1-1-2019

Power System Resilience Enhancement Using Artificial Intelligence

Rozhin Eskandarpour
University of Denver

Power System Resilience Enhancement Using Artificial Intelligence

Abstract

Extreme weather events and natural disasters are the major cause of power outages in the United States. An accurate forecast of component outages and the resultant load curtailment in response to extreme events is an essential task in pre- and post-event planning, recovery and hardening of power systems. Power system resilience improvement is investigated in this work from component outage prediction to identifying the potential power outages in the system to estimating probable load curtailment due to these outages and offering methods for grid hardening. Initially, two machine learning based prediction methods are proposed to determine the potential outage of power grid components in response to an imminent hurricane, namely a second order logistic regression model and a three-dimensional Support Vector Machine (SVM). The logistic regression model defines the decision boundary, which partitions the components' states into two sets of damaged and operational. Two metrics are examined to validate the performance of the obtained decision boundary in efficiently predicting component outages. The proposed three-dimensional SVM furthermore leverages its accuracy-uncertainty tradeoff to achieve highly accurate results, which can be further used to schedule system resources in a predictive manner with the objective of maximizing its resilience. The performance of the model is tested through numerical simulations and validated based on well-defined and commonly-used performance measures.

After training the outage estimation model, the predicted component outages are plugged into a load curtailment minimization model to estimate the nodal load curtailments in the system. The standard IEEE 30-bus system with a combination of hurricane path and intensity scenarios are used to study the model where the results demonstrate that the proposed modelling framework is capable of effectively capturing the dynamics of load curtailment estimation in response to extreme events.

Furthermore, a machine learning based grid hardening model is proposed with the objective of improving power grid resilience. The predictions from previous stages are fed into the proposed grid hardening model, which determines strategic locations for placement of distributed generation (DG) units. In contrast to existing literature in hardening and resilience enhancement, this work co-optimizes grid economic and resilience objectives by considering the intricate dependencies of the two. The numerical simulations on the standard IEEE 118-bus test system illustrate the merits and applicability of the proposed model. The results further indicate that the proposed hardening model through decentralized and distributed local energy resources can produce a more robust solution that can protect the system significantly against multiple component outages.

Finally, a probabilistic load curtailment estimation model is proposed through a three-step sequential method. At first, to determine a deterministic outage state of the grid components in response to a forecasted hurricane, a machine learning model based on TWSVM is proposed. Then, to convert the deterministic results into probabilistic outage states, a posterior probability sigmoid model is trained on the obtained results from the previous step. Finally, the obtained component outages are integrated into a load curtailment estimation model to determine the potential load curtailments in the system.

The simulation results on a standard test system illustrate the high accuracy performance of the proposed method.

Document Type

Dissertation

Degree Name

Ph.D.

Department
Electrical Engineering

First Advisor
Amin Khodaei, Ph.D.

Second Advisor
Ali Besharat, Ph.D.

Third Advisor
Mohammad Matin, Ph.D.

Keywords
Support Vector Machine, Distributed generation units, Power system resilience improvement

Subject Categories
Electrical and Computer Engineering | Engineering

Power System Resilience Enhancement Using

Artificial Intelligence

A Dissertation

Presented to

the Faculty of the Daniel Felix Ritchie School Engineering and Computer Science

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Rozhin Eskandarpour

August 2019

Advisor: Dr. Amin Khodaei

Author: Rozhin Eskandarpour
Title: Power System Resilience Enhancement Using Artificial Intelligence
Advisor: Dr. Amin Khodaei
Degree Date: August 2019

ABSTRACT

Extreme weather events and natural disasters are the major cause of power outages in the United States. An accurate forecast of component outages and the resultant load curtailment in response to extreme events is an essential task in pre- and post-event planning, recovery and hardening of power systems. Power system resilience improvement is investigated in this work from component outage prediction to identifying the potential power outages in the system to estimating probable load curtailment due to these outages and offering methods for grid hardening. Initially, two machine learning based prediction methods are proposed to determine the potential outage of power grid components in response to an imminent hurricane, namely a second order logistic regression model and a three-dimensional Support Vector Machine (SVM). The logistic regression model defines the decision boundary, which partitions the components' states into two sets of damaged and operational. Two metrics are examined to validate the performance of the obtained decision boundary in efficiently predicting component outages. The proposed three-dimensional SVM furthermore leverages its accuracy-uncertainty tradeoff to achieve highly accurate results, which can be further used to schedule system resources in a predictive manner with the objective of maximizing its resilience. The performance of the model is tested through numerical simulations and validated based on well-defined and commonly-used performance measures.

After training the outage estimation model, the predicted component outages are plugged into a load curtailment minimization model to estimate the nodal load curtailments in the system. The standard IEEE 30-bus system with a combination of hurricane path and intensity scenarios are used to study the model where the results demonstrate that the proposed modelling framework is capable of effectively capturing the dynamics of load curtailment estimation in response to extreme events.

Furthermore, a machine learning based grid hardening model is proposed with the objective of improving power grid resilience. The predictions from previous stages are fed into the proposed grid hardening model, which determines strategic locations for placement of distributed generation (DG) units. In contrast to existing literature in hardening and resilience enhancement, this work co-optimizes grid economic and resilience objectives by considering the intricate dependencies of the two. The numerical simulations on the standard IEEE 118-bus test system illustrate the merits and applicability of the proposed model. The results further indicate that the proposed hardening model through decentralized and distributed local energy resources can produce a more robust solution that can protect the system significantly against multiple component outages.

Finally, a probabilistic load curtailment estimation model is proposed through a three-step sequential method. At first, to determine a deterministic outage state of the grid components in response to a forecasted hurricane, a machine learning model based on TWSVM is proposed. Then, to convert the deterministic results into probabilistic outage states, a posterior probability sigmoid model is trained on the obtained results from the previous step. Finally, the obtained component outages are integrated into a load curtailment estimation model to determine the potential load curtailments in the system.

The simulation results on a standard test system illustrate the high accuracy performance of the proposed method.

ACKNOWLEDGEMENTS

One amazing scholar, one privilege opportunity, and one endless support have changed my life.

It is my greatest pleasure to express my deepest gratitude and appreciation to my adviser, Dr. Amin Khodaei for supporting me in this incredible field of research and his endless commitment. All my achievements would have been unreachable without my adviser's tactful support and his never-ending compassion. I truly thank him for his kind everlasting help, generous advice and dedication throughout my study.

Additionally and foremost, I would like to show my sincerest appreciations to Dr. Amin Khodaei one more time, not only because he is an excellent and knowledgeable professor but also because he is a considerate and supportive chair and I have always admired his devotion to students.

I am also very grateful to my oral defense committee Dr. Ali Besharat, Dr. Mohammad Matin and Dr. Wenzhong Gao for their assistance to improve this project; and truly appreciate their time and consideration.

Last but not least, this thesis is dedicated to my parents and my valuable friends, who have been a constant source of support and encouragement during the challenges of this research. I truly thank them for their warm everlasting help, compassionate advice and support through my study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Extreme weather events and natural disasters are the major cause of power outages in the United States, resulting in significant economic, social, and physical disruptions and cause considerable inconvenience for residents living in disaster areas [1]. It is estimated that only storm related outages cost the U.S. economy between $20 billion and $55 billion annually [2]. Various events have different characteristics and behaviour, however, the aftermath of all these events on the power grid is the loss of components and potential power outages.

Utilities and local governments are dealing with rising expectations of uninterrupted service from electricity consumers to effectively respond to the outcome of these catastrophic occurrences. With the purpose of improving the power grid resilience, electric utilities in the U.S. are spending billions of dollars on proactive and preventive responses such as grid hardening [3].

An efficient prediction of the probable damages to power grid components due to extreme weather events is a key step for developing efficient response and recovery models and performing preventive actions to encounter minimum damage. Among all types of extreme events, hurricanes are notably recognized as one of the most recurring events in the United States, mostly occurred by the Atlantic Ocean throughout Gulf of Mexico, from Maine to Texas [1]. In this work, hurricanes are explored not only because they cause the most widespread and long-lasting outages in the United States [4], but also because weather forecasting approaches that can predict a hurricane's arrival and characteristics (wind-speed, hurricane type, duration etc.) are optimally advanced to determine the probable impact in a localized region [5]. This work tackles the important problem of power grid resilience improvement in response to extreme weather events, in particular hurricanes, using machine learning. Different classification approaches such as Logistic Regression, Support Vector Machine (SVM) on different features are trained and evaluated in this work. The model is trained on artificial data and historical data from storm-related damages to predict component outages.

If the impacts of these events on the power grid are accurately predicted, grid operators can deploy a range of mitigation, response, and recovery actions to considerably reduce the undesirable socioeconomic aftermath. This work proposes a computationally-efficient and economically-viable grid hardening model in response to ongoing challenges and urgent needs in designing more resilient power grids. First, the state of each component is predicted using a SVM which is trained on historical data. Then, these predictions are fed to a hardening model, which takes grid resilience and economic needs into consideration. Different from existing literature in hardening and resilience enhancement, this work identifies that investments targeted at resilience enhancement would indeed impact power grid resilience and economic operations. The proposed grid hardening model determines the economically optimal set of candidates to be deployed for enhancing system resilience under prevailing uncertainties, while ensuring an adequate and secure supply of forecasted loads under normal, contingency, and extreme conditions.

The rest of the chapter is organized as follow: Section 1.1 reviews the importance of power system resilience and introduces some of the existing work on improving power system resilience. Section 1.2 presents the literature on machine learning approaches in system resilience and introduces logistic regression and SMV methods to estimate and model the system components that can potentially fail during a predicted hurricane. The importance of grid hardening in power system resilience is presented in Section 1.3. Finally, an overview of the contributions in this thesis are presented in Section 1.5.

1.1. Power System Resilience

Resilience denotes the capability of a system to absorb and to adapt to external shocks, which is an important characteristic expected from critical lifeline systems such as electric power grids [6]. There are several types of external shocks to the power grid, most notably extreme events which include adverse weather events and natural disasters that are known to cause considerable negative impacts not only on the system itself but also on the society in general. Among these extreme events, hurricanes are known to be the most frequent extreme event in the United States, mainly occurred along the Atlantic Ocean and Gulf of Mexico [1]. The devastating aftermath of these events calls for disruptive strategies to ensure that the power grid can still supply electricity to customers, or even if considerably impacted, can quickly bounce back from the contingency state to its normal operational condition. In this case, an accurate forecasting of the likely hurricane impacts on the power grid can be of significant value as it can be leveraged in achieving enhanced grid resilience. This work proposes a machine learning based method for predicting the state of the power grid components in response to upcoming hurricane strikes.

The concept of resilience for complex systems was originally introduced by Holling [7] in the ecology area. Holling defined the resilience of a system as the rate and speed of returning to normal conditions after an extreme event. The intent of resilience study is to anticipate the unexpected change due to failure, considering that systems have limits and gaps, and the atmosphere constantly affects both regarding design and external shocks [8]. Improving resilience in power systems is extensively discussed in the literature including research work on system modelling, resource allocation, and optimal scheduling for enhancing grid resilience, among others.

In [9], the significance of geographic and cascading interdependencies are highlighted which are associated with urban infrastructure, and a general method to describe infrastructure interdependencies is proposed. In [10] the impact of resilient systems on diminishing the probabilities of failure in urban infrastructure is analyzed. This concept was extended into other systems including the power grids. In [11] an approach for calculating the resilience of a single infrastructure and its components is proposed. In [12] a proactive resource allocation method aiming to repair and recover power grid after extreme events is proposed. In [9] and [10] a proactive recovery framework of power grid components is introduced which develops a stochastic model for operating the components prior to the event, followed by a deterministic recovery model for managing resources after the event. In [15] a restoration model is proposed based on power flow constraints which identifies an optimal schedule using the macroeconomic concept of the value of lost load (VOLL) in order to minimize the economic loss due to load interruptions in the post-disaster phase. A decision-making model, based on unit commitment solution and system configuration, is proposed in [16] to find the optimal repair schedule after a hurricane and in the restoration phase of a damaged power grid.

In [17], a power grid resilience index is proposed by analyzing the process of generation, transmission, and consumption of electricity in various countries. The geometric mean of several factors such as the generation efficiency of non-renewable fuel dependence, the distribution efficiency, the carbon intensity, and the diversity are considered to develop the resilience index. However, an index for individual components in the system is not considered in the methodology. In [18], a methodology to calculate resilience index of power delivery systems in post-event infrastructure recovery is proposed. A multi-infrastructure system including electric power delivery, telecommunications, and transportation is considered and the resilience measures of fragility and quality are combined with the input-output model of these infrastructures. The proposed index is evaluated by the data collected from post-landfall of Hurricane Katrina to assess the resilience and interdependence of a multi-system networked infrastructure during natural extreme events. The study in [19] proposes a framework for resilience enhancement of urban infrastructure systems. The time-dependent expected resilience metric is built on performance and response of the power grid following an extreme event. The process is performed in the stages of disaster prevention, damage propagation, and assessment and recovery. The hurricane resilience of electric power grids is quantified through a probabilistic modeling approach in [20], using a Poisson process model for hurricane occurrence, component fragility models, and a grid restoration model with component repair priority. The model is then calibrated using actual customer outage and power grid restoration data in Harris County, Texas in the aftermath of Hurricane Ike in 2008.

1.2. Machine Learning Approaches in Power System Resilience

Machine learning is an application of Artificial Intelligence (AI) that provides the system the ability to learn from historical data and to make predictions without being explicitly programmed. In many problems, a closed formulation of the problem and its solution cannot be easily derived. Machine learning investigates the algorithms that are capable of learning from and making forecasts from data. These algorithms can categorize the observed data for classification (supervised learning), combine similar patterns for clustering (unsupervised learning), and predict the output of the system based on its past behavior and historical data (regression modeling) [21]. Figure 1-1 shows the different aspect of Artificial Intelligence and machine learning.

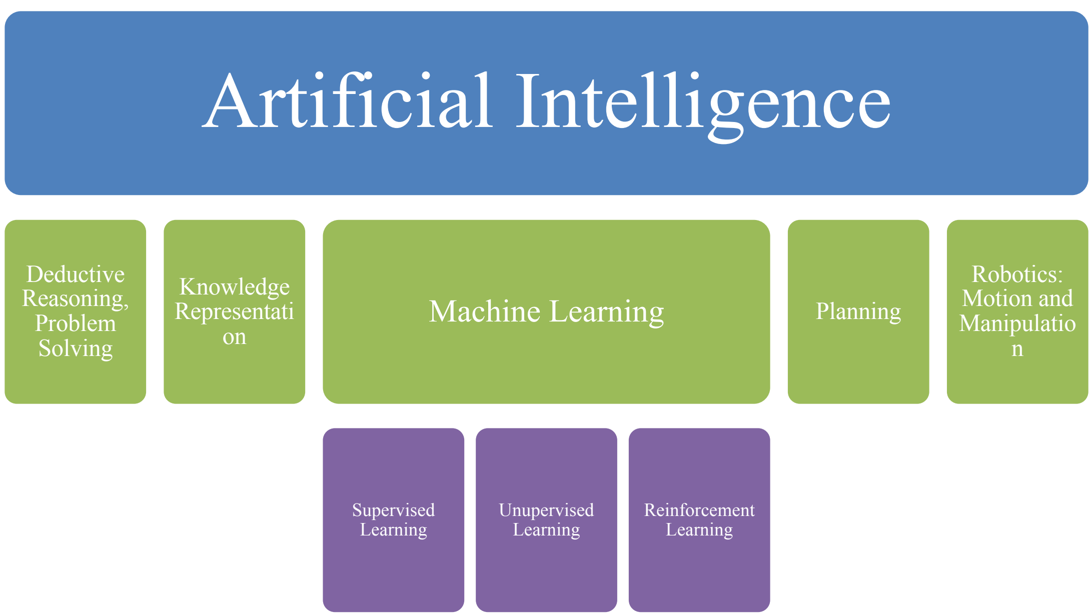


FIGURE 1-1- ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Predictive analytics and emerging applications of machine intelligence tools are shaping every aspect of our daily lives. Data has become the epicenter of the modern decision making by policy makers, corporations, and enterprises. Utilities and local governments are facing increasing expectations from their customers and constituencies to effectively respond to the aftermath of the catastrophic events such as hurricanes that can affect the quality of life of the communities and interrupt the business continuity. In this climate, the concept of resilience enhancement has become an important risk management measure in addressing these challenges.

Machine learning approaches have been utilized in a considerable number of research efforts in the power and energy sector [22]. Machine learning has been applied to several power grid related problems such as forecasting (using extreme learning artificial neural networks) [23], security assessment (using decision tree induction, multilayer perceptions, and nearest neighbor classifiers) [24], risk analysis (using parametric, semi-parametric, and non-parametric regression models, artificial neural network, and support vector machine) [25], distribution fault identification (using artificial neural network and support vector machine) [26], and power outage duration prediction (using regression models, regression trees, Bayesian additive regression trees, and multivariate additive regression splines) [27].

Security assessment is one of the most versatile machine learning applications in power grids with the applications from pattern recognition [22], decision tree induction, and nearest neighbor classifiers [28], to name a few. Forecasting arises as another popular application of machine learning. A number of Artificial Neural Networks (ANNs) have been proposed for short-term load forecasting [29] and wind power forecasting [30]. Some other examples of machine learning applications in power grids include risk analysis using regression models, ANNs, and Support Vector Machine (SVM) [25], distribution fault detection applying ANNs and SVM [26], and power outage duration prediction using regression models and regression trees/splines [27].

1.3. Grid Hardening

With the purpose of improving the power grid resilience, electric utilities in the U.S. are spending billions of dollars on proactive and preventive responses such as grid hardening [3]. Grid hardening represents the physical and nonphysical improvement to the electricity infrastructure to make it less susceptible to adverse extreme events improving grid resilience and enabling the grid to withstand the impacts of extreme events with the least possible outages [31]. Physical hardening refers to installing new facilities and modifying the current grid topology. Nonphysical hardening options represent adjustments in consumption, generation, and power flow patterns. Current electric power grid hardening practices merely focus on the aspect of improving system resilience in responding efficiently to an extreme event.

There are a limited number of studies on the efficient hardening of electric power grids in response to extreme events. In [32] a comprehensive strategy for mitigating hazards is proposes which aims at creating resilient cities that are able to withstand disasters. In [33], hurricane damage predictions and topological assessment are combined to characterize the impact of hurricanes on power grid reliability. Component fragility models are applied to predict failure probability for individual transmission and distribution components. The research shows that topological features, such as network mesh structure, centrality, and the compact irregular ring mesh topology, need to be considered in hurricane hardening activities. A comprehensive survey of models and algorithms for emergency response logistics in electric distribution systems is presented in [34], [35].

Analysis of cost-effectiveness of engineering solutions to harden the electric power infrastructure is another area which has been covered in the literature. In [36], a probabilistic model for analyzing electric power infrastructure risk mitigation investments is proposed which aim to evaluate the tradeoffs between wetland restoration and infrastructure hardening for the electric power grid. The results indicate that wetland restoration and undergrounding of power infrastructure is not preferred over keeping them without wetland protection. The current practice of utilities and government agencies for hardening the power grid has been reflected in several publications and presentations. For example, the hurricane hardening efforts in state of Florida is described in [37], which presents an overview of storm hardening strategies and a discussion on the progress of a utility's hardening initiative and current research efforts on cost/benefit analysis for hurricane.

In practice, multiple grid hardening options may be available for system planners. Finding the most suitable option is a challenging task as several factors are involved in the modelling, and furthermore mathematical approaches may not be able to fully capture the behaviour and aftermath of the events. Given the amount of data that exists on previous hurricanes and the complexity of the system, machine learning can be a viable approach to tackle this problem. Machine learning approaches can learn from historical data and to make predictions without being explicitly programmed. Machine learning approaches are utilized in a considerable number of research efforts in the power and energy sector, such as security assessment [22], load forecasting [29], distribution fault detection [26], and power outage duration prediction [27][38][39][40].

1.4. Probabilistic load curtailment estimation

Having a precise prediction of the potential impacts of an upcoming hurricane plays a vital role in improving the power system resilience by helping identify the most efficient resource allocation [41]. Resource allocation before and after a hurricane is a well-studied topic in power systems. In [42], a proactive resource allocation model is proposed to repair and recover power system infrastructure located in a hurricane-impacted region, attempting to develop a decision-making tool which ensures the least potential damages in an efficient manner. In [43][44], a proactive recovery framework of power system components is presented based on a stochastic model for operating the components prior to the event, followed by a deterministic recovery model to manage the available resources after the event. In [45], an optimal restoration model is proposed to minimize the economic loss due to power supply interruptions during the post-disaster phase. In [46], a decision-making model is introduced based on unit commitment constraints and system configuration. The objective of the proposed model is to determine the optimal repair schedule after an extreme event and during the restoration phase.

Pre-hurricane scheduling specifically plays an important role in improving system resilience. A resilience-constrained unit commitment (RCUC) model is proposed in [47] which ensures a resilient supply of loads even in case of multiple component outages. In many of the related works on hurricane modeling, the impact of the hurricane on the power system is the input to the model or determined by a stochastic model. Machine learning, however, is recognized as an efficient method in predictive analytics and data analysis to identify the likelihood of future outcomes based on historical data [21]. In particular, SVM is a popular machine learning method for data classification (supervised learning) which is developed on the basis of statistical learning theory and structural risk minimization [32, 33]. SVM has numerous advantages such as providing a global solution for data classification as well as great generalization capability. The achieved results in several studies illustrate SVM as one of the most accurate methods in several applications such as generation forecasting [34, 35], load forecasting [36], fault detection [37], power quality disturbance monitoring [38], and transient stability analysis [39]. SVM has also shown a superior performance in predicting possible outages of power system components in response to extreme events [40]. In [41], a three-dimensional SVM is proposed to predict the outage of power system components in response to an extreme event, where its accuracy–uncertainty tradeoff is leveraged to achieve more precise results.

Despite the good performance of SVM in several applications, the performance of SVM drops significantly when faced with imbalanced datasets, for example when the number of negative instances far outnumbers the positive instances, or vice versa [42]. Twin support vector machine (TWSVM) is the answer to this, as an efficient machine learning approach which is suitable for complex classification problems. TWSVM classifies the patterns of two classes by using two non-parallel hyperplanes [43]. Since two hyperplanes are defined as representatives of each class, TWSVM can handle imbalanced datasets much better than the traditional SVM [44].

In this paper, a TWSVM classification method is trained to find the operational state of each component by considering the path and the intensity of the hurricane, as well as the distance of each component from the center of the hurricane. A posterior probability model is consequently applied to the output of the TWSVM model to estimate the outage probability of each component. Having an accurate estimation of probable outages plays a vital role in responding to an upcoming hurricane.

Unlike the existing work on outage prediction and extended outage consideration in security-constrained unit commitment (SCUC), including the previous work of authors in [5, 6], this paper considers the probability of outage obtained by a machine learning approach in scheduling. TWSVM is chosen for its performance in complex intertwined classification problems and when dealing with imbalanced datasets. This can be potentially problematic since the data of past hurricanes are imbalanced, i.e., the number of non-operational components is far less than the number of operational components. The merit behind proposed probabilistic load curtailment estimation model is that it considers all contingency scenarios with their probability and hence the most probable scenario or the scenarios with most load curtailment can be recognized. The predicted outage and estimated outage probability can be useful for electric utilities to assess their risk and allocate necessary resources and repair crews to prepare for and recover from hurricanes in a considerably shorter time-frame.

1.5. Contributions

The contributions of this work are as follows:

1.5.1. Logistic Regression Based Power Grid Outage Prediction

In this work, an outage prediction model based on logistic regression is proposed to determine the probable outage of power grid components in response to an imminent hurricane. The proposed logistic regression model is used as a viable machine learning method to determine the decision boundary between damaged (on outage) and operational (in service) components in response to a hurricane. The logistic regression method is simple, fast, robust, and can efficiently handle the complexity of the decision boundary in terms of characteristic parameters. The regression model is applied considering the wind speed and the distance of each component from the center of the hurricane as two major features to find the state of each component after an extreme event.

1.5.2. SVM Based Power Grid Outage Prediction

Despite the acceptable performance of the proposed logistic regression model, logistic regression requires much more data to achieve stable and meaningful results compared to other prediction models, such as support vector machine. In addition, the characteristic parameters of logistic regression increase exponentially as number of features increases. Hence, an SVM-based method is proposed and adopted to predict the state of each component in the aftermath of an imminent hurricane. Particularly, a multi-dimensional SVM is proposed which considers the associated resilience index, i.e., the infrastructure quality level and the time duration that each component can withstand the event, as well as predicted path and intensity of the upcoming extreme event. The outcome of the proposed model is the classified component state data to two categories of outage and operational, which can be further used to schedule system resources in a predictive manner with the objective of maximizing its resilience.

Furthermore, a new three-dimensional Support Vector Machine (SVM) for power grid component outage prediction is proposed which leverages its accuracy-uncertainty tradeoff to achieve highly accurate result. The new proposed SVM considers the component deterioration level as an additional critical and decisive factor. The objective of this model is to tailor the gap made by the decision boundary to increase prediction accuracy. The proposed SVM model is used to define a clear gap between the outage and operational states. This gap is considered as an uncertain area, which is further utilized to improve the accuracy of the predicted states. It should be noted that such capability is not available using a logistic regression.

1.5.3. Load Curtailment Estimation in Response to Extreme Events

A minimum load curtailment problem is proposed and formulated to estimate the amount of load curtailment considering the predicted outage states. The predictions are integrated into a minimum load curtailment model to estimate the potential nodal load curtailments—which are of utmost importance for grid operators in order to identify critical and prone-to-curtailment areas to proactively mobilize the restoration resources.

The proposed framework enables one to effectively identify the critical components in the power system and prioritize the limited restoration resources. Given the crucial importance of accurate power grid outage prediction, this model provides a practical forward-looking framework for utilities, local governments, and policy makers for a risk-informed operations management, emergency response planning, humanitarian logistics, and restoration of the life-line power grid infrastructure in both strategic level and real-time basis.

1.5.4. Machine Learning Assisted Power Grid Hardening

A new hardening a machine learning based grid hardening model is proposed with the objective of improving power grid resilience in response to extreme weather events. The proposed hardening model determines strategic locations for placement of distributed generation (DG) units. In contrast to existing literature in hardening and resilience enhancement, this model co-optimizes grid economic and resilience objectives by considering the intricate dependencies of the two. This proposes approach is a computationally-efficient and economically-viable grid hardening model in response to ongoing challenges and urgent needs in designing more resilient power grids. Particularly, the predictions from previous contributions are fed to a hardening model, which takes grid resilience and economic needs into consideration. Different from existing literature in hardening and resilience enhancement, this model identifies that investments targeted at resilience enhancement would indeed impact power grid resilience and economic operations. The proposed grid hardening model determines the economically optimal set of candidates to be deployed for enhancing system resilience under prevailing uncertainties, while ensuring an adequate and secure supply of forecasted loads under normal, contingency, and extreme conditions.

1.5.5. Probabilistic load curtailment estimation

A three-step sequential method in identifying such load curtailments prior to hurricane. This work considers the probability of outage obtained by a machine learning approach in scheduling. TWSVM is chosen for its performance in complex intertwined classification problems and when dealing with imbalanced datasets. This can be potentially problematic since the data of past hurricanes are imbalanced, i.e., the number of non-operational components is far less than the number of operational components. The merit behind proposed probabilistic load curtailment estimation model is that it considers all contingency scenarios with their probability and hence the most probable scenario or the scenarios with most load curtailment can be recognized. The predicted outage and estimated outage probability can be useful for electric utilities to assess their risk and allocate necessary resources and repair crews to prepare for and recover from hurricanes in a considerably shorter time-frame.

In the first step, a twin support vector machine (TWSVM) model is trained on path/intensity information of previous hurricanes to enable a deterministic outage state assessment of the grid components in response to upcoming events. The TWSVM model is specifically used as it is suitable for handling imbalanced datasets. In the second step, a posterior probability sigmoid model is trained on the obtained results to convert the deterministic results into probabilistic outage states. These outage states enable formation of probability-weighted contingency scenarios. Finally, the obtained component outages are integrated into a load curtailment estimation model to determine the expected potential load curtailments in the grid. The simulation results, tested on the standard IEEE 118-bus system and based on synthetic datasets, illustrate the high accuracy performance of the proposed method.

CHAPTER TWO: POWER GRID OUTAGE PREDICTION

In this chapter, the model outline and formulation of the proposed approaches to predict power outages in response to hurricane is presented. For this purpose, two machine learning approaches are studied in this work. Section 2.12 introduces the proposed logistic regression-based approach and evaluate the performance the performance of the obtained decision boundary in efficiently predicting component outages. Despite the acceptable performance of the proposed logistic regression model, it requires much more data to achieve stable and meaningful results compared to other prediction models, such as support vector machine. Section 2.2 introduces an SVM-based method which is proposed and adopted to predict the state of each component in the aftermath of an imminent hurricane. The model is developed based on three distinct features of component deterioration, distance from the extreme event, and the intensity of the extreme event, and is analytically investigated to exhibit its acceptable performance.

2.1. Logistic Regression-Based Power Grid Outage Prediction

Consider the power grid in which a subset of its components is located in the path of an upcoming hurricane. The path and the intensity of the hurricane can be forecasted based on the weather data obtained from weather forecasting agencies. Two states are considered for each component in the path of the hurricane: damaged (on outage) and operational (in service). The decisive factors to determine these states are the hurricane wind speed (which also determines the category of the hurricane) and the component distance from the center of the hurricane, respectively represented here by parameters x_1 and x_2. Figure 2-1 depicts a schematic of the damaged and operational states (shown by crosses and circles, respectively) from historical hurricane data, as well as the decision boundary separating these two states. The probability of damage increases as the wind speed increases or the distance to the center of the hurricane decreases. Based on the available data, there should be a minimum wind speed to result in an impact to components (hence the intersection of the decision boundary with the x_1 axis). The goal is to determine the function representing the decision boundary, thus outages in response to imminent hurricanes can be effectively predicted.

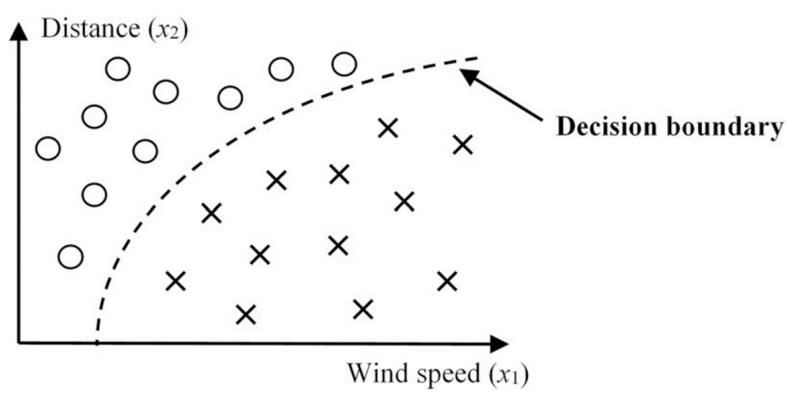


FIGURE 2-1- DAMAGED/OPERATIONAL STATES OF ELECTRIC POWER GRID COMPONENTS SEPARATED BY THE DECISION BOUNDARY

2.1.1. Logistic Regression

The logistic regression method [48] is used to determine the decision boundary. The decision boundary is defined by a second order polynomial based on the wind speed and the distance (1):

$$h(x,\ k) = k_0 + k_1x_1 + k_2x_2 + k_3x_1^2 + k_4x_2^2 + k_5x_1x_2 \qquad (1)$$

where k_j , $j = 1,\dots, 5,$ is the characteristic parameter to be determined. A second order function is considered for the function h to prevent overfitting. The classification function is denoted by *f(x, k)* and defined as a Sigmoid function, i.e.,

$$f(x,\ k) = \frac{1}{1\ +\ e^{-h(x,k)}} \qquad (2)$$

The Sigmoid function is depicted in Figure 2-2, which ensures that for positive values of *h(x, k)* a value of 1 is reached, while for its negative values, a value of 0 is reached (3).

$$f(x,k) = \begin{cases} 0 & t \le 0 \\ 1 & t > 0 \end{cases} \qquad (3)$$

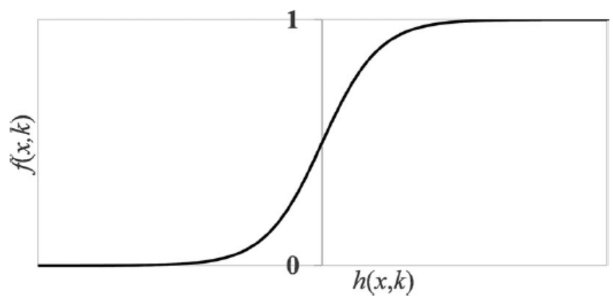


FIGURE 2-2- THE SIGMOID FUNCTION

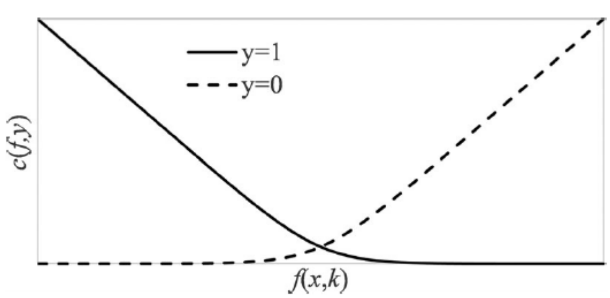


FIGURE 2-3- PROPOSED COST FUNCTION

This function nicely classifies the data based on the obtained function. If *h(x, k)*=0, the value of f will be 0.5, which shows the data is exactly on the decision boundary. To determine the characteristic parameter k_j , the cost function (4) is defined to minimize the errors between the fitted curve and the realized values from historical data:

$$J(k) = \frac{1}{m}\sum_{i=1}^{m} c\big(f(x,\ k),\ y\big) + \frac{\lambda}{2m}\sum_{j=1}^{5} k_j^2 \tag{4}$$

$$c(f, y) = -y\ log\big(f(x,\ k)\big) - (1-y)\ log\big(1 - f(x,\ k)\big) \tag{5}$$

where *m* is the number of training data points, and y is the actual state (*y* = 0 for damaged and *y* = 1 for operational).

This cost function, as shown in Figure 2-3, efficiently evaluates the classification function based on the obtained characteristic parameters by becoming equal to zero when the prediction is correct (i.e., *f(x, k)*=0 when *y* = 0, or *f(x, k)*=1 when *y* = 1) while becoming a very large number when the prediction is wrong (i.e., *f(x, k)*=0 when *y* = 1, or *f(x, k)*=1 when *y* = 0). The second term in (4) is added for regularization, which would ensure small values for characteristic parameters and accordingly a simpler decision boundary. Using regularization, some of the terms will be automatically eliminated if the second order function results in overfitting. The regularization parameter, λ, controls the tradeoff between keeping a small number of parameters and overfitting, which however is problem-dependent and needs to be carefully determined.

Once the cost function *J(k)* is minimized, the characteristic parameters are determined, hence we would have the decision boundary. The outcome of this method is the prediction function in the form of *f(x, k)*, with given values for k_j , that can predict the damaged/operational state of any power grid component based on the wind speed of an imminent hurricane as well as the distance of the component from the center of the hurricane.

To test the performance of the obtained decision boundary, the F_1-Score (6) will be examined on the test data:

$$F_1 = \frac{2PR}{P+R} \tag{6}$$

where P is the number of positive predictions divided by the total number of positive class values predicted (i.e., precision), and R is the number of positive predictions divided by the number of positive class values in the test data (i.e., recall). For example, in the case of the outage prediction problem, precision (P) is the number of correctly predicted outages divided by the total number of predicted outages, and recall (R) is the number of correctly predicted outages divided by the total number of actual outages. The F_1-Score will be a value between 0 and 1, where higher values represent a better prediction and justify the acceptable performance of the obtained decision boundary.

2.1.2. Numerical Simulation

A set of 1000 artificial data points is generated, based on a normal distribution, and used for training (80%), and validation (20%). The proposed method results in the following solution for the characteristics parameters in (1): $k_0 = 1.47$, $k_1 = -2.85$, $k_2 = 0.59$, $k_3 = -2.05$, $k_4 = 0.70$, and $k5 = -0.36$. Table 2-1 shows the obtained confusion matrix based on the calculated decision boundary on validation set. The F_1-Score is calculated as 0.9027 ($R = 0.8759$, $P = 0.9311$) which shows the acceptable performance of the proposed method.

TABLE 2-1- CONFUSION MATRIX BASED ON THE LOGISTIC REGRESSION CALCULATED DECISION BOUNDARY

	Predicted	
Actual	Operational	Damaged
Operational	425	67
Damaged	35	473

2.2. SVM Based Power Grid Outage Prediction

Despite the acceptable performance of the proposed logistic regression model, logistic regression requires much more data to achieve stable and meaningful results compared to other prediction models, such as support vector machine. In addition, the characteristic parameters of logistic regression increase exponentially as number of features increases. Hence, an SVM-based method is proposed and adopted to predict the state of each component in the aftermath of an imminent hurricane.

2.2.1. Support Vector Machines

SVM is a discriminative classifier that defines a separating hyperplane between two classes. The best hyperplane in SVM is considered as the hyperplane with the widest gap between the classes which decreases the risk of miss-classifying and increases the generalization of the classifier. This gap is usually referred to as margin, where SVM intends to maximize this margin between the classes.

The details of the SVMs are fully described in the literature [49], so only a brief introduction to SVM in three-dimensional space is presented in this section. Consider m training samples $x_i \in \mathrm{R}^3$, $i=1,\dots,m$ in a binary classification problem. . The linear decision is function $f(x)=\mathrm{sign}(w^T x+b)$, $x_i \in \mathrm{R}^3$, where w is the weight vector which defines a direction perpendicular to the hyperplane of the decision function, while $b \in \mathrm{R}$ is a bias which moves the hyperplane parallel to itself. The optimal decision function given by support vectors is the solution of the following optimization problem:

$$\begin{aligned} &\min \frac{1}{2}\|w\|^2 + c\sum_{\beta=1}^{m} \varepsilon_\beta \\ s.t. \quad & y_\beta\left(w^T x_\beta + g\right) \geq 1, -\varepsilon_\beta, \qquad \beta = 1, \ldots\ldots, m \\ & \varepsilon_\beta \geq 0, \qquad\qquad\qquad \beta = 1, \ldots\ldots, m \end{aligned} \tag{7}$$

where w is the normal vector to the hyperplane separating training examples, $|g|/||w||$ is the perpendicular distance of the hyperplane from the origin, and c is a penalty parameter. When $c \rightarrow \infty$, SVM does not allow any training errors (hard margin classification) and when $0 < c < \infty$, the model allows some training errors, and hence allowing separating nonlinear examples (soft margin). This is a quadratic programming problem which can be solved for the problem's Lagrange duality multiplier $\alpha \in \mathrm{R}^3$ as follows:

$$\begin{aligned} &\max_{\alpha} \quad -\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \left(x_i . x_i\right) + \sum_{i=1}^{m} \alpha_i \\ &s.t. \quad 0 \leq \alpha \leq C, \ \sum_{\mathrm{i=1}}^{\mathrm{m}} \alpha_i y_i = 0 \end{aligned} \tag{8}$$

In order to solve the duality problem, many analytical approaches are proposed in the literature, depending on the size of dataset and memory limitation considerations. Sequential Minimal Optimization (SMO) [50] is one of the analytic approaches that is used to solve the quadratic programming (QP) problem (2) in many SVM toolboxes such as LIBSVM tool in MATLAB [51]. SMO breaks the QP problem into multiple smaller subproblems, which are then solved analytically. SMO picks two support vectors, finds corresponding Lagrange multipliers and repeats this process until reaching convergence (within a user-defined tolerance) or a maximum number of iterations.

By solving the duality problem (8), the final hyperplane only depends on the support vectors (i.e., sample points that are in the margin) and SVM needs to find only the inner products between the test samples and the support vectors. Figure 2-4 shows the support vectors and optimal hyperplane in a separable two-class classification of SVM. In regards to the objective of this work, Figure 2-4 also shows the support vectors and optimal hyperplane to separate outage from operational components based on the associated resiliency index, distance from the center of the hurricane, and the wind speed.

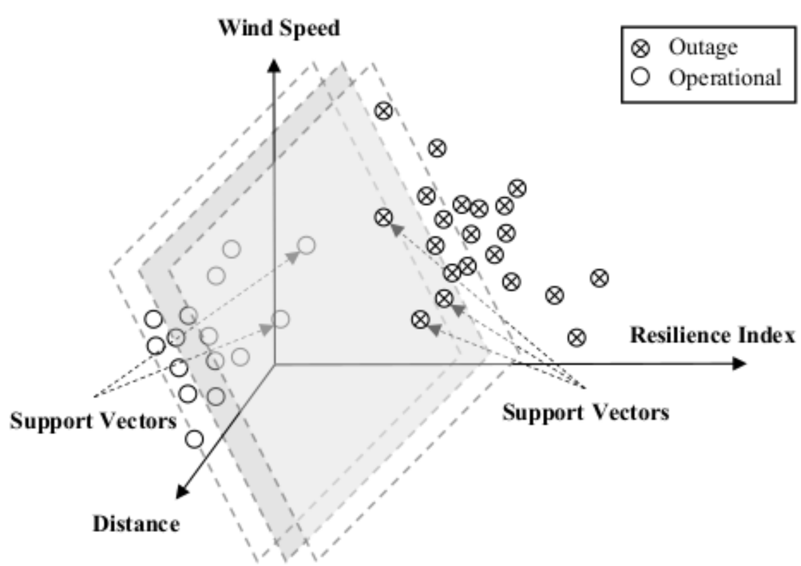


FIGURE 2-4- SUPPORT VECTORS AND OPTIMAL MARGIN IN SVM

The idea of the maximum-margin hyperplane, which is discussed above, is based on the assumption that training data are linearly separable. To apply SVM to nonlinear data (which often is the case, especially in the case of the hurricane data), kernel methods [49] can be used. The idea of a kernel method (or as sometime called kernel trick) is to map the input space into a linear separable feature space, usually a higher dimension, where the linear classifiers can separate two classes (Figure 2-5). As shown in Figure 2-5, the linearly inseparable data in a two-dimensional space can be linearly separable in higher dimensions (three dimensions in this figure). Kernel trick simply states that for all x_1 and x_2 in the input space, a certain function $k(x_1,x_2)$ can be replaced as inner product of x_1 and x_2 in another space. For example, a Gaussian kernel can be defined as:

$$k\left(x_i, x_j\right) = e^{-\frac{1}{2\sigma^2}\left\|x_i - x_j\right\|^2} \tag{9}$$

where σ^2 is the parameter of the kernel defined by the user. In practice, the best kernel is found by experiment while adjusting kernel parameters via a search method to minimize the error on a test set.

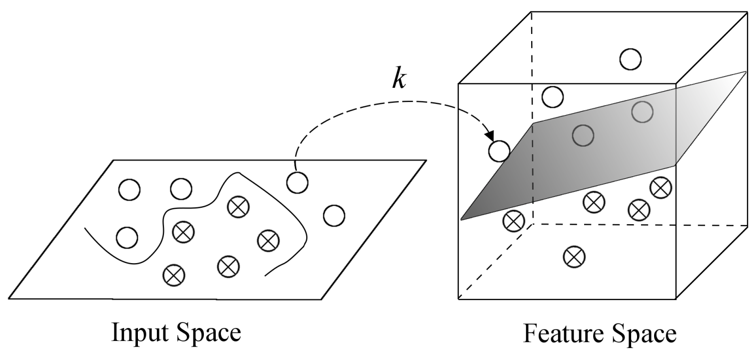


FIGURE 2-5- THE KERNEL METHOD IN SVM.

2.2.2. Resilience index as a Component Features

A feature, in machine learning, is defined as an individual measurable property of a phenomenon being observed [21]. Selection of discriminating, independent, and informative features plays a critical role in the performance of the classification method. Various features can be defined to determine the state of the components in response to a hurricane strike. In [38], the wind speed and the distance of the each component from the center of the hurricane are proposed as response to a hurricane.

Although these features are obviously adequately informative, they do not provide information about the component itself. Resilience index of components is also an important factor during weather-related events. Similar to [20], we quantify the hurricane resilience of the electric power grid using a probabilistic modeling approach. For the sake of illustration, only the Poisson process model of hurricane occurrence during a given time period along with fragility models are considered in this work. Other factors used in [20] such as DC power flow, power grid restoration and component repair priority are not considered in this index. However, the proposed model is a general framework and can be extended to other resilience indices. Based on this, hurricanes are described by a Poisson process of constant rate λ_h such that the time interval between successive hurricane events has an exponential distribution with a probability function of

$$f(t) = \begin{cases} \lambda_h e^{-\lambda_h t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{10}$$

Similar to [20] and based on historical data from 1900 to 1999 [52], the annual occurrence rate of hurricanes is considered as λ_h = 1/7 per year, and the probability of a hurricane belonging to each category is respectively calculated as 0.53, 0.19, 0.15, 0.08, and 0.05. In this work, we consider resilience index for four components: a) generation units, b) transmission lines, c) distribution lines, and d) substations. For their flexible analytical properties, similar fragility models following a normal distribution, are considered for all four categories with probabilities of low, moderate, severe, and complete. Resilience index is then considered as the average of fragility model and the probability of the hurricane. The category of hurricane, the distance of each component from the center of the hurricane, and the calculated components resilience index are investigated as three main features to predict the state of each component in response to the hurricane.

2.2.3. Leveraging Accuracy-Uncertainty Tradeoff

SVM defines a clear margin of support vectors. The majority of miss-classification happens in the area near the decision boundary. In SVM, the optimal margin is found by checking each and every data point against the condition stated in (7), then the vectors of data points that lie on either side of the hyperplanes become the support vectors. This is usually found using a numerical approach such as Sequential Minimal Optimization (SMO) [50]. The margin is defined as the distance between two closest support vectors, as in (11):

$$\text{distance} = \frac{w^T x}{\|x\|^2}. \tag{11}$$

In this work, the area between the support vectors (margin from the decision hyperplane) is considered as an uncertain area. To improve the classification accuracy, the SVM gap is extended by decreasing the penalty coefficient, so the estimated states in the certain area will become more accurate. Figure 2-6 depicts the optimal margin and the uncertain area for a two-dimensional classification (for the mere purpose of clarity). The figure on the right has a smaller penalty coefficient, hence a larger margin allowing miss-classification, and thus, a higher prediction accuracy. As shown, by increasing the margin, more missclassification occurs, in which the miss-classified data are located within the uncertain area. Allowing a wider gap significantly increases the accuracy of the model at the small expense of few miss-classified data.

The aim of the SVM is to fit a hyperplane based on the data points at the edge of each class, or "support vectors." One of the advantages of SVM over other classification techniques is that it only considers support vectors (i.e., data points on the border of the boundary) when defining the separating hyperplanes and therefore it can offer a better generalization compared to other techniques such as logistic regression [4]. Also, SVM approximates the structural risk minimization principle in statistical learning theory rather than the empirical risk minimization method [8]. This property makes the SVM less prone to overfitting the training dataset. Figure 2-7 depicts the flowchart of the proposed method (components in the margin of SVM are considered as uncertain).

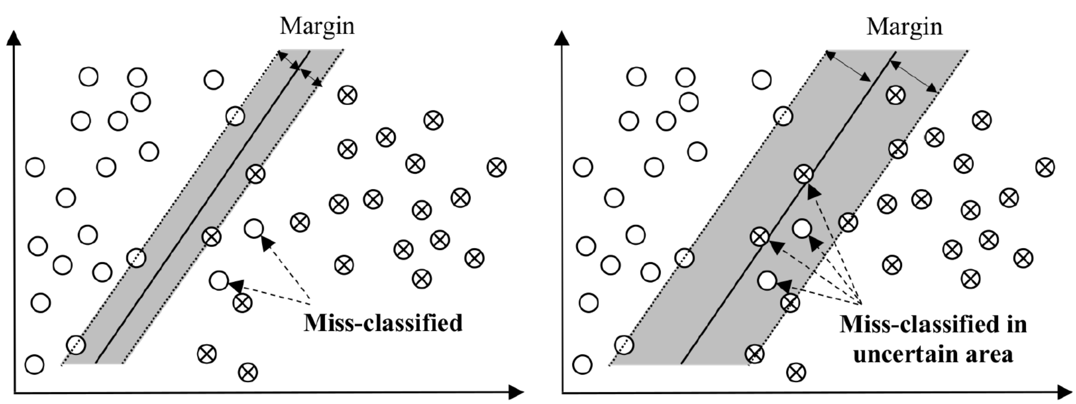


FIGURE 2-6- THE OPTIMAL MARGIN IN SVM AND MISS-CLASSIFIED SAMPLES.

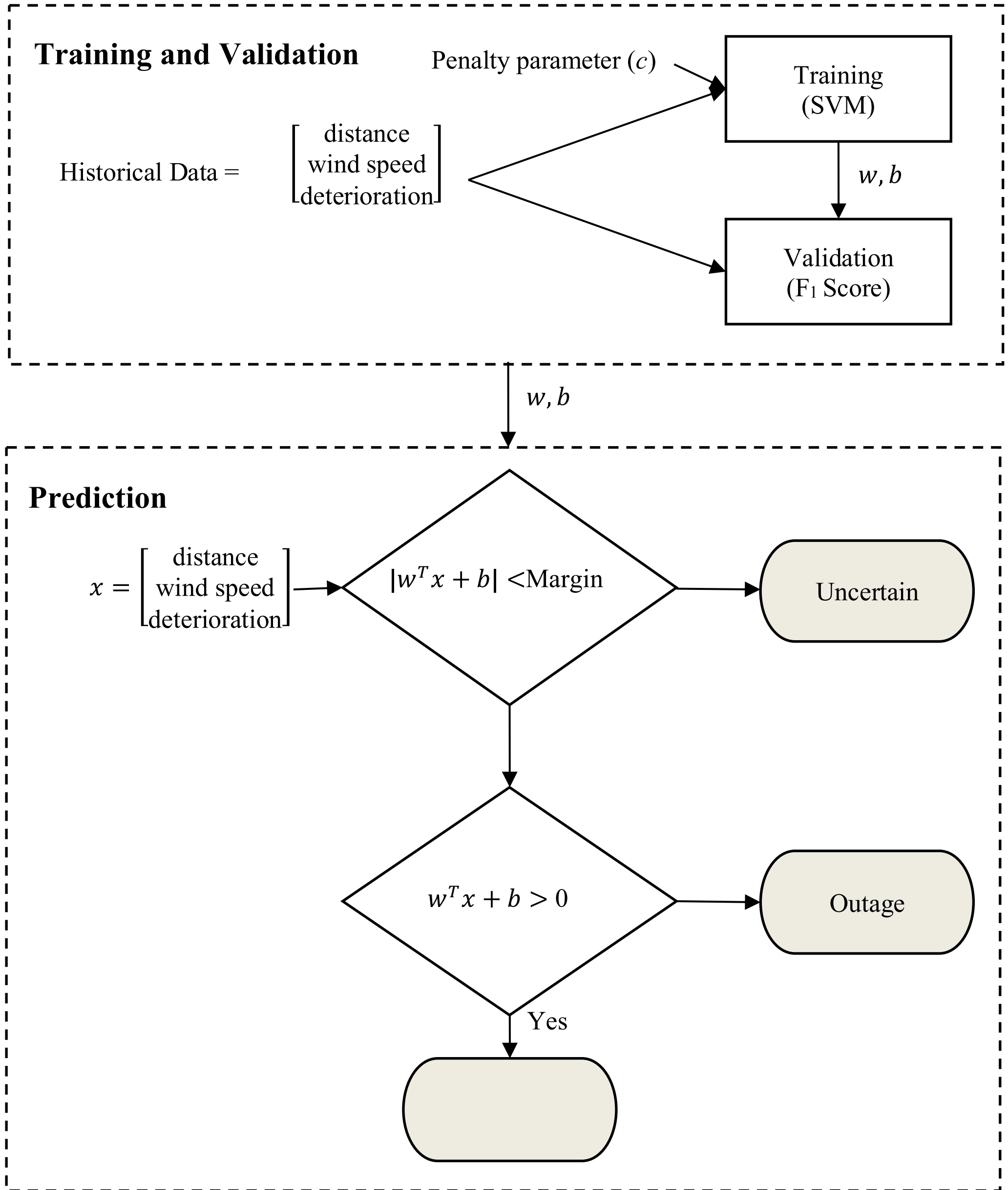


FIGURE 2-7- FLOWCHART OF DETERMINING THE COMPONENT STATE.

2.2.4. Numerical Simulation

Scarcity of readily available datasets still remains a challenge for research community and industry practitioners. However, the limited historical data on past extreme hurricanes at the component granularity level shall not preclude methodological developments in critical areas including in machine learning systems. Therefore, in this work, a synthetic set of 1000 sample data is generated to train the SVM model, considering half of the samples in outage state and the other half in the operational state. The generated samples follow a normal distribution function of one-minute sustained wind speed of different Saffir-Simpson Hurricane Scale categories with a small Gaussian noise. The features are normalized to [0,1] based on the maximum considered values of wind speed and distance. Figure 2-8 shows the generated synthetic data.

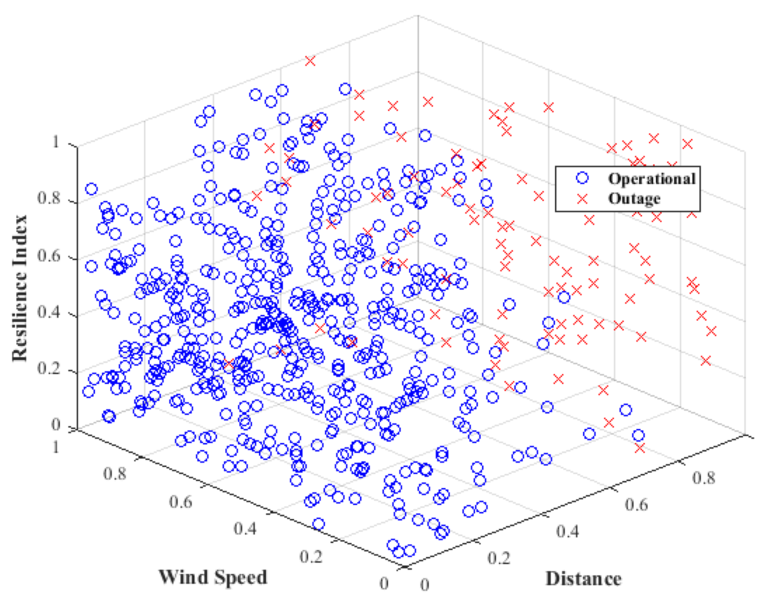


FIGURE 2-8- GENERATED SYNTHETIC DATA FOR SVM TRAINING AND VALIDATION

To evaluate the performance of the classifier, usually a subset of the historical dataset is reserved as holdout sample for model validation. k-fold cross-validation is a common validation technique for assessing the results of a classification system and evaluating how well it can generalize on a dataset [53]. In k-fold cross-validation, the dataset is randomly partitioned into k equal sized subsamples. A single subsample is reserved as the validation/test set, and the other $k-1$ subsamples are used as training data for the model. This process is iterated for k times (the number folds), where each of the k subsamples is used only once for the validation. The k results from the folds are accordingly averaged to obtain a single estimation.

2.2.4.a) Multi-Dimensional SVM with Resilience index as Component Feature

In this case study, the proposed SVM is trained on historical data with three features, namely the resiliency index of the component, the distance of the component from the center of the hurricane, and the category of the hurricane which is determined based on the wind speed.

A k-fold cross validation (k=5) is performed to measure the performance of the proposed model. Different kernels (linear, polynomial Quadratic, Cubic, and Gaussian) with various penalty parameters (c=0.01, 0.1, 1, 10, 100) are examined. Since the considered dataset is relatively small, an off-the-shelf SVM model implemented in LibSVM [51] is used in this work. In the proposed work, the SMO tolerance for convergence is set to 1e-3 and the maximum number of iterations is set to a large value (15000 iterations). In practice, since the considered dataset is relatively small, it converges in about 350 iterations for different folds. Table 2-2 shows the average F_1-Score for various penalty parameters and kernel shapes. As it is shown, SVM with Gaussian kernel and c=1 offers the best performance among other settings.

A third order polynomial logistic regression model is also trained and examined in the same fashion (i.e., k-fold cross-validation with k=5) to predict the component outages. Table 2-3 compares evaluation metrics of SVM with different kernels (using penalty parameter c=1) and a third order polynomial logistic regression model. As shown, among the trained models, Gaussian kernel SVM had the best overall classification accuracy with a precision of 0.893, a recall of 0.826, and overall F_1-Score of 0.858. Comparing the result of logistic regression with the proposed SVM indicates that the proposed SVM approach has a better performance in both accuracy and F_1-Score.

Table 2-4 shows confusion matrix of predicting components as operational and outage using Gaussian kernel SVM. The proposed model can predict outage and operational states with the accuracy of 90.2% and 82.6%, respectively.

TABLE 2-2- AVERAGE F1-SCORE OF SVM WITH VARIOUS PENALTY PARAMETERS "C" AND KERNELS USING 5-FOLD CROSS-VALIDATION

Kernel	c=0.1	c=1	c=10	c=100
Linear	0.845	0.845	0.846	0.846
Quadratic	0.858	0.856	0.855	0.857
Cubic	0.855	0.854	0.840	0.754
Gaussian	0.857	**0.858**	0.850	0.847

TABLE 2-3- COMPARISON OF THE PERFORMANCE OF SVM WITH VARIOUS KERNELS AND THE LOGISTIC REGRESSION METHOD.

	Accuracy	**Precision**	**Recall**	**F1-Score**
Linear SVM	0.847	0.853	0.838	0.845
Quadratic SVM	0.863	0.898	0.818	0.856
Cubic SVM	0.861	0.896	0.816	0.854
Gaussian SVM	**0.864**	0.893	0.826	0.858
Logistic Reg.	0.809	0.815	0.798	0.806

TABLE 2-4- CONFUSION MATRIX OF CLASSIFYING SYSTEM COMPONENTS USING GAUSSIAN KERNEL SVM (NUMBER OF SAMPLES AND PERCENTAGE)

		Predicted	
		Normal	**Outage**
Actual	**Normal**	451 (90.2%)	49 (9.8%)
	Outage	87 (17.4%)	413 (82.6%)

2.2.4.b) Leveraging Accuracy-Uncertainty Tradeoff with Multi-Dimensional SVM

In this case study, the area between the support vectors (margin from the decision hyperplane) is considered as an uncertain area. To improve the classification accuracy, the SVM gap is extended by decreasing the penalty coefficient, so the estimated states in the certain area will become more accurate. A k-fold cross validation with $k = 5$ is used to evaluate the performance. Particularly, the dataset is randomly partitioned into five subsamples each containing 120 samples. A single subsample is retained as the validation/test set, and the remaining subsamples are used for training. This process is then repeated five times (i.e., the number of folds).

Table 2-5 shows the performance of SVM and the number of components in uncertain area without considering the component deterioration. Table 2-6 shows the improvement when component deterioration is considered as a feature of the trained model. Comparing the results of the proposed approach with and without considering deterioration level indicates the benefit and the importance of this factor. As it is shown, the F_1-score is improved, in both cases of base and certain, for all considered penalty coefficients. This improvement can be as high as 7.4% which is obtained for the case of $c = 0.1$. In addition, the number of components in uncertain area is reduced in all cases, especially when penalty coefficient c is larger than 1.

TABLE 2-5- PERFORMANCE OF SVM AND THE NUMBER OF COMPONENTS IN UNCERTAIN AREA WITHOUT CONSIDERING DETERIORATION LEVEL OF THE COMPONENT

	Penalty Coefficient (c)				
	100	10	1	0.1	0.01
F_1 -score	84.16	85.66	87.16	83.66	84.33
F_1 -score certain	90.00	91.67	91.81	90.63	94.43
Margin Size	0.111	0.111	0.115	0.121	0.254
No. of uncertain	25	26	28	35	61

TABLE 2-6- PERFORMANCE OF SVM AND THE NUMBER OF COMPONENTS IN UNCERTAIN AREA WITH CONSIDERING DETERIORATION LEVEL OF THE COMPONENT

	Penalty Coefficient (c)				
	100	10	1	0.1	0.01
F_1 -score	89.67	89.50	89.33	90.17	89.67
Improvement (%)	6.55	4.48	2.49	7.78	6.33
F_1 -score certain	95.34	95.52	95.61	95.36	95.37
Improvement (%)	5.93	4.20	4.14	7.43	4.17
Margin Size	0.079	0.082	0.097	0.157	0.300
Change (%)	-28.83	-26.13	-15.65	29.75	18.11
No. of uncertain	17	17	19	30	60

The obtained results advocate that by decreasing the penalty coefficient, the margin of SVM becomes larger and thus more components will be located in the uncertain area. However, the F1-Score of components is significantly improved (from 95.34 in $c = 100$ to 98.37 in $c = 0.01$). The final decision can be considered as a tradeoff between the prediction accuracy and the number of components in the uncertain area.

Figure 2-9 shows the relationship of penalty coefficient (c) and regularization weight (ε) of miss-classified data points inside the margin. By increasing the penalty coefficient, regularization weight decreases. Figure 2-10 illustrates optimal hyperplane in a three-dimensional feature space for the studied case.

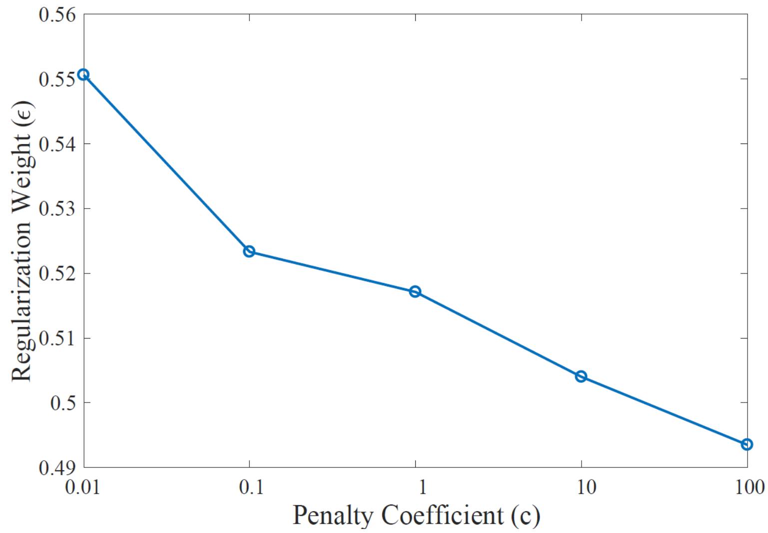


FIGURE 2-9- RELATIONSHIP OF PENALTY COEFFICIENT (C) AND REGULARIZATION WEIGHT OF MISS-CLASSIFIED DATA POINTS INSIDE THE MARGIN.

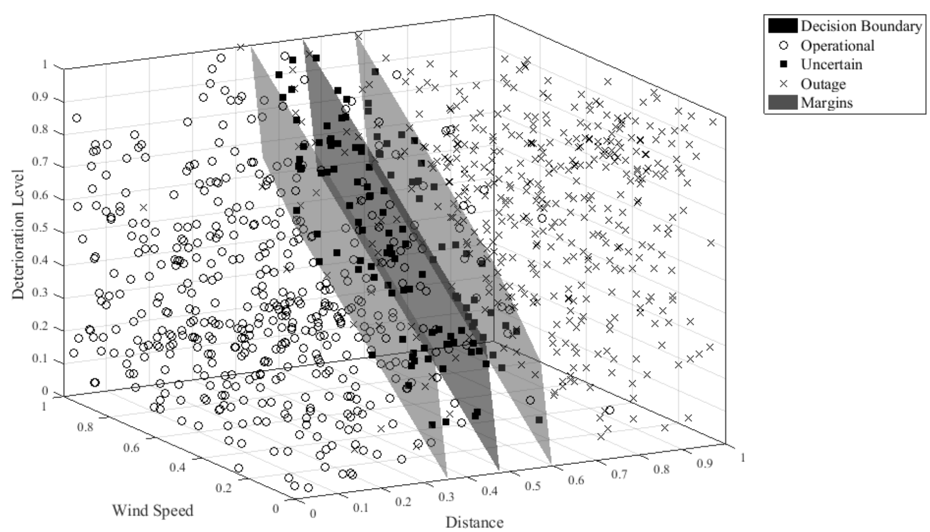


FIGURE 2-10- OPTIMAL HYPERPLANE IN A THREE-DIMENSIONAL FEATURE SPACE USING SVM ON REAL DATA

A third order polynomial logistic regression model is also developed and trained with these three features to predict the component outage and to further show improvement over existing work in Section 2.1.2. The logistic regression model is evaluated in the same fashion (i.e., using cross k-fold validation with k = 5), which offers an overall F1-score of 0.885. Comparing the results of the logistic regression model with the SVM (shown in Table 2-6), it can be clearly seen that the proposed SVM offers a considerably better performance.

CHAPTER THREE: LOAD CURTAILMENT ESTIMATION AND GRID HARDENING

The predicted component outages from previous Chapter are then plugged into a load curtailment minimization model to estimate the nodal load curtailments in the system. The formulation of the proposed load curtailment minimization model is discussed in Section 3.1. The standard IEEE 30-bus system with a combination of hurricane path and intensity scenarios are used to study the model. The results demonstrate that the proposed modelling framework is capable to effectively capture the dynamics of load curtailment estimation in response to extreme events.

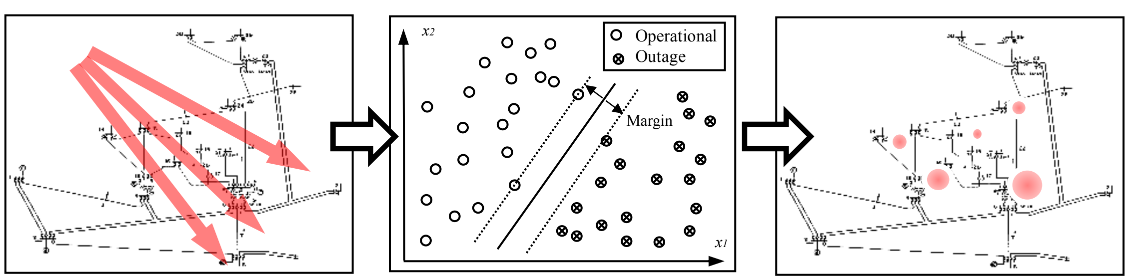
Once the probable damages to system components are estimated, these predictions are fed into a hardening model, which determines strategic locations for placement of distributed generation (DG) units, which is presented in Section 3.2. The numerical simulations on the standard IEEE 118-bus test system illustrate the merits and applicability of the proposed hardening model. The results indicate that the proposed hardening model through decentralized and distributed local energy resources can produce a more robust solution that can protect the system significantly against multiple component outages due to an extreme event.

3.1. Load Curtailment Estimation in Response to Extreme Events

The Load Curtailment Estimation problem is solved in three consecutive stages as illustrated in Figure 3-1. First, the category and the path of an upcoming hurricane are predicted, as shown in Figure 3-1(a). The category and path are used to identify the intensity of the hurricane and the potentially impacted regions, respectively. These data are obtained from weather forecasting agencies. Next, the speed of the hurricane, and the distance of each power grid component from the center of the hurricane— denoted by x_1 and x_2, respectively—are used to predict the state of a component, as shown in Figure 3-1(b). An SVM method is used in this stage to classify the components into two states of damaged (on outage) and operational (in service). The SVM model is trained on historical data. Finally, a minimum load curtailment problem considering the predicted state of each component to estimate the potential nodal load curtailments is solve, as shown in Figure 3-1(c).

3.1.1. Proposed Load Curtailment Estimation Model

The objective of the minimum load curtailment problem is defined as the value-weighted cost of load curtailment in the system, as follows:



(a) Forecasting (b) Component Outage Prediction (c) Load Curtailment Estimation

FIGURE 3-1- THE SCHEMATIC VIEW OF THE PROPOSED LOAD CURTAILMENT ESTIMATION MODEL

$$\min \sum_{t} \sum_{s} \sum_{b} VOLL_b \times LC_{bts} \tag{12}$$

where $VOLL_b$ is the Value of Lost Load at bus b, and LC_{bts} is the amount of load curtailment at bus b at time t during contingency scenarios s. The Value of Lost Load represents the average cost that each customer is willing to pay in order to avoid any load interruptions [13]. Assuming UX_{its} as the outage state of unit i at time t in scenario s (where *operational* state equals to 1 and *outage* state equals to 0) and UY_{lts} as the outage state of line l at time t in scenario s, the proposed objective function is subject to the following physical constraints:

$$\sum_{i \in B_b} P_{its} + \sum_{l \in B_b} PL_{lts} + LC_{bts} = D_{bt} \qquad \forall b, \forall s, \forall t \tag{13}$$

$$P_i^{\min} I_{it} UX_{its} \le P_{its} \le P_i^{\max} I_{it} UX_{its} \qquad \forall i, \forall s, \forall t \tag{14}$$

$$\left| P_{it0} - P_{its} \right| \in \Delta_i \qquad \forall i, \forall s, \forall t \tag{15}$$

$$-PL_l^{\max} UY_{lts} \le PL_{lts} \le PL_l^{\max} UY_{lts} \qquad \forall l, \forall s, \forall t \tag{16}$$

$$\left| PL_{lts} - \frac{\sum_b a_{lb} \theta_{bts}}{x_l} \right| \le M \left(1 - UY_{lts} \right) \qquad \forall l, \forall s, \forall t \tag{17}$$

where b, i, and l are the indices for buses, generation units, and lines, respectively; B_b is the set of components connected to bus b, s is index for scenarios, and t is index for time; $P_i^{\max}$ and $P_i^{\min}$ represent the maximum and minimum generation capacity of unit i, respectively; PL_{lts} is the real power flow of line l at time t in scenario s, θ_{bts} is the phase angle of bus b at time t in scenario s, and M is a large positive constant. The parameter a_{lb} is the element of line l and bus b at line-bus incidence matrix, and D_{bt} is the load at bus b at time t.

The total injected power to each bus from generation units and line flows is equal to the nodal load which can be ensured by load balance equation (13). Load curtailment variable (LC_{bts}) ensures a feasible solution in case of component outages when there is not sufficient generation and/or transmission capacity to supply loads. Generation unit output power is limited to its capacity limit and will be set to zero depending on its commitment and outage states (14). The change in unit generation is further limited by the maximum permissible limit between normal and contingency scenarios (15). Transmission line capacity and power flow constraints are modeled by (16) and (17), respectively, where the outage state variable is effectively incorporated in order to model the line outages in contingency scenarios.

3.1.2. Numerical Simulation

Due to the scarcity of structured historical data at components level from the recent hurricanes, a set of synthetic data is generated to train the SVM model. The data includes 300 samples in *outage* state and 300 samples in the *operational* state. To define the synthetic data, Saffir-Simpson Hurricane Scale [55] is used to generate wind speed features of the synthetic data. These generated scenarios are used in the pre-process stage for training the proposed machine learning model, ensuring relevant outage scenario generation. A subset of data (80%) is sampled for training purpose, and the remaining 20% is held out to validate the model. The output of this model (i.e., the outage state of the power grid components) can be used as an input not only for load curtailment estimation application of this study, but also to enhance the accuracy of the scenarios and reduction of *model risk* in other applications such as those presented in [12], [13].

In this word, in order to find the best kernel and its penalty parameters, a set of linear, polynomial quadratic, and Gaussian kernels with different ranges of penalty parameter (i.e., $c = 0.01, 0.1, 1, 10$) are also examined in training process. Table 3-1 shows the accuracy of SVM with aforementioned combinations of penalty parameters and kernels. As shown, the polynomial kernel SVM with c=1 outperforms other models in terms of classification accuracy. The margin size of the SVM with polynomial kernel is 0.1131, and the average ε (regularization weight) is 0.4558.

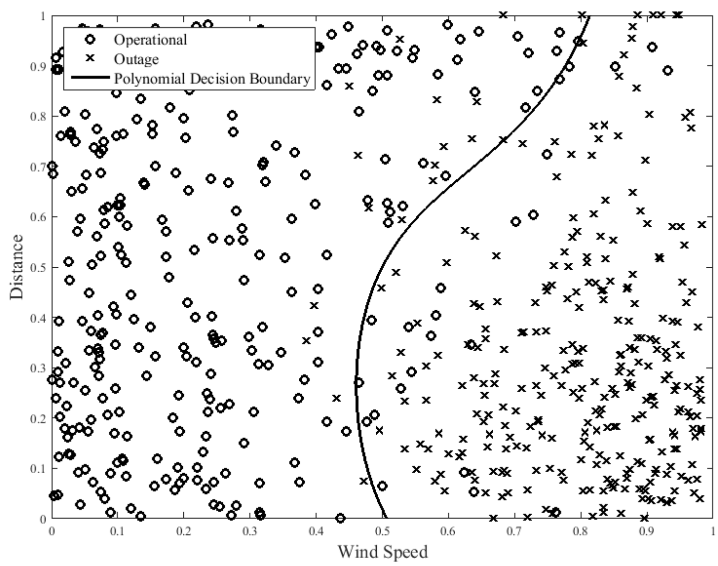


FIGURE 3-2- DECISION BOUNDARY OF THE POLYNOMIAL KERNEL WITH PENALTY PARAMETER C=1

Figure 3-2 shows the decision boundary of the polynomial kernel with penalty parameter c=1, separating outage from operational components based on wind speed and distance from the center of the hurricane. As shown, the instances are not linearly separable, and a nonlinear kernel is necessary to better classify the components. Table 3-2 shows the confusion matrix of this classification. As shown, the proposed method can effectively classify the components into *outage* and *operational* classes.

TABLE 3-1- ACCURACY (%) OF SVM WITH VARIOUS PENALTY-PARAMETERS AND KERNELS

Kernel	*c*=0.1	*c*=1	*c*=10
Linear	91.0	91.4	91.2
Quadratic	91.3	91.2	91.2
Polynomial	92.3	**92.8**	92.7
Gaussian	91.3	91.2	91.8

TABLE 3-2- CONFUSION MATRIX OF CLASSIFYING SYSTEM COMPONENTS

Actual	Predicted	
	Normal	Outage
Normal	91.7%	8.3%
Outage	6.0%	94.0%

The proposed minimum load curtailment model is applied to the standard IEEE 30-bus test system. A hurricane passes through three hypothetical paths with different intensities. Particularly, based on the available hurricane data and the estimated distance from the center of the hurricane, the state of each component in the system is predicted using the trained SVM model. This study estimates how much load curtailment is expected to occur due to an imminent hurricane. Table 3-3 shows the load curtailment of each contingency scenario based on the predicted outages.

TABLE 3-3- LOAD CURTAILMENT OF BUS OUTAGES ALONG THREE HURRICANE PATHS

Bus number	Total Load (MWh)	LC Scenario 1 (MWh)	LC Scenario 2 (MWh)	LC Scenario 3 (MWh)
2	423.08	0	0	4.91
3	46.79	44.95	0	1.62
15	159.87	0	0	0.37
18	62.39	0	59.94	2.10
19	185.22	0	177.95	0
20	42.89	0	41.21	0
23	62.39	0	0	9.92
24	169.62	0	0	162.97
29	46.79	0	0	0.31

As shown, buses 3 and 18 are shown to be the most sensitive buses, since in both Scenarios 2 and 3 these two buses are predicted to be in *outage* state. In addition, buses 18, 19, and 20 are the most critical buses as more than 95% of the total load curtailments are expected to take place in these buses. The predicted outages and load curtailment estimation are of crucial for utilities to effectively mobilize their restoration resources in prior- and post-hurricane phases.

3.2. Machine Learning Assisted Power Grid Hardening

The outline of the proposed grid hardening model is depicted in Figure 3-3. The problem is solved in three consecutive steps. In step 1, an SVM model is trained to classify the components into two states of damaged (on outage) and operational (in service) based on historical data. In step 2, the category and the path of an upcoming hurricane are forecasted which can be obtained from a weather forecasting channel. The category and path are used to identify the intensity of the hurricane and the potentially impacted regions, respectively. The speed of the hurricane and the distance of each power grid component from the center of the hurricane are used to predict the state of each component using the model trained in step 1. These predictions can subsequently help determine a set of suitable hardening candidates. Step 3 solves a grid hardening problem to ensure a secure supply of loads in response to the forecasted extreme event based on the predicted state of the components from step 2 and through strategic placement of utility-owned DGs. The proposed hardening model takes grid resilience and economic needs into consideration with the objective of minimizing the total system upgrade cost as well as system operation costs, subject to prevailing investment and operation constraints.

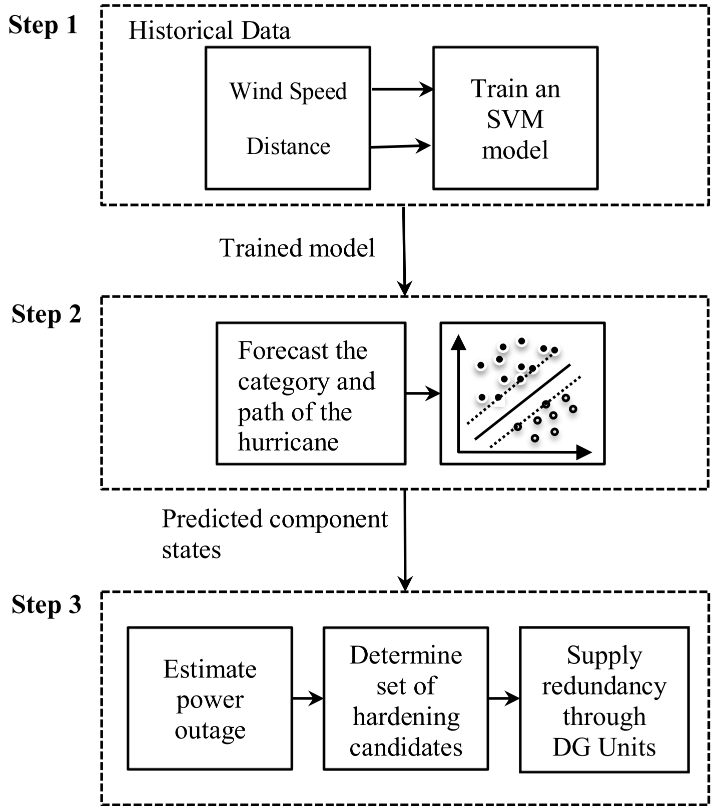


FIGURE 3-3- PROPOSED GRID HARDENING MODEL

This work focuses on physical hardening options, as resilience events are mainly triggered by outages and displacements of physical power grid facilities. Supply redundancy is considered as a valuable hardening approach. Supply redundancy decentralizes the electricity generation, thus instead of relying on large-scale power plants and bulk transmission network for power supply and delivery, a localized supply of power is utilized in certain regions to improve resilience. In this case, if power transfer and delivery from centralized generation is interrupted, a local supply of loads will be provided via available DGs.

3.2.1. Proposed Grid Hardening Model

The proposed grid hardening model minimizes the total investment cost of the grid hardening candidates as well as system operation costs, subject to prevailing investment and operation constraints. For reliability studies in power systems, it is common to use the *N*-1 criterion. The *N*-1 criterion simply states that the system needs to adequately and reliably supply loads in case of a single component outage at any given time. However, after an extreme event, it is anticipated that more than one component is affected and becomes unavailable. Hence, different contingency scenarios are considered in neighboring locations along the hurricane path in which more than one component can be in outage state. Assuming s is the contingency scenario, the problem objective is defined as:

$$\min \sum_{t}\sum_{i} F_i\left(P_{it0}, I_{it}\right) + \sum_{t}\sum_{s}\sum_{b} v_b LC_{bts} + \sum_{b} IC_b P_b^{G,\max} \tag{18}$$

where $F_i(.)$ is the operation cost of unit i in normal operation, v is the value of lost load, LC_{bts} is the amount load curtailment, and IC_b is the investment cost associated with system upgrades by a DG unit with the capacity of $P_b^{G,max}$ at bus b. The value of lost load, v, is defined as the average cost that each type of customer, i.e., residential, commercial, or industrial, is willing to pay in order to avoid load interruptions [54]. Assuming UX_{its} as the operation state of unit i at time t in scenario s (1 when operating and 0 when on outage), and UY_{lts} as the operation state of line l at time t in scenario s (1 when operating and 0 when on outage), the following operational constraints are defined:

$$\sum_{i \in B} P_{its} + \sum_{b \in B} P_{bts}^{G} + \sum_{l \in B} PL_{lts} + LC_{bts} = D_{bt} \qquad \forall b, \forall s, \forall t \tag{19}$$

$$P_i^{\min} I_{it} UX_{its} \le P_{its} \le P_i^{\max} I_{it} UX_{its} \qquad \forall i, \forall s, \forall t \qquad (20)$$

$$\left| P_{it0} - P_{its} \right| \in \Delta_i \qquad \forall i, \forall s, \forall t \qquad (21)$$

$$\left| PL_{lts} - \frac{\sum_b a_{lb} \theta_{bts}}{x_l} \right| \le M\left(1 - UY_{lts}\right) \qquad \forall l, \forall s, \forall t \qquad (22)$$

$$-PL_l^{\max} UY_{lts} \le PL_{lts} \le PL_l^{\max} UY_{lts} \qquad \forall l, \forall s, \forall t \qquad (23)$$

$$0 \le P_b^G \le P_b^{G,\max} \qquad \forall b, \forall s, \forall t \qquad (24)$$

$$\sum_b IC_b P_b^{G,\max} \le \Delta \qquad \forall b \qquad (25)$$

Constraint (19) represents nodal load balance. The load balance ensures that the total injected power to each bus from generation units, supply redundancies through DGs, and line flows is equal to the total consumed load at that bus. The load curtailment variable, *LC*, is added to the load balance equation to ensure a feasible solution when there is not sufficient generation to supply loads (due to component outages). Load curtailment is zero under normal operation conditions. Generation unit output power is limited by its capacity limit and is set to zero depending on its commitment and operation states (20). The change in a unit generation is further limited by the maximum permissible limit between normal and contingency scenarios (21). Transmission line capacity limits and power flow constraints are modeled by (22) and (23), respectively, in which the operation state is included to effectively model the line outages in contingency scenarios. P^G_{bts} is the DG output power which is limited by its capacity limit and is set to zero depending on supply redundancy decision at bus *b* (24). Furthermore, the sum of the investment cost of all installed DGs in the system cannot exceed the available budget set by the system planner (25).

3.2.2. Numerical Simulation

The proposed hardening model is applied to the standard IEEE 118-bus test system. A hurricane is assumed to pass through three hypothetical paths as shown in Figure 3-4. The components in each path and its neighboring areas are classified into two categories of operational and outage according to the wind speed and the distance to the center of the hurricane, using the SVM model trained in the previous section. The trained model classified 48, 56, and 55 components as outage in paths 1, 2 and 3, respectively.

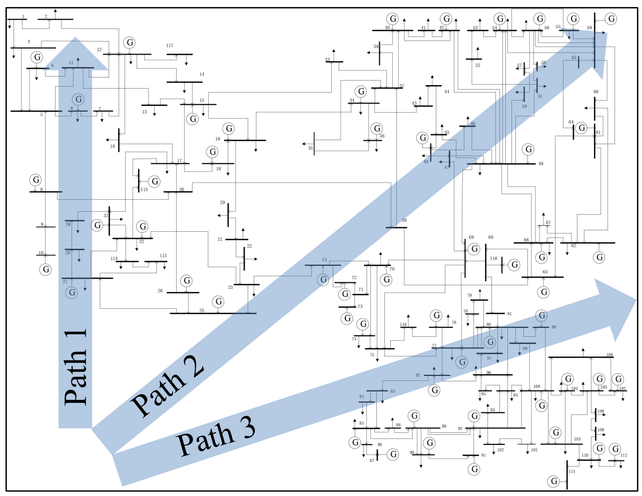


FIGURE 3-4- IEEE 118-BUS TEST SYSTEM AND THE FORECASTED HURRICANE PASSING THROUGH THREE HYPOTHETICAL PATHS

The proposed hardening model and the optimal scheduling problem is solved for one year (8760 hours). The value of lost load is considered \$100/MWh at all buses. The investment cost associated with installing a DG unit (supply redundancy) at any given bus is assumed to be \$50/MW. The following cases are studied:

Case 1: In this case, power grid scheduling is performed without hardening (supply redundancy). The optimal operation cost is obtained as \$366,277,300. A total of 43338, 47143, and 44393 MWh load curtailment occurs in paths 1, 2, and 3, respectively. The average cost of unserved energy is calculated as $449,580,000.

Case 2: In this case, power grid scheduling is solved using the proposed hardening model. It is assumed that there is no constraint on investment budget. The annual optimal operation cost is obtained as $492,307,700. No load curtailment has occurred in this case, so the cost of unserved energy is zero and the system is secure against considered component outage scenarios. The proposed model advocates on hardening options at buses 33, 37, 39, 41, 42, 54, 59, and 80 to avoid load curtailments.

Case 3: This case discusses the effect of system hardening investment budget on the solution when all other parameters are kept unchanged. The results are summarized in Table 3-4. As shown, the average unserved energy decreases by increasing the amount of budget.

TABLE 3-4- EFFECT OF INVESTMENT BUDGET ON OPERATION COST AND LOAD CURTAILMENT

Budget	Load Curtailment (MWh)			Average Unserved Energy Cost
	Path 1	Path 2	Path3	
$0M	43,338	47,143	44,393	$449,580,000
$1M	-	22,341	3155	$84,986,666
$10M	-	20,138	2,751	$76,296,666
$100M	-	5294	-	$17,646,666
$126M	-	-	-	$0

As Table 3-4 suggests the relationship between the investment budget and average unserved energy cost reduction is not linear. For instance, the unserved energy cost reduced drastically ($364,593,334) with $1M investment, but to zero out the unserved energy cost (from $84,986,666 to zero), the system requires $125 M additional budget. The final decision is a trade-off between hardening budget and load curtailment reduction based on planner's discretion.

CHAPTER FOUR: PROBABILISTIC LOAD CURTAILMENT ESTIMATION USING POSTERIOR PROBABILITY MODEL AND TWIN SUPPORT VECTOR MACHINE

In this chapter, a TWSVM classification method is trained to find the operational state of each component by considering the path and the intensity of the hurricane, as well as the distance of each component from the center of the hurricane. A posterior probability model is consequently applied to the output of the TWSVM model to estimate the outage probability of each component. Having an accurate estimation of probable outages plays a vital role in responding to an upcoming hurricane.

Unlike the existing work on outage prediction and extended outage consideration in security-constrained unit commitment (SCUC), including the previous work of authors in [56] [41], this chapter considers the probability of outage obtained by a machine learning approach in scheduling. TWSVM is chosen for its performance in complex intertwined classification problems and when dealing with imbalanced datasets. This can be potentially problematic since the data of past hurricanes are imbalanced, i.e., the number of non-operational components is far less than the number of operational components. The merit behind proposed probabilistic load curtailment estimation model is that it considers all contingency scenarios with their probability and hence the most probable scenario or the scenarios with most load curtailment can be recognized. The predicted outage and estimated outage probability can be useful for electric utilities to assess their risk and allocate necessary resources and repair crews to prepare for and recover from hurricanes in a considerably shorter time-frame.

The rest of the chapter is organized as follows: Section 2 presents the model outline and formulation of the proposed machine learning method for outage prediction. Section 3 presents simulation results on a test system, and Section 4 concludes the chapter.

4.1. Proposed model

The goal of this section is to determine the probable load curtailments in a power system as a result of hurricane-caused component outages. The considered components include, but are not limited to, transmission lines, generation units, and substations. The problem is solved in three consecutive steps. In Step 1, a TWSVM model [57][58] is trained on historical outage data to help classify the operational state of components after the hurricane.

The speed of the hurricane and the distance of each component from the center of the hurricane are used to predict the probability of outage for each component. The output of the TWSVM model will be a list of 0/1 values, showing whether each component is operational or on outage, however it provides no information on the outage probability. To estimate the outage probability for each component, a posterior probability sigmoid model [59] is applied in the Step 2 to the output of the first step. The category and the path of the upcoming hurricane in this step are obtained from weather forecasting agencies. In Step 3, the obtained component outages and their associated probabilities are integrated into a probabilistic load curtailment estimation model to estimate the nodal load curtailments and thus help identify the areas that will potentially be impacted by the hurricane.

4.1.1. TWSVM

The SVM method has numerous advantages including the ability to provide a global solution for data classification. It generates a unique global hyperplane by solving a quadratic programming problem (QPP) to separate the data samples of different classes rather than local boundaries as compared to other existing data classification approaches. Due to its performance, SVM is one of the most widely-used classification techniques in data mining. One of the main challenges with the traditional SVM, however, is that it solves only one QPP problem to classify the data, which may not be suitable in cases of imbalanced data.

Although SVM often produces effective solutions for balanced datasets, it is sensitive to imbalance in datasets and produces suboptimal results [60]. In other words, the separating hyperplane of an SVM model trained with an imbalanced dataset can be skewed towards the minority class [61], and hence the performance of that model is degraded with respect to the minority class. Several approaches in literature have been proposed to improve the SVM performance when dealing with imbalanced dataset classification [60]. These approaches can be categorized as data processing approaches (such as resampling methods [62] and ensemble learning methods [63]), algorithmic approaches (such as different error cost [61] or z-SVM [64]), and hybrid approaches (such as hybrid kernel machine ensemble [65]). Despite the performance improvement of these approaches, the suboptimality of the soft-margin is an inherited problem of SVM and majority of these approaches require an expert understanding of data shape and empirical parameter tuning, e.g., setting a proper weight for each class, or finding best ensemble size.

A viable alternative to SVM is TWSVM, as a machine learning approach suitable for complex intertwined classification problems, which classifies the patterns of two classes by using two non-parallel hyperplanes [66]. The biggest advantage of TWSVM, in addition to the training speed, is its ability to handle imbalanced datasets [57]. This is because each class has its own representative hyperplane instead of one hyperplane separating two classes from each other, and therefore TWSVM can classify underrepresented classes better than traditional SVM, especially when the classes are intertwined. Since TWSVM classifies the data using two hyperplanes, it solves a pair of QPPs instead of a single complex QPP as in traditional SVM. Comparing to a traditional SVM over benchmark datasets, TWSVM has shown comparable performance while being approximately four times faster [57][58]. TWSVM has shown improvement in several practical applications such as classification of biomedical data [67], gesture classification [68] speaker recognition (i.e., personal identity from the speech signal) [69], and image analysis [70], to name a few. Figure 4-1 illustrates a traditional linear classifier SVM and TWSVM in separating two classes. As shown, traditional SVM does not take the data skewness into account and the separating hyperplane is the one that represents the largest margin between two classes.

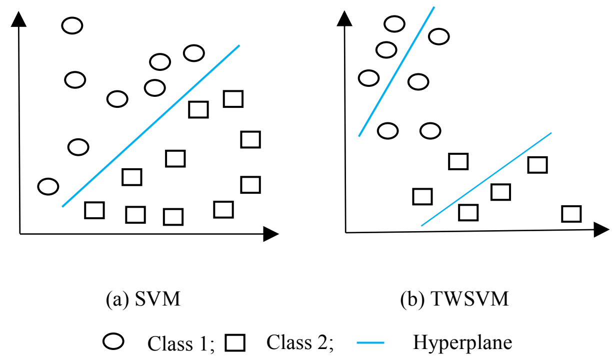


FIGURE 5-1- SVM AND TWSVM FOR IMBALANCED DATASET IN TWO-DIMENSIONAL FEATURE SPACE.

The goal of TWSVM in a binary classification problem is to construct two non-parallel planes for each class such that each hyperplane is closer to the data samples of its representative class while distant from the samples of the other class [66]. The distances between the samples and both non-parallel hyperplanes are compared to determine the category of each sample.

Consider a binary classification problem that classifies m_1 training samples belonging to positive class and m_2 training samples belonging to negative class in an n-dimensional real space $\mathbf{R}^n$, where $m_1+m_2=m$. Let matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ represent the training samples of the positive and negative classes respectively. Since a linear TWSVM seeks two non-parallel hyperplanes, two hyperplanes $\boldsymbol{h}_1(\boldsymbol{x})$ and $\boldsymbol{h}_2(\boldsymbol{x})$ are defined as:

$$\boldsymbol{h}_i\left(\boldsymbol{x}\right)=\boldsymbol{w}_i^{\mathrm{T}}\boldsymbol{x}+\boldsymbol{d}_i=\mathbf{0} \qquad \forall i\in\{1,2\} \qquad (1)$$

where $\boldsymbol{w}_i$ is the normal vector to the hyperplane representing training examples of class i; and $\boldsymbol{d}_i$ is the bias vector of the separating hyperplanes representing class i. $|\boldsymbol{d}_i|/||\boldsymbol{w}_i||$ is the perpendicular distance from the hyperplanes to the origin. To find hyperplanes $\boldsymbol{h}_1(\boldsymbol{x})$ and $\boldsymbol{h}_2(\boldsymbol{x})$, such that $\boldsymbol{h}_1(\boldsymbol{x})$ is closest to the training samples of the positive class and far from the training samples of the negative class, and $\boldsymbol{h}_2(\boldsymbol{x})$ is closest to the training samples of the negative class and far from the training samples of the positive class, the following QPP is solved for each class:

$$\min_{\boldsymbol{w}_i, \boldsymbol{d}_i, \xi_i} \left(\frac{1}{2} \left\| \boldsymbol{A}_i \boldsymbol{w}_i + \boldsymbol{e}_i \boldsymbol{d}_i \right\|^2 + c_i \boldsymbol{e}_j^{\mathrm{T}} \xi_i \right) \tag{2}$$

s.t.

$$-\rho_i \left(\boldsymbol{A}_j \boldsymbol{w}_i + \boldsymbol{e}_j \boldsymbol{d}_i \right) + \xi_i \geq \boldsymbol{e}_j \qquad \xi_i \geq 0, i \neq j \tag{3}$$

where c_i>0 is the regularization term to control overfitting of class i; $\boldsymbol{e}_i$ is a vector of ones of appropriate dimension; $\|.\|^2$ denotes Euclidean distance; ξ_i is slack variable of class i; and ρ_i is the coefficient of each class where ρ_1=1 for the positive class and ρ_2=−1 for the negative class. TWSVM solves two QPPs problem (2) and (3) separately for each class. If sample sizes of both classes are approximately equal to $m/2$, the complexity of solving these two QPPs in TWSVM will be $O(2\times(m/2)^3)$. Comparing with the standard SVM with computational complexity of $O(m^3)$ which solves one QPP problem for both classes at the same time, TWSVM is approximately four times faster [66]. The objective function seeks the distance from the sample to the hyperplane by the square distances (L_2-norm), and minimizes the distance to ensure the hyperplane is as close as possible to the samples of its own class. The sample x is assigned to class i if:

$$\frac{\left| \boldsymbol{w}_i^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{d}_i \right|}{\left| \boldsymbol{w}_j^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{d}_j \right|} + \frac{\left\| \boldsymbol{w}_j \right\|}{\left\| \boldsymbol{w}_i \right\|} < 1 \qquad \forall i \in \{1,2\}, \forall j \in \{1,2\}, i \neq j \tag{4}$$

where $\|\boldsymbol{w}_i\|$ is the Euclidean length of vector $\boldsymbol{w}_i$.

Similar to SVM, kernel method [49] can be applied to TWSVM. The idea of a kernel method (or as sometimes called kernel trick) is to map the input feature vector into a higher-dimension space where the classes are linearly separable. To apply kernel to TWSVM, the QPP problem of (2) and (3) is formulated as:

$$\min_{\boldsymbol{w}_i,\boldsymbol{d}_i,\xi_i} \frac{1}{2}\left\| K(\boldsymbol{A}_i,\boldsymbol{B}^{\mathrm{T}})\boldsymbol{w}_i + \boldsymbol{e}_i\boldsymbol{d}_i \right\|^2 + c_i\boldsymbol{e}_j^{\mathrm{T}}\xi_i \tag{5}$$

s.t.

$$-\rho_i\left(K(\boldsymbol{A}_j,\boldsymbol{B}^{\mathrm{T}})\boldsymbol{w}_i + \boldsymbol{e}_j\boldsymbol{d}_i\right) + \xi_i \geq \boldsymbol{e}_j \qquad \xi_i \geq 0,\, i \neq j \tag{6}$$

where $B=[\boldsymbol{A}_1, \boldsymbol{A}_2]^{\mathrm{T}}$ and K is the kernel function. Finding a proper value of penalty parameter c and the best kernel depends on the shape of classes, which are often found via a search method to minimize the error on the test set.

4.1.2. Posterior probability estimation

To determine the likelihood of a sample belonging to a specific class, two normalized distances, to each hyperplane $\boldsymbol{h}_i$, are defined as:

$$D_i(\boldsymbol{x}) = \frac{\left|\boldsymbol{w}_i^T\boldsymbol{x} + \boldsymbol{d}_i\right|}{\left\|\boldsymbol{w}_i\right\|} \qquad \forall i \in \{1,2\} \tag{7}$$

Given the distance between two representative hyperplanes h_1 and h_2, two new relative distances can be defined as:

$$D_+(\boldsymbol{x}) = D_1(\boldsymbol{x}) + D_2(\boldsymbol{x}) = 0 \tag{8}$$

$$D_-(\boldsymbol{x}) = D_1(\boldsymbol{x}) - D_2(\boldsymbol{x}) = 0 \tag{9}$$

Intuitively, the probability of a sample x belonging to a certain class depends on its relative distance to the positive class D_+ and the negative class D_-. Two relevant quantities $D_{\min}(\boldsymbol{x})$ and $D_{\max}(\boldsymbol{x})$ are then defined by:

$$D_{\min}(\boldsymbol{x}) = \min\left\{D_+(\boldsymbol{x}), D_-(\boldsymbol{x})\right\} \tag{10}$$

$$D_{\max}(\boldsymbol{x}) = \max\{D_{+}(\boldsymbol{x}), D_{-}(\boldsymbol{x})\} \qquad (11)$$

Figure 4-2 shows a sample x and its corresponding relative distances $D_{+}(\boldsymbol{x})$ and $D_{-}(\boldsymbol{x})$.

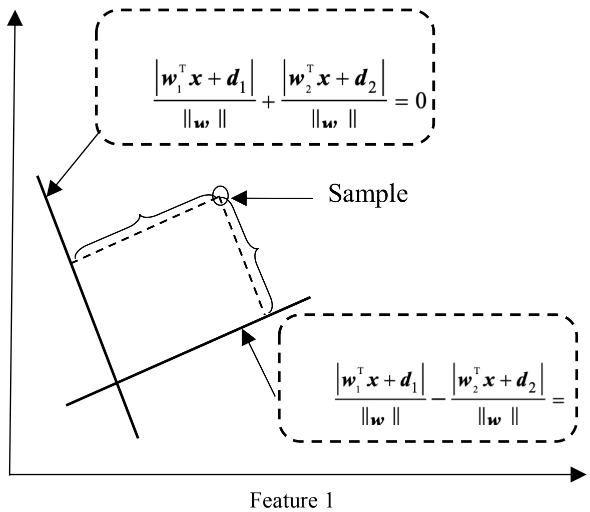


FIGURE 4-2- AN EXAMPLE INDICATING MEANING OF RELATIVE DISTANCES OF SAMPLE X TO THE POSITIVE AND NEGATIVE SEPARATING HYPERPLANES IN A TWO-DIMENSIONAL FEATURE SPACE.

As it is shown, the quantities $D_{\min}(\boldsymbol{x})$ and $D_{\max}(\boldsymbol{x})$ are the factors influencing the probability of belonging to the positive class. In other words, the probability of belonging to the positive class increases when either $D_{\min}(\boldsymbol{x})$ or $D_{\min}(\boldsymbol{x})/D_{\max}(\boldsymbol{x})$ becomes larger. Hence, a score function $f(\boldsymbol{x})$ can be define as:

$$f(\boldsymbol{x}) = \begin{cases} D_{\min}(\boldsymbol{x})\left(\dfrac{D_{\min}(\boldsymbol{x})}{D_{\max}(\boldsymbol{x})}\right)^{\lambda} & D_1(\boldsymbol{x}) > D_2(\boldsymbol{x}) \\ 0 & D_1(\boldsymbol{x}) = D_2(\boldsymbol{x}) \\ -D_{\min}(\boldsymbol{x})\left(\dfrac{D_{\min}(\boldsymbol{x})}{D_{\max}(\boldsymbol{x})}\right)^{\lambda} & D_1(\boldsymbol{x}) < D_2(\boldsymbol{x}) \end{cases} \qquad (12)$$

If $D_1 > D_2$, then the sample belongs to the positive class, otherwise to the negative class. If $D_{\min}$ is small and $D_{\max}$ is large, it means that the sample is very close to one of the planes and far away from the other. Hence, the probability is large, i.e., $f(\boldsymbol{x})$ becomes a very large positive number for the positive class and a very large negative number for the negative class. If $D_{\min} \approx D_{\max}$, then it means the sample is relatively in the same distance between these classes and the $f(\boldsymbol{x})$ is small. Constant λ is the weight parameter. This parameter can be determined on a validation set. The data is split into three subsets, training, validation and test. The training set is used to find separating hyperplanes. Then different values of λ in the score functions $f(\boldsymbol{x})$ will be evaluated on the validation set and the best parameter will be tested on the test subset.

The above formulation can be easily extended to nonlinear TWSVM by considering the kernel-generated surfaces instead of the hyperplanes as:

$$D_i(\boldsymbol{x}) = \frac{\left|\boldsymbol{w}_i^{\mathrm{T}} K(\boldsymbol{x}, \boldsymbol{B}^{\mathrm{T}}) + \boldsymbol{d}_i\right|}{\left\|\boldsymbol{w}_i^{\mathrm{T}} K(\boldsymbol{B}, \boldsymbol{B}^{\mathrm{T}}) \boldsymbol{w}_i^{\mathrm{T}}\right\|} \qquad \forall i \in \{1, 2\} \tag{13}$$

Since $D_{\min}$ and $D_{\max}$ can be any arbitrary value, the range of the score function $f(\boldsymbol{x})$ is $(-\infty, +\infty)$. Platt scaling or Platt calibration is a way of transforming the score of a classification model into a probability distribution over classes [71]. Platt scaling finds the parameters of a sigmoid function which converts the scoring output of $(-\infty, +\infty)$ to a probability of [0, 1]. It has been shown that Platt method yields probability estimates that are at least as accurate as ones obtained by training a SVM, while being expedient [72]. Similar to the continuous output in an SVM, the following posterior probability function is constructed over the values of score function $f(\boldsymbol{x})$ as:

$$P(y = +1 \mid f(\boldsymbol{x})) = \frac{1}{1 + \mathrm{e}^{\alpha f(\boldsymbol{x}) + \beta}} \tag{14}$$

$$P(y = -1 \mid f(\boldsymbol{x})) = 1 - P\left(y = +1 \mid f(\boldsymbol{x})\right) \tag{15}$$

where α and $ß$ are the scaling weights of the sigmoid function calculated using the maximum likelihood estimation (i.e., Platt scaling) [71], by minimizing the following function:

$$\min_{\alpha,\beta}\left\{-\sum_{k=1}^{m}\left[t_k \lg p_k + \left(1-t_k\right)\lg\left(1-p_k\right)\right]\right\} \tag{16}$$

s.t.

$$p_k = \Pr\left(y_k = +1 \mid f(\boldsymbol{x}_k)\right) \tag{17}$$

$$t_k = \begin{cases} \dfrac{m_1+1}{m_1+2} & y_k = +1 \\ \dfrac{1}{m_2+2} & y_k = -1, k = 1,2,\ldots,m \end{cases} \tag{18}$$

where t_k is the target probability of a particular sample of $\boldsymbol{x}_k$; p_k is the predicted probability of that sample; and m, m_1 and m_2 are the numbers of total training samples, positive training samples and negative training samples, respectively.

4.1.3. Evaluation criteria

1) Evaluation of classifier

To evaluate the performance of the classifier, a cross-fold validation is used. The cross-fold validation splits the data into q subsets, in which the classifier is trained on q-1 subsets and evaluated on the subset that is left in the training. This process is performed q times (such that the classifier is evaluated on all samples). The final classification accuracy is the average of classification accuracies on all folds. Reporting the general accuracy of prediction cannot be sufficient as the number of samples may not be balanced in the test set. The F_1-score is a common and reliable measure of classification performance [21] defined as:

$$F_1 = \frac{2PR}{P+R} \tag{19}$$

where P (precision) is the number of correct positive results divided by the number of all positive results returned by the classifier; and R (recall) is the number of correct positive results divided by the number of all relevant samples. In case of outage estimation, P is defined as the ratio of number of correctly predicted outages to total number of predicted outages, and R is defined as the ratio of number of correctly predicted outages to total number of actual outages.

A higher value of the F_1-score, which is a number between 0 and 1, indicates a better classification and justifies the viable performance of the existing decision boundary.

2) Evaluation of posterior probability estimation

A common way to determine how well a posterior probability estimator model fits the data is the area under receiver operating characteristic (ROC) curve [21]. A ROC curve is a graph showing the performance of a classification model at all classification thresholds. The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. In this chapter, since the goal is to estimate outage probability, the outage state is considered as positive and the operational state is considered as negative class. The TPR is the number of correctly predicted samples in outage state divided by the total number of samples in outage state, and FPR is the number of incorrectly predicted samples in operational state divided by the total number of samples in operational state.

The area under the ROC curve (AU-ROC) measures the entire two-dimensional area underneath the entire ROC curve as:

$$A_{\text{AU-ROC}} = \int_{-\infty}^{+\infty} TPR(\tau)FPR(\tau)\mathrm{d}\tau \tag{20}$$

where τ is a threshold indicating that an instance is classified as positive class if the posterior probability is greater than τ, and negative otherwise. AU-ROC provides an aggregate measure of performance across all possible classification thresholds. It is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one [21].

4.1.4. Probabilistic load curtailment estimation

The objective function of the probabilistic load curtailment estimation problem is defined as:

$$\min \sum_{t}\sum_{g} F_g\left(P_{gt0}, I_{gt}\right) + \sum_{t}\sum_{s}\sum_{b} \pi_s v_b L_{C,bts} \tag{21}$$

where π_s is the probability of each hurricane scenario where $\sum\pi_s$=1; $F_g(.)$ is the operation cost function, which includes the generation cost and startup/shutdown costs, P_{gt0} is the real power generation of unit g at time t in scenario zero (i.e., normal operation), I_{gt} is the commitment state of unit g at time t, v is the value of lost load, and $L_{C,bts}$ is the amount of nodal load curtailment at bus b at time t in scenario s. The value of lost load is defined as the average cost that each type of customer, i.e., residential, commercial, or industrial, is willing to pay in order to avoid power supply interruptions. Assuming U_X and U_Y as outage states for generation units and transmission lines, respectively, the proposed objective function is subject to the following operational constraints:

$$\sum_{g\in B} P_{gts} + \sum_{l\in B} P_{L,lts} + L_{C,bts} = D_{bt} \qquad \forall b, \forall t, \forall s \tag{22}$$

$$P_g^{\min} I_{gt} U_{X,gts} \le P_{gts} \le P_g^{\max} I_{gt} U_{X,gts} \qquad \forall g, \forall t, \forall s \tag{23}$$

$$P_{gts} - P_{g(t-1)s} \le U_{R,g} \qquad \forall g, \forall t, \forall s \qquad (24)$$

$$P_{g(t-1)s} - P_{gts} \le D_{R,g} \qquad \forall g, \forall t, \forall s \qquad (25)$$

$$T_{gt}^{\text{on}} \ge U_{T,g} \left(I_{gt} - I_{g(t-1)} \right) \qquad \forall g, \forall t \qquad (26)$$

$$T_{gt}^{\text{off}} \ge D_{T,g} \left(I_{g(t-1)} - I_{gt} \right) \qquad \forall g, \forall t \qquad (27)$$

$$0 \le L_{C,bts} \le D_{bt} \qquad \forall b, \forall t, \forall s \qquad (28)$$

$$-P_{L,l}^{\max} U_{Y,lts} \le P_{L,lts} \le P_{L,l}^{\max} U_{Y,lts} \qquad \forall l, \forall t, \forall s \qquad (29)$$

$$\left| P_{L,lts} - \frac{\sum_b a_{lb} \theta_{bts}}{X_l} \right| \le M \left(1 - U_{Y,lts} \right) \qquad \forall l, \forall t, \forall s \qquad (30)$$

where P_{gts} is the real power generation of unit g at time t in scenario s, $P_{L,lts}$ is the real power flow of line l at time t in scenario s, D_{bt} is the load at bus b at time t, $P_g^{\min}$ and $P_g^{\max}$ are respectively the minimum and maximum generation capacity of unit g, $U_{R,g}$ and $D_{R,g}$ are respectively ramp up and ramp down rates of unit g, T_{gt}^{on} and T_{gt}^{off} are respectively the number of successive ON and OFF hours of unit g at time t, $U_{T,g}$ and $D_{T,g}$ are respectively the minimum up time and down time of unit g, $P_{L,l}^{\max}$ is the maximum power flow of line l, a_{lb} is the element of line l and bus b in line-bus incidence matrix, θ_{bts} is the phase angle of bus b at time t in scenario s, X_l is the reactance of line l, and M is a large positive constant.

Load balance equation (22) ensures that the total injected power to each bus from generation units and line flows is equal to the total load at that bus. Load curtailment variable ($L_{C,bts}$) is further added to the load balance equation to ensure a feasible solution when there is not sufficient generation to supply loads (due to component outages). Generation unit output power is limited by its capacity limit and will be set to zero depending on its commitment and outage states (23). Generation units are further subject to prevailing technical constraints including ramp up and down rate limits (24), (25), and minimum up and down time limits (26), (27). The load curtailment at each bus is constrained by the total load on that bus (28). Transmission line capacity limits and power flow constraints are modeled by (29) and (30), respectively, in which the outage state is included to model the line outages in contingency scenarios. Note that (21)-(30) is effectively a SCUC problem with weighted scenarios and simultaneous component outages.

4.2. Numerical simulations

The standard IEEE 118-bus test system is used for testing the proposed model, by assuming that a hurricane is predicted to pass through the system. The system characteristics, including generation, line, and load data, can be found in [73].

4.2.1. TWSVM performance

As historical data for the past hurricanes at component level are limited, 550 samples are synthetically generated (500 samples of component in operational state and 50 samples in outage state) following a normal distribution function with a small Gaussian noise. To ensure that these samples fit a practical situation, the models proposed in [56] are used for hurricane modeling and the models in [74] are used for identifying the response of each component to the modeled hurricanes. The features are normalized to [0, 1] range based on the maximum considered values of wind speed and distance. These samples are shown in Figure 4-3.

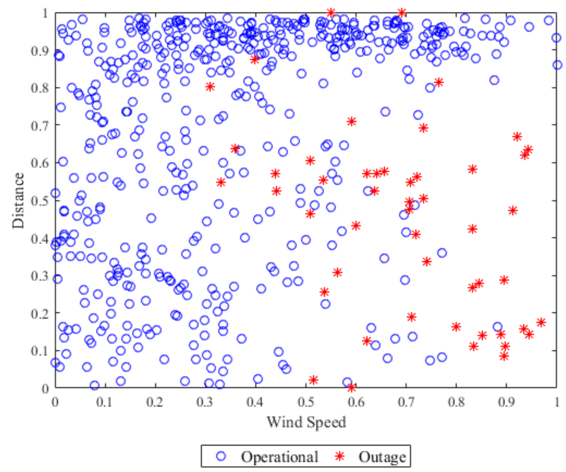


FIGURE 4-3- GENERATED SAMPLES FOR EACH CLASS (OPERATIONAL AND OUTAGE)

Although several other features can be defined, when the dimension increases, typically a significant amount of training data is required to ensure that the samples cover all combinations of feature values. As gathering component level data is not trivial, a limited number of samples is synthesized in the studied dataset and only the two most important/salient features (i.e., wind speed and distance) are used in the outage estimation problem.

To measure the performance of the proposed method, a series of penalty parameters (c=0.01, 0.1, 1, 10, 100) with various common kernels are examined. In each setting. A weighted soft-margin SVM [61] (wSVM) is used to compare the performance. The wSVM adjusts the class sensitivity (penalty of missclassifying) of each class inversely proportional to the frequencies of the class in the training set. In other words, the penalty of missclassifying outage samples are 0.91 (50/550) and the penalty of missclassifying of operational samples are 0.09 (500/550). Table 4-1 shows the average F_1-score of both wSVM and TWSVM over a 5-fold cross validation. On average, TWSVM took 0.0148 seconds to solve the problem and SVM took 0.0320 seconds to find proper separating hyperplane over 5-fold cross validation.

As it is shown, TWSVM with quadratic kernel and c=1 offers the best performance among other settings with the average overall precision of 0.932, recall of 0.912 and F_1-score of 0.922. The relatively small variance (about 3%) in the F_1-score of the SVM and TWSVM under various hyper-parameters indicates that both methods are insensitive to hyper-parameters and are not over-fitted to the training data in the studied case. A third order polynomial logistic regression model is also trained and examined in the same fashion (i.e., 5-fold cross validation) to predict the component outages. The logistic regression model has an F_1-score of 0.856 on the test set which advocates on the superior performance of both SVM and TWSVM in solving this problem.

TABLE 4-1- F1-SCORE OF CLASSIFYING SYSTEM COMPONENTS INTO TWO CLASSES OF OUTAGE AND OPERATIONAL WITH VARIOUS KERNELS AND PENALTY PARAMETERS

c	Linear Kernel		Quadratic Kernel		Gaussian Kernel	
	wSVM	TWSVM	wSVM	TWSVM	wSVM	TWSVM
0.01	0.871	0.891	0.862	0.892	0.851	0.881
0.1	0.871	0.899	0.871	0.901	0.852	0.891
1	0.879	0.915	0.88	0.922	0.851	0.891
10	0.881	0.904	0.869	0.912	0.842	0.880
100	0.879	0.902	0.869	0.899	0.844	0.872

4.2.2. Evaluating posterior probability estimation

To determine the likelihood of a sample belonging to each class, a sigmoid posterior probability function is constructed over the values of score function (12) of the trained model with quadratic kernel and penalty parameter c=1. The scaling weights of sigmoid function are calculated as α=-25.93 and β=2.12 by solving (16). The trained model probability weight λ=0.5 has overall AU-ROC of 0.89 on the test subset. Other weight parameters (λ=0, 0.5, 1.0, 1.5) are further tested on the validation set, however λ =0.5 produces the best result in terms of AU-ROC. Figure 4-4 demonstrates posterior probability for different weight parameters. As shown, by increasing λ the posterior probability function becomes smoother and the classes become less distinguishable. A small value of weight parameter, e.g., λ=0, makes the probabilistic model very sharp where probabilities are either zero or one depending on the predicted class, and hence the model doesn't generalize well for the sample in the area between the two classes.

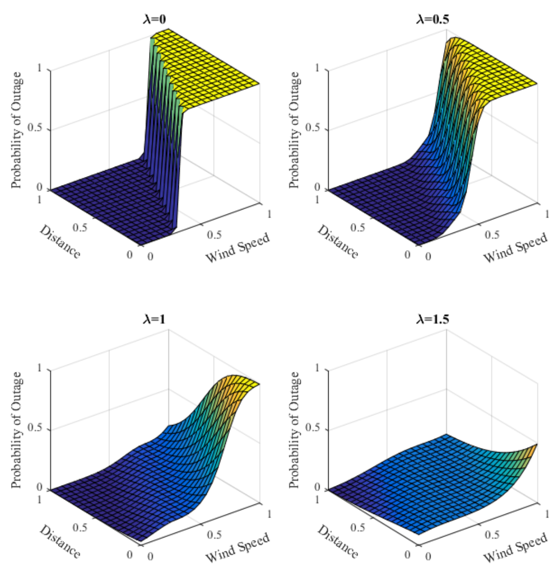


FIGURE 4-4- POSTERIOR PROBABILITY MODELS FOR VARIOUS VALUES OF λ

3.3 Evaluating probabilistic load curtailment estimation

Eight components are considered to be damaged in the path of the upcoming hurricane. The outage probability of these components is calculated based on estimated wind speed and distance from the center of the hurricane and through the proposed posterior probability estimation. Table 4-2 shows the distance and wind speed of each component, normalized based on the highest wind speed (obtained from the category of the hurricane) and the distance of the furthest impacted component from the center of the hurricane (line 44). The calculated outage probability is also shown in this table for each impacted component. As the results suggest, the components that are closer to the hurricane and experience higher wind speeds, such as line 46, show a very high probability of outage, here as much as 99.5%. On the other hand, the components far from the hurricane and subject to lower wind speeds may show very small chances of outages, such as line 44 which only has a 1.7% outage probability.

The obtained outage probabilities show a promising improvement compared to the existing work in this area which only provide a 0/1 output, i.e., showing whether each component is operational or on outage. Identifying outage probabilities would provide significant opportunities in better managing the available resources as the system response and recovery studies can shift from deterministic models to probabilistic models.

TABLE 4-2- COMPONENTS ALONG HURRICANE PATH AND THEIR PREDICTED OUTAGE PROBABILITIES

Component	Wind speed	Distance	Outage probability
Line 44	0.471	1.000	0.017
Line 45	0.471	0.873	0.032
Line 48	0.509	0.571	0.091
Line 50	0.509	0.555	0.077
Line 49	0.509	0.492	0.183
Line 47	0.644	0.444	0.220
Line 30	0.962	0.142	0.971
Line 46	0.994	0.120	0.995

These outage probabilities are used to define 2^8=256 scenarios, where all possible combinations of outage/operational sets of these components are considered. These scenarios are fed into the load curtailment estimation problem which is formulated using mixed-integer linear programming (MILP) and solved by CPLEX 12.6 [75]. A value of lost load of $1000/MWh is considered.

The problem objective is calculated as $1054507 in which $1024226 is the operation cost and the rest is the aggregated cost of load curtailment in all scenarios. The highest load curtailment is experienced in scenario 129, in which line 30 is in service and all other lines are on outage. The expected load curtailment in this scenario is 434 MWh, however the probability of this scenario is only 1.25×10^{-9}. The highest probability, 0.59, occurs in scenario 112 in which lines 30 and 46 are on outage and other lines are in service. However, there is no load curtailment in this scenario. The focus of this chapter is to estimate potential load curtailments in response to imminent hurricanes, however, other probabilistic factors, such as renewable energy generation can be easily formulated and integrated into the proposed model.

4.3. Conclusion

In this chapter, a probabilistic load curtailment estimation model was proposed through a three-step sequential method. At first, to determine a deterministic outage state of the grid components in response to a forecasted hurricane, a machine learning model based on TWSVM was proposed. Then, to convert the deterministic results into probabilistic outage states, a posterior probability sigmoid model was trained on the obtained results from the previous step. Finally, the obtained component outages were integrated into a load curtailment estimation model to determine the potential load curtailments in the system. The simulation results on a standard test system illustrated the high accuracy performance of the proposed method.

The work concludes that the probabilistic load curtailment estimation offers a viable prospect to understand the most impactful outage scenarios in the system, as well as the severity of their impact, in response to an upcoming hurricane, and opens significant opportunities in better planning for those events. In this work, since historical data for hurricanes at component level are limited, a synthetic data is used to show the effectiveness of the proposed method. In future, more detailed historical data for hurricanes will be requested from some of the utility companies affected by hurricanes. In addition, the authors are currently investigating applying the proposed probabilistic outage estimation model for renewable energy integration and accordingly studying the impact of growing renewable penetration on system resilience in response to hurricanes.

CHAPTER FIVE: CONCLUSION AND FUTURE WORK

Prediction of a component state in response to an extreme event is a challenging task in practice. An outage prediction model based on logistic regression was proposed to determine the probable outage of power grid components in response to an imminent hurricane. The acceptable performance of the proposed model was validated in this work. The logistic regression method is simple, fast, robust, and can efficiently handle the complexity of the decision boundary in terms of characteristic parameters. This method, however, requires much more data to achieve stable and meaningful results compared to other prediction models, such as support vector machine. Hence, a three-dimensional SVM was proposed to categorize system components into two classes of damaged and operational in response to an upcoming hurricane.

The proposed SVM was trained on historical data with three features related to each grid component—i.e., the resilience index, the distance of the component from the center of the hurricane, and the category of the hurricane (the wind speed). A synthetic set of data was generated to train the SVM, as the publicly available data on the impact of hurricanes on power grid components is limited. High accuracy was obtained by allowing some data points to enter an uncertain area by increasing the SVM margin, thus increasing the estimation accuracy for other components. Practicality was ensured by considering component deterioration in addition to other prevailing factors, and efficiency was guaranteed by outperforming other existing methods.

After training the SVM model, a minimum load curtailment problem was formulated to estimate the amount of load curtailment. The predictions obtained from the SVM model were integrated into a minimum load curtailment model and the potential nodal load curtailments—which are of utmost importance for grid operators in order to identify critical and prone-to-curtailment areas to proactively mobilize the restoration resources—were estimated. Finally, an electric power grid hardening model was proposed through localized and decentralized supply of power in certain regions. In contrast to existing literature in hardening and resilience enhancement, this model co-optimizes grid economic and resilience objectives by considering the intricate dependencies of the two.

Simulation results showed the effectiveness of the proposed SVM model compared to the results obtained from Logistic Regression, as a popular benchmark for two-class classification problem, and further demonstrated its acceptable performance in reaching high accuracy estimations. The proposed model can greatly help grid operators in estimating the components availability in response to extreme events, and therefore, better plan their resources for mitigation, response, and recovery.

The effectiveness of the proposed load curtailment estimation model were tested on IEEE 30-bus system with a combination of hurricane path and intensity scenarios. The results demonstrate that the proposed modelling framework is capable to effectively capture the dynamics of load curtailment estimation in response to extreme events. The results indicated that the proposed framework enables one to effectively identify the critical components in the power system, and prioritize the limited restoration resources.

The numerical simulations on the standard IEEE 118-bus test system illustrated the merits and applicability of the proposed hardening model. The results indicated that the proposed hardening model can produce a robust solution that can protect the system against multiple component outages due to a hurricane. Given the crucial importance of accurate power grid outage prediction, this model provides a practical forward-looking framework for utilities, local governments, and policy makers for a risk-informed operations management, emergency response planning, humanitarian logistics, and restoration of the life-line power grid infrastructure in both strategic level and real-time basis.

Finally, a probabilistic load curtailment estimation model was proposed through a three-step sequential method. At first, to determine a deterministic outage state of the grid components in response to a forecasted hurricane, a machine learning model based on TWSVM was proposed. Then, to convert the deterministic results into probabilistic outage states, a posterior probability sigmoid model was trained on the obtained results from the previous step. Finally, the obtained component outages were integrated into a load curtailment estimation model to determine the potential load curtailments in the system. The simulation results on a standard test system illustrated the high accuracy performance of the proposed method.

5.1. Future Work

The SVM method has numerous advantages including the ability to provide a global solution for data classification. It generates a unique global hyper-plane by solving a Quadratic Programming Problem (QPP) to separate the data samples of different classes rather than local boundaries as compared to other existing data classification approaches. Due to its better performance, SVM is one of the most widely-used classification techniques in data mining. One of the main challenges with the traditional SVM, however, is that it solves only one QPP problem to classify the data, which may not be suitable in cases of intertwined data. In addition, despite the good performance of SVM in several applications, the performance of SVM drops significantly when faced with imbalanced datasets, for example when the number of negative instances far outnumbers the positive instances, or vice versa [76]. This can be potentially problematic since the data of past hurricanes are imbalanced (i.e., the number of non-operational components is far less than the number of operational components).

The work concludes that the probabilistic load curtailment estimation offers a viable prospect to understand the most impactful outage scenarios in the system, as well as the severity of their impact, in response to an upcoming hurricane, and opens significant opportunities in better planning for those events. In this work, since historical data for hurricanes at component level are limited, a synthetic data is used to show the effectiveness of the proposed method. In future, more detailed historical data for hurricanes will be requested from some of the utility companies affected by hurricanes. In addition, the authors are currently investigating applying the proposed probabilistic outage estimation model for renewable energy integration and accordingly studying the impact of growing renewable penetration on system resilience in response to hurricanes.

REFERENCES

[1] Executive Office of the President, “Economic Benefits of Increasing Electric Grid Resilience to Weather Outages-August 2013.”

[2] R. J. Campbell, “Weather-related power outages and electric system resiliency,” 2012.

[3] “Green Tech Media, ‘New Jersey Launches $200M Energy Resilience Bank for Microgrids and Distributed Generation,’Availabl[online:]http://www.greentechmedia.com/articles/read/ New-Jersