*, vol. 99, no. 1, 2011, doi: 10.1109/JPROC.2010.2060450.

[23] Y. Wu, L. Badesa, M. T. Musavi, and P. Lerley, “Monitoring power system transient stability using synchrophasor data,” in *IEEE Power and Energy Society General Meeting*, 2015, vol. 2015-Septe, doi: 10.1109/PESGM.2015.7286142.

[24] N. Dahal, O. Abuomar, R. King, and V. Madani, “Event stream processing for improved situational awareness in the smart grid,” *Expert Syst. Appl.*, vol. 42, no. 20, 2015, doi: 10.1016/j.eswa.2015.05.003.

[25] D. Song, B. Wen, X. Yang, Y. Gu, S. Wei, and S. Ma, “A New Method for Processing and Application of Wide Area Measurement Big Data in a Power System,” 2015, doi: 10.1109/CyberC.2015.69.

[26] C. Tu, X. He, Z. Shuai, and F. Jiang, “Big data issues in smart grid – A review,” *Renewable and Sustainable Energy Reviews*, vol. 79. 2017, doi: 10.1016/j.rser.2017.05.134.

[27] P. Kundur, *Power System Stability And Control.pdf*. 1994.

[28] P. Kundur, J. Paserba, and S. Vitet, “Overview on definition and classification of power system stability,” 2003, doi: 10.1109/QSEPDS.2003.159786.

[29] R. Bessa, C. Moreira, B. Silva, and M. Matos, “Handling renewable energy variability and uncertainty in power systems operation,” *Wiley Interdiscip. Rev. Energy Environ.*, vol. 3, no. 2, 2014, doi: 10.1002/wene.76.

[30] L. Meegahapola and D. Flynn, “Impact on transient and frequency stability for a power system at very high wind penetration,” 2010, doi: 10.1109/PES.2010.5589908.

[31] M. Beaudin, H. Zareipour, A. Schellenberglabe, and W. Rosehart, “Energy storage for mitigating the variability of renewable electricity sources: An updated review,” *Energy for Sustainable Development*, vol. 14, no. 4. 2010, doi: 10.1016/j.esd.2010.09.007.

[32] H. Saad, S. Dennetière, and B. Clerc, “Interactions investigations between power electronics devices embedded in HVAC network,” in *IET Conference Publications*, 2017, vol. 2017, no. CP709, doi: 10.1049/cp.2017.0055.

[33] J. Fang, H. Li, Y. Tang, and F. Blaabjerg, “On the Inertia of Future More-Electronics Power Systems,” *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 7, no. 4, 2019, doi: 10.1109/JESTPE.2018.2877766.

[34] N. Hatziargyriou *et al.*, “Stability Definitions and Characterization of Dynamic Behavior in Systems with High Penetration of Power Electronic Interfaced Technologies.” 2020.

[35] ENTSO-E, “High Penetration of Power Electronic Interfaced Power Sources (HPoPEIPS)

ENTSO-E Guidance document for national implementation for network codes on grid connection. ENTSO-E Guidance document for national implementation for network codes on grid connection," 2017.

[36] "Workshop on High Penetration of Power Electronic Interfaced Power Sources and the Potential Contribution of Grid Forming Converters."

[37] P. Bhui and N. Senroy, "Real-Time Prediction and Control of Transient Stability Using Transient Energy Function," *IEEE Trans. Power Syst.*, vol. 32, no. 2, 2017, doi: 10.1109/TPWRS.2016.2564444.

[38] J. Yusuf, F. Sakib, and Q. Ahsan, "Study of blackout prevention of a power system," 2017, doi: 10.1109/ICECE.2016.7853971.

[39] NERC Steering Group, "Technical Analysis of the August 14, 2003, Blackout: What Happened, Why, and What Did We Learn?," *NERC Rep.*, no. July, 2004.

[40] A. Berizzi, "The Italian 2003 blackout," in *2004 IEEE Power Engineering Society General Meeting*, 2004, vol. 2, doi: 10.1109/pes.2004.1373159.

[41] ENTSO-E, "Oscillation Event 03.12.2017 – System Protection and Dynamics WG." Brussels, Belgium, 2018.

[42] K. Prasertwong, N. Mithulananthan, and D. Thakur, "Understanding low-frequency oscillation in power systems," *Int. J. Electr. Eng. Educ.*, vol. 47, no. 3, 2010, doi: 10.7227/IJEEE.47.3.2.

[43] L. L. Grigsby, *Power system stability and control*. CRC Press, 2017.

[44] H. D. Chiang, C. C. Chu, and G. Cauley, "Direct Stability Analysis of Electric Power Systems Using Energy Functions: Theory, Applications, and Perspective," *Proc. IEEE*, vol. 83, no. 11, 1995, doi: 10.1109/5.481632.

[45] B. Bonvini, S. Massucco, A. Morini, and T. Siewierski, "Comparative analysis of power system transient stability assessment by direct and hybrid methods," in *Proceedings of the Mediterranean Electrotechnical Conference - MELECON*, 1996, vol. 3, doi: 10.1109/melcon.1996.551253.

[46] A. Llamas, J. De La Ree Lopez, L. Mili, A. G. Phadke, and J. S. Thorp, "Clarifications of the BCU method for transient stability analysis," *IEEE Trans. Power Syst.*, vol. 10, no. 1,

1995, doi: 10.1109/59.373944.

[47] J. J. . Sanchez-Gasca and et al., “Identification of Electromechanical Modes in Power Systems,” 2012.

[48] A. Almunif, L. Fan, and Z. Miao, “A tutorial on data-driven eigenvalue identification: Prony analysis, matrix pencil, and eigensystem realization algorithm,” *International Transactions on Electrical Energy Systems*, vol. 30, no. 4. 2020, doi: 10.1002/2050-7038.12283.

[49] I. Nagel, L. Fabre, M. Pastre, F. Krummenacher, R. Cherkaoui, and M. Kayal, “High-speed power system transient stability simulation using highly dedicated hardware,” *IEEE Trans. Power Syst.*, vol. 28, no. 4, 2013, doi: 10.1109/TPWRS.2013.2259185.

[50] RG-CE System Protection & Dynamics Sub Group, “SPD DSA Task Force - Dynamic Security Assessment (DSA),” 2017.

[51] H. Bosetti and S. Khan, “Transient Stability in Oscillating Multi-Machine Systems Using Lyapunov Vectors,” *IEEE Trans. Power Syst.*, vol. 33, no. 2, 2018, doi: 10.1109/TPWRS.2017.2720724.

[52] W. Du *et al.*, “A Comparative Study of Two Widely Used Grid-Forming Droop Controls on Microgrid Small-Signal Stability,” *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 2, 2020, doi: 10.1109/JESTPE.2019.2942491.

[53] L. Huang, H. Xin, Z. Wang, L. Zhang, K. Wu, and J. Hu, “Transient Stability Analysis and Control Design of Droop-Controlled Voltage Source Converters Considering Current Limitation,” *IEEE Trans. Smart Grid*, vol. 10, no. 1, 2019, doi: 10.1109/TSG.2017.2749259.

[54] H. Wu and X. Wang, “Design-oriented transient stability analysis of grid-connected converters with power synchronization control,” *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, 2019, doi: 10.1109/TIE.2018.2875669.

[55] H. Wu and X. Wang, “Transient angle stability analysis of grid-connected converters with the first-order active power loop,” in *Conference Proceedings - IEEE Applied Power Electronics Conference and Exposition - APEC*, 2018, vol. 2018-March, doi: 10.1109/APEC.2018.8341529.

[56] L. Cheng and T. Yu, “A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems,” *Int. J. Energy Res.*, vol. 43, no. 6, 2019, doi: 10.1002/er.4333.

[57] L. Duchesne, E. Karangelos, and L. Wehenkel, “Recent Developments in Machine Learning for Energy Systems Reliability Management,” *Proc. IEEE*, vol. 108, no. 9, 2020, doi: 10.1109/JPROC.2020.2988715.

[58] D. W. Gao *et al.*, “Application of AI techniques in monitoring and operation of power systems,” *Front. Energy*, vol. 13, no. 1, 2019, doi: 10.1007/s11708-018-0589-4.

[59] S. Zhao, F. Blaabjerg, and H. Wang, “An overview of artificial intelligence applications for power electronics,” *IEEE Trans. Power Electron.*, vol. 36, no. 4, 2021, doi: 10.1109/TPEL.2020.3024914.

[60] L. Wehcnkel, T. Van Cutsem, and M. Ribbens-Pavella, “An artificial intelligence framework for on-line transient stability assessment of power systems,” *IEEE Trans. Power Syst.*, vol. 4, no. 2, 1989, doi: 10.1109/59.193853.

[61] B. Jayasekara and U. D. Annakkage, “Derivation of an accurate polynomial representation of the transient stability boundary,” *IEEE Trans. Power Syst.*, vol. 21, no. 4, 2006, doi: 10.1109/TPWRS.2006.881111.

[62] J. Lv, M. Pawlak, and U. D. Annakkage, “Prediction of the transient stability boundary using the lasso,” *IEEE Trans. Power Syst.*, vol. 28, no. 1, 2013, doi: 10.1109/TPWRS.2012.2197763.

[63] J. Lv, M. Pawlak, and U. D. Annakkage, “Prediction of the Transient Stability Boundary Based on Nonparametric Additive Modeling,” *IEEE Trans. Power Syst.*, vol. 32, no. 6, 2017, doi: 10.1109/TPWRS.2017.2669839.

[64] A. Sharifian and S. Sharifian, “A new power system transient stability assessment method based on Type-2 fuzzy neural network estimation,” *Int. J. Electr. Power Energy Syst.*, vol. 64, 2015, doi: 10.1016/j.ijepes.2014.07.007.

[65] M. Arefi and B. Chowdhury, “Ensemble adaptive neuro fuzzy support vector machine for prediction of transient stability,” 2017, doi: 10.1109/NAPS.2017.8107168.

[66] C. Zheng, V. Malbasa, and M. Kezunovic, “Regression tree for stability margin prediction

using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, 2013, doi: 10.1109/TPWRS.2012.2220988.

[67] A. Karami, "Power system transient stability margin estimation using neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 4, 2011, doi: 10.1016/j.ijepes.2011.01.012.

[68] A. Karami and S. Z. Esmaili, "Transient stability assessment of power systems described with detailed models using neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 45, no. 1, 2013, doi: 10.1016/j.ijepes.2012.08.071.

[69] R. Diao, V. Vittal, and N. Logic, "Design of a real-time security assessment tool for situational awareness enhancement in modern power systems," *IEEE Trans. Power Syst.*, vol. 25, no. 2, 2010, doi: 10.1109/TPWRS.2009.2035507.

[70] T. Liu, Y. Liu, J. Liu, Y. Yang, G. A. Taylor, and Z. Huang, "Multi-indicator inference scheme for fuzzy assessment of power system transient stability," *CSEE J. Power Energy Syst.*, vol. 2, no. 3, pp. 1--9, 2016, doi: 10.17775/CSEEJPES.2016.00029.

[71] Y. Xu, Z. Y. Dong, J. Zhao, K. Meng, and K. P. Wong, "Transient stability assessment based on data-structure analysis of operating point space," 2010, doi: 10.1109/PES.2010.5590223.

[72] B. A. Archer, U. D. Annakkage, B. Jayasekara, and P. Wijetunge, "Accurate prediction of damping in large interconnected power systems with the aid of regression analysis," *IEEE Trans. Power Syst.*, vol. 23, no. 3, 2008, doi: 10.1109/TPWRS.2008.926708.

[73] F. Sulla, E. Masback, and O. Samuelsson, "Linking damping of electromechanical oscillations to system operating conditions using neural networks," in *IEEE PES Innovative Smart Grid Technologies Conference Europe*, 2015, vol. 2015-Janua, no. January, doi: 10.1109/ISGTEurope.2014.7028818.

[74] L. S. Moulin, A. P. Alves Da Silva, M. A. El-Sharkawi, and R. J. Marks, "Support vector machines for transient stability analysis of large-scale power systems," *IEEE Trans. Power Syst.*, vol. 19, no. 2, 2004, doi: 10.1109/TPWRS.2004.826018.

[75] D. You, K. Wang, L. Ye, J. Wu, and R. Huang, "Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs," *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, 2013, doi: 10.1016/j.ijepes.2012.07.057.

[76] W. Zhang, W. Hu, Y. Min, L. Chen, L. Zheng, and X. Liu, “A novel stability classifier based on reformed support vector machines for online stability assessment,” in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2016, vol. 2016-Janua, doi: 10.1109/APPEEC.2015.7380884.

[77] D. Yuanhang, C. Lei, Z. Weiling, and M. Yong, “Multi-support vector machine power system transient stability assessment based on relief algorithm,” in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2016, vol. 2016-Janua, doi: 10.1109/APPEEC.2015.7381006.

[78] N. G. Baltas, P. Mazidi, J. Ma, F. De Asis Fernandez, and P. Rodriguez, “A comparative analysis of decision trees, support vector machines and artificial neural networks for on-line transient stability assessment,” 2018, doi: 10.1109/SEST.2018.8495872.

[79] I. Kamwa, S. R. Samantaray, and G. Joos, “Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices,” *IEEE Trans. Smart Grid*, vol. 1, no. 2, 2010, doi: 10.1109/TSG.2010.2052935.

[80] Y. Chen, S. M. Mazhari, C. Y. Chung, S. O. Faried, B. Wang, and B. Hu, “Power system on-line transient stability prediction by margin indices and random forests,” 2019, doi: 10.1109/EPEC47565.2019.9074804.

[81] C. Zhang, Y. Li, Z. Yu, and F. Tian, “Feature selection of power system transient stability assessment based on random forest and recursive feature elimination,” in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2016, vol. 2016-Decem, doi: 10.1109/APPEEC.2016.7779696.

[82] V. Vittal, “Application of phasor measurements for dynamic security assessment using decision trees,” 2012, doi: 10.1109/PESGM.2012.6344580.

[83] W. D. Oliveira, J. P. A. Vieira, U. H. Bezerra, D. A. Martins, and B. das G. Rodrigues, “Power system security assessment for multiple contingencies using multiway decision tree,” *Electr. Power Syst. Res.*, vol. 148, 2017, doi: 10.1016/j.epsr.2017.03.029.

[84] J. An, J. Yu, Z. Li, Y. Zhou, and G. Mu, “A Data-driven Method for Transient Stability Margin Prediction Based on Security Region,” *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 6, 2020, doi: 10.35833/MPCE.2020.000457.

[85] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, “Power System Transient Stability

Assessment Based on Big Data and the Core Vector Machine," *IEEE Trans. Smart Grid*, vol. 7, no. 5, 2016, doi: 10.1109/TSG.2016.2549063.

[86] A. B. Mosavi, A. Amiri, and H. Hosseini, "A Learning Framework for Size and Type Independent Transient Stability Prediction of Power System Using Twin Convolutional Support Vector Machine," *IEEE Access*, vol. 6, 2018, doi: 10.1109/ACCESS.2018.2880273.

[87] B. P. Soni, A. Saxena, V. Gupta, and S. L. Surana, "Assessment of Transient Stability through Coherent Machine Identification by Using Least-Square Support Vector Machine," *Model. Simul. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/5608591.

[88] G. N. Baltas, H. R. Chamorro, F. Gonzalez-Longatt, and P. Rodriguez, "Coherency Groups Analysis based on Self Organizing Maps," in *IEEE Power and Energy Society General Meeting*, 2019, vol. 2019-Augus, doi: 10.1109/PESGM40551.2019.8973411.

[89] G. N. Baltas, N. Bao Lai, L. Marin, A. Tarraso, and P. Rodriguez, "A Growing Self-Organising Maps Implementation for Coherency Identification in a Power Electronics Dominated Power System," 2020, doi: 10.1109/ECCE44975.2020.9235611.

[90] N. W. A. Lidula and A. D. Rajapakse, "A pattern recognition approach for detecting power islands using transient signals - Part I: Design and implementation," *IEEE Trans. Power Deliv.*, vol. 25, no. 4, 2010, doi: 10.1109/TPWRD.2010.2053724.

[91] N. W. A. Lidula and A. D. Rajapakse, "A pattern-recognition approach for detecting power islands using transient signals-part II: Performance evaluation," *IEEE Trans. Power Deliv.*, vol. 27, no. 3, 2012, doi: 10.1109/TPWRD.2012.2187344.

[92] Q. Cui, K. El-Arroudi, and G. Joos, "Islanding detection of hybrid distributed generation under reduced non-detection zone," *IEEE Trans. Smart Grid*, vol. 9, no. 5, 2018, doi: 10.1109/TSG.2017.2679101.

[93] M. R. Alam, K. M. Muttaqi, and A. Bouzerdoum, "Evaluating the effectiveness of a machine learning approach based on response time and reliability for islanding detection of distributed generation," *IET Renew. Power Gener.*, vol. 11, no. 11, 2017, doi: 10.1049/iet-rpg.2016.0987.

[94] N. Amjady and S. F. Majedi, "Transient stability prediction by a hybrid intelligent system," *IEEE Trans. Power Syst.*, vol. 22, no. 3, 2007, doi:

10.1109/TPWRS.2007.901667.

[95] N. Amjady and S. A. Banihashemi, “Transient stability prediction of power systems by a new synchronism status index and hybrid classifier,” *IET Gener. Transm. Distrib.*, vol. 4, no. 4, 2010, doi: 10.1049/iet-gtd.2009.0255.

[96] T. Guo, J. He, Z. Li, and J. V. Milanović, “Evaluation of classification methods for on-line identification of power system dynamic signature,” 2014, doi: 10.1109/PSCC.2014.7038430.

[97] S. A. Siddiqui, K. Verma, K. R. Niazi, and M. Fozdar, “Artificial neural network based early detection of real-time transient instability for initiation of emergency control through wide-area synchrophasor measurements,” 2016, doi: 10.1109/IC4.2015.7375733.

[98] M. Mahdi and V. M. I. Genc, “Artificial neural network based algorithm for early prediction of transient stability using wide area measurements,” 2017, doi: 10.1109/SGCF.2017.7947611.

[99] Y. Chen, S. M. Mazhari, C. Y. Chung, S. O. Faried, and B. C. Pal, “Rotor Angle Stability Prediction of Power Systems with High Wind Power Penetration Using a Stability Index Vector,” *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4632–4643, 2020, doi: 10.1109/TPWRS.2020.2989725.

[100] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, “A topology adaptive high-speed transient stability assessment scheme based on multi-graph attention network with residual structure,” *Int. J. Electr. Power Energy Syst.*, vol. 130, p. 106948, 2021, doi: https://doi.org/10.1016/j.ijepes.2021.106948.

[101] Y. Zhou, J. Wu, L. Hao, L. Ji, and Z. Yu, “Transient Stability Prediction of Power Systems Using Post-disturbance Rotor Angle Trajectory Cluster Features,” *Electr. Power Components Syst.*, vol. 44, no. 17, 2016, doi: 10.1080/15325008.2016.1204373.

[102] M. R. Aghamohammadi and M. Abedi, “DT based intelligent predictor for out of step condition of generator by using PMU data,” *Int. J. Electr. Power Energy Syst.*, vol. 99, 2018, doi: 10.1016/j.ijepes.2018.01.001.

[103] F. Darbandi, A. Jafari, H. Karimipour, A. Dehghantanha, F. Derakhshan, and K. K. R. Choo, “Real-time stability assessment in smart cyberphysical grids: A deep learning approach,” *IET Smart Grid*, vol. 3, no. 4. 2020, doi: 10.1049/iet-stg.2019.0191.

[104] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. Gu, and V. O. K. Li, “Intelligent time-adaptive transient stability assessment system,” *IEEE Trans. Power Syst.*, vol. 33, no. 1, 2018, doi: 10.1109/TPWRS.2017.2707501.

[105] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv*, 2014.

[106] F. Pan *et al.*, “Stacked-GRU based power system transient stability assessment method,” *Algorithms*, vol. 11, no. 8, 2018, doi: 10.3390/a11080121.

[107] Y. Xu, Z. Y. Dong, Z. Xu, K. Meng, and K. P. Wong, “An intelligent dynamic security assessment framework for power systems with wind power,” *IEEE Trans. Ind. Informatics*, vol. 8, no. 4, 2012, doi: 10.1109/TII.2012.2206396.

[108] P. N. Papadopoulos, T. Guo, and J. V. Milanović, “Probabilistic framework for online identification of dynamic behavior of power systems with renewable generation,” *IEEE Trans. Power Syst.*, vol. 33, no. 1, 2018, doi: 10.1109/TPWRS.2017.2688446.

[109] T. Guo and J. V. Milanovic, “On-line prediction of transient stability using decision tree method - Sensitivity of accuracy of prediction to different uncertainties,” 2013, doi: 10.1109/PTC.2013.6652106.

[110] I. Genc, R. Diao, and V. Vittal, “Computation of transient stability related security regions and generation rescheduling based on decision trees,” 2010, doi: 10.1109/PES.2010.5589364.

[111] Y. Zhou, J. Wu, L. Ji, Z. Yu, K. Lin, and L. Hao, “Transient stability preventive control of power systems using chaotic particle swarm optimization combined with two-stage support vector machine,” *Electr. Power Syst. Res.*, vol. 155, 2018, doi: 10.1016/j.epsr.2017.10.007.

[112] V. Tjeng, K. Xiao, and R. Tedrake, “Evaluating robustness of neural networks with mixed integer programming,” 2019.

[113] A. Venzke, D. T. Viola, J. Mermet-Guyennet, G. S. Misyris, and S. Chatzivasileiadis, “Neural networks for encoding dynamic security-constrained optimal power flow to mixed-integer linear programs,” *arXiv*. 2020.

[114] Y. Zhang and K. Tomsovic, “Real-time transient instability detection based on decision

trees," 2009, doi: 10.1109/ISAP.2009.5352947.

[115] E. M. Voumvoulakis and N. D. Hatziargyriou, "Decision trees-aided self-organized maps for corrective dynamic security," *IEEE Trans. Power Syst.*, vol. 23, no. 2, 2008, doi: 10.1109/TPWRS.2008.920194.

[116] P. Gopakumar, M. J. B. Reddy, and D. K. Mohanta, "Stability control of smart power grids with artificial intelligence and wide-area synchrophasor measurements," *Electr. Power Components Syst.*, vol. 42, no. 10, 2014, doi: 10.1080/15325008.2014.913745.

[117] O. Soares, H. Gonçalves, A. Martins, and A. Carvalho, "Nonlinear control of the doubly-fed induction generator in wind power systems," *Renew. Energy*, vol. 35, no. 8, 2010, doi: 10.1016/j.renene.2009.12.008.

[118] K. H. Lu, C. M. Hong, and Q. Xu, "Recurrent wavelet-based Elman neural network with modified gravitational search algorithm control for integrated offshore wind and wave power generation systems," *Energy*, vol. 170, 2019, doi: 10.1016/j.energy.2018.12.084.

[119] M. Ren, J. Zhang, Y. Tian, and G. Hou, "A neural network controller for variable-speed variable-pitch wind energy conversion systems using generalized minimum entropy criterion," *Math. Probl. Eng.*, vol. 2014, 2014, doi: 10.1155/2014/412027.

[120] W. Qiao, G. K. Venayagamoorthy, and R. G. Harley, "DHP-based wide-area coordinating control of a power system with a large wind farm and multiple FACTS devices," 2007, doi: 10.1109/IJCNN.2007.4371281.

[121] J. Duan, Z. Yi, D. Shi, H. Xu, and Z. Wang, "A Neuron-Network-Based Optimal Control of Ultra-Capacitors with System Uncertainties," 2019, doi: 10.1109/ISGT.2019.8791559.

[122] X. Fu and S. Li, "Control of Single-Phase Grid-Connected Converters with LCL Filters Using Recurrent Neural Network and Conventional Control Methods," *IEEE Trans. Power Electron.*, vol. 31, no. 7, 2016, doi: 10.1109/TPEL.2015.2490200.

[123] X. Yu, F. Gao, and G. Ding, "Deep Learning Based Transient Stability Assessment for Grid-Connected Inverter," 2018, doi: 10.1109/PEAC.2018.8590332.

[124] Z. Dai, Z. Zhang, Y. Yang, F. Blaabjerg, Y. Huangfu, and J. Zhang, "A Fixed-Length Transfer Delay Based Adaptive Frequency-Locked Loop for Single-Phase Systems," *IEEE Trans. Power Electron.*, vol. 34, no. 5, 2019, doi: 10.1109/TPEL.2018.2871032.

[125] A. R. Tavakoli, A. R. Seifi, and M. M. Arefi, “Designing a self-constructing fuzzy neural network controller for damping power system oscillations,” *Fuzzy Sets Syst.*, vol. 356, 2019, doi: 10.1016/j.fss.2018.01.006.

[126] L. Ji, J. Wu, Y. Zhou, and L. Hao, “Using trajectory clusters to define the most relevant features for transient stability prediction based on machine learning method,” *Energies*, vol. 9, no. 11, 2016, doi: 10.3390/en9110898.

[127] M. He, J. Zhang, and V. Vittal, “Robust online dynamic security assessment using adaptive ensemble decision-tree learning,” *IEEE Trans. Power Syst.*, vol. 28, no. 4, 2013, doi: 10.1109/TPWRS.2013.2266617.

[128] B. Tan, J. Yang, X. Pan, J. Li, P. Xie, and C. Zeng, “Representational learning approach for power system transient stability assessment based on convolutional neural network,” *J. Eng.*, vol. 2017, no. 13, 2017, doi: 10.1049/joe.2017.0651.

[129] S. M. Mazhari, B. Khorramdel, C. Y. Chung, I. Kamwa, and D. Novosel, “A Simulation-Based Classification Approach for Online Prediction of Generator Dynamic Behavior under Multiple Large Disturbances,” *IEEE Trans. Power Syst.*, vol. 36, no. 2, 2021, doi: 10.1109/TPWRS.2020.3021137.

[130] 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, 2011, doi: 10.1109/TPWRS.2010.2082575.

[131] J. Geeganage, U. D. Annakkage, T. Weekes, and B. A. Archer, “Application of Energy-Based Power System Features for Dynamic Security Assessment,” *IEEE Trans. Power Syst.*, vol. 30, no. 4, 2015, doi: 10.1109/TPWRS.2014.2353048.

[132] E. A. Frimpong, P. Y. Okyere, and J. Asumadu, “On-line determination of transient stability status using MLPNN,” 2017, doi: 10.1109/PowerAfrica.2017.7991194.

[133] E. A. Frimpong, P. Y. Okyere, and J. Asumadu, “Wavelet Analysis and Neural Network Technique for Predicting Transient Stability Status,” *Carpathian J. Electr. Eng.*, vol. 14, no. 1, pp. 42–56, 2020.

[134] N. Shenoy and R. Ramakumar, “Power systems Dynamic Security Assessment using fisher information metric,” in *IEEE Power and Energy Society General Meeting*, 2016,

vol. 2016-Novem, doi: 10.1109/PESGM.2016.7741551.

[135] A. Venzke and S. Chatzivasileiadis, “Verification of Neural Network Behaviour: Formal Guarantees for Power System Applications,” *IEEE Trans. Smart Grid*, vol. 12, no. 1, 2021, doi: 10.1109/TSG.2020.3009401.

[136] D. Tsipras, S. Santurkar, L. Engstrom, A. Turner, and A. Mądry, “Robustness May Be at Odds with Accuracy,” *arXiv*. 2018.

[137] F. Thams, A. Venzke, R. Eriksson, and S. Chatzivasileiadis, “Efficient Database Generation for Data-Driven Security Assessment of Power Systems,” *IEEE Trans. Power Syst.*, vol. 35, no. 1, 2020, doi: 10.1109/TPWRS.2018.2890769.

[138] S. You *et al.*, “Build Smart Grids on Artificial Intelligence -- A Real-world Example,” 2020.

[139] “Demonstration 8: Robust conventional generation plant through Power System Stabilizer | FLEXITRANSTORE.” http://www.flexitranstore.eu/Demo-8 (accessed Apr. 12, 2021).

[140] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *J. Comput. Phys.*, vol. 378, 2019, doi: 10.1016/j.jcp.2018.10.045.

[141] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis, “Physics-informed neural networks for power systems,” in *IEEE Power and Energy Society General Meeting*, 2020, vol. 2020-Augus, doi: 10.1109/PESGM41954.2020.9282004.

[142] A. L. Samuel, “Some studies in machine learning using the game of checkers,” *IBM J. Res. Dev.*, vol. 44, no. 1–2, 2000, doi: 10.1147/rd.441.0206.

[143] A. Mellit and S. A. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review,” *Progress in Energy and Combustion Science*, vol. 34, no. 5. 2008, doi: 10.1016/j.pecs.2008.01.001.

[144] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, vol. 10. 2001.

[145] L. Rokach and O. Maimon, *Data Mining With Decision Trees: Theory and Applications*. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 2014.

[146] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” *Data*

Min. Knowl. Discov., vol. 2, no. 2, 1998, doi: 10.1023/A:1009715923555.

[147] M. A. Hearst, B. Scholkopf, S. Dumais, E. Osuna, and J. Platt, "Supprot vector machines," *IEEE Intell. Syst. their Appl.*, vol. 13, no. 4, 1998.

[148] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. 2000.

[149] "A Complete List of Kernels Used in Support Vector Machines," *Biochem. Pharmacol. Open Access*, vol. 04, no. 05, 2015, doi: 10.4172/2167-0501.1000195.

[150] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Fourth. Academic Press, 2008.

[151] T. Schaul, D. Horgan, K. Gregor, and D. Silver, "Universal value function approximators," in *32nd International Conference on Machine Learning, ICML 2015*, 2015, vol. 2.

[152] D. . Rumelhart, G. . Hinton, and R. . Williams, "Learning Internal Representations By Error Propagation (original)," *Explorations in the Micro-Structure of Cognition Vol. 1 : Foundations*, no. V. 1986.

[153] Aurélien Géron, "Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems," *O'Reilly Media*. p. 851, 2019.

[154] M. T. Haque and A. M. Kashtiban, "Application of neural networks in power systems; A review," in *Proceedings - Wec 05: Fourth World Enformatika Conference*, 2005, vol. 6.

[155] B. Wang and J. Pineau, "Online Bagging and Boosting for Imbalanced Data Streams," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 12, 2016, doi: 10.1109/TKDE.2016.2609424.

[156] T. Guo, P. Papadopoulos, P. Mohammed, and J. V. Milanovic, "Comparison of ensemble decision tree methods for on-line identification of power system dynamic signature considering availability of PMU measurements," 2015, doi: 10.1109/PTC.2015.7232364.

[157] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[158] G. Rogers, *Power System Oscillations*. Springer US, 2000.

[159] J. J. Sanchez-Gasca, "Identification of Electromechanical Modes in Power Systems,"

IEEE PES Resour. Center, IEEE PES Tech. Rep. TR-15, no. June 2012, pp. 1–282, 2012.

[160] N. C. Chakraborty, A. Banerji, and S. K. Biswas, "Survey on major blackouts analysis and prevention methodologies," in *IET Conference Publications*, 2015, vol. 2015, no. CP683, pp. 297–302, doi: 10.1049/cp.2015.1647.

[161] C. Canizares *et al.*, "Benchmark Models for the Analysis and Control of Small-Signal Oscillatory Dynamics in Power Systems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, 2017, doi: 10.1109/TPWRS.2016.2561263.

[162] G. E. P. Box and M. E. Muller, "A Note on the Generation of Random Normal Deviates," *Ann. Math. Stat.*, vol. 29, no. 2, 1958, doi: 10.1214/aoms/1177706645.

[163] F. Marcham, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems (Preliminary White Paper, November 9, 2015)," *Library (Lond).*, vol. s4-X, no. 3, 1929.

[164] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, 2011.

[165] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*, vol. 9781107057135. 2013.

[166] J. Menčík and J. Menčík, "Latin Hypercube Sampling," in *Concise Reliability for Engineers*, InTech, 2016.

[167] E. P. Zafiropoulos and E. N. Dialynas, "Reliability and cost optimization of electronic devices considering the component failure rate uncertainty," *Reliab. Eng. Syst. Saf.*, vol. 84, no. 3, 2004, doi: 10.1016/j.ress.2003.11.012.

[168] W. Zhang, A. Tarraso, J. Rocabert, A. Luna, J. Ignacio Candela, and P. Rodriguez, "Frequency Support Properties of the Synchronous Power Control for Grid-Connected Converters," in *IEEE Transactions on Industry Applications*, 2019, vol. 55, no. 5, doi: 10.1109/TIA.2019.2928517.

[169] J. Rocabert, A. Luna, F. Blaabjerg, and P. Rodríguez, "Control of power converters in AC microgrids," *IEEE Trans. Power Electron.*, vol. 27, no. 11, 2012, doi: 10.1109/TPEL.2012.2199334.

[170] S. Yazdani, M. Ferdowsi, M. Davari, and P. Shamsi, "Advanced Current-Limiting and

Power-Sharing Control in a PV-Based Grid-Forming Inverter under Unbalanced Grid Conditions," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 2, 2020, doi: 10.1109/JESTPE.2019.2959006.

[171] A. Tayyebi, D. Grob, A. Anta, F. Kupzog, and F. Dorfler, "Frequency Stability of Synchronous Machines and Grid-Forming Power Converters," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 2, 2020, doi: 10.1109/JESTPE.2020.2966524.

[172] M. G. Taul, X. Wang, P. Davari, and F. Blaabjerg, "Current Limiting Control with Enhanced Dynamics of Grid-Forming Converters during Fault Conditions," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 2, 2020, doi: 10.1109/JESTPE.2019.2931477.

[173] P. Rodriguez, C. Citro, J. I. Candela, J. Rocabert, and A. Luna, "Flexible Grid Connection and Islanding of SPC-Based PV Power Converters," in *IEEE Transactions on Industry Applications*, 2018, vol. 54, no. 3, doi: 10.1109/TIA.2018.2800683.

[174] A. Tarraso, N. B. Lai, G. N. Baltas, and P. Rodriguez, "Power quality services provided by virtually synchronous facts," *Energies*, vol. 12, no. 17, 2019, doi: 10.3390/en12173292.

[175] P. Rodriguez, I. Candela, J. Rocabert, and R. Teodorescu, "Virtual Controller of Electromechanical Characteristics for Static Power Converters," 2012.

[176] G. N. Baltas, N. B. Lai, L. Marin, A. Tarrasó, and P. Rodriguez, "Grid-Forming Power Converters Tuned through Artificial Intelligence to Damp Subsynchronous Interactions in Electrical Grids," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2995298.

[177] K. Diakos, Q. Wu, and A. H. Nielsen, "Phasor measurement unit and phasor data concentrator test with real time digital simulator," in *Asia-Pacific Power and Energy Engineering Conference, APPEEC*, 2014, vol. 2015-March, no. March, doi: 10.1109/APPEEC.2014.7066096.

[178] T. S. Ustun, S. M. Farooq, and S. M. Suhail Hussain, "Implementing Secure Routable GOOSE and SV Messages Based on IEC 61850-90-5," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2971011.

[179] T. S. Ustun, S. M. Farooq, and S. M. S. Hussain, "A novel approach for mitigation of replay and masquerade attacks in smartgrids using IEC 61850 Standard," *IEEE Access*,

vol. 7, 2019, doi: 10.1109/ACCESS.2019.2948117.

[180] M. Abdollahi, J. I. Candela, J. Rocabert, M. A. Elsaharty, and P. Rodriguez, “Novel Analytical Method for Dynamic Design of Renewable SSG SPC Unit to Mitigate Low-Frequency Electromechanical Oscillations,” *IEEE Trans. Power Electron.*, vol. 35, no. 7, 2020, doi: 10.1109/TPEL.2019.2956397.

[181] A. Stativă and F. Gonzalez-Longatt, “Peer-to-Peer (P2p) MATLAB–powerfactory communication: Optimal placement and setting of power system stabilizer,” in *Green Energy and Technology*, vol. 0, no. 9783319505312, 2018.

[182] G. N. Baltas, N. B. Lai, A. Tarraso, L. Marin, F. Blaabjerg, and P. Rodriguez, “AI-Based Damping of Electromechanical Oscillations by Using Grid-Connected Converter,” *Front. Energy Res.*, vol. 9, 2021, doi: 10.3389/fenrg.2021.598436.

[183] M. Weixelbraun, H. Renner, O. Kirkeluten, and S. Lovlund, “Damping low frequency oscillations with hydro governors,” 2013, doi: 10.1109/PTC.2013.6652341.

[184] I. Kamwa, R. Grondin, and G. Trudel, “IEEE PSS2B versus PSS4B: The limits of performance of modern power system stabilizers,” *IEEE Trans. Power Syst.*, vol. 20, no. 2, 2005, doi: 10.1109/TPWRS.2005.846197.

[185] Z. Assi Obaid, L. M. Cipcigan, and M. T. Muhssin, “Power system oscillations and control: Classifications and PSSs’ design methods: A review,” *Renew. Sustain. Energy Rev.*, vol. 79, pp. 839–849, Nov. 2017, doi: 10.1016/j.rser.2017.05.103.

[186] F. Rashidi and M. Rashidi, “Robust and adaptive tuning of power system stabilizers using artificial neural networks,” in *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 2004, vol. 3029, doi: 10.1007/978-3-540-24677-0_105.

[187] F. Zhuang *et al.*, “A Comprehensive Survey on Transfer Learning,” *Proceedings of the IEEE*, vol. 109, no. 1. 2021, doi: 10.1109/JPROC.2020.3004555.

[188] 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, 2020, doi: 10.1109/TPWRS.2019.2947781.

[189] I. W. Tsang, J. T. Kwok, and P. M. Cheung, “Core vector machines: Fast SVM training on very large data sets,” *J. Mach. Learn. Res.*, vol. 6, 2005.

[190] Jayadeva, R. Khemchandani, and S. Chandra, “Twin support vector machines for pattern classification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 5, 2007, doi: 10.1109/TPAMI.2007.1068.

[191] G. N. Baltas, C. Perales-Gonzalez, P. Mazidi, F. Fernandez, and P. Rodriguez, “A Novel Ensemble Approach for Solving the Transient Stability Classification Problem,” 2018, doi: 10.1109/ICRERA.2018.8566815.

Tesis depositada en Universidad Loyola
Sevilla, 2021

Universidad
LOYOLA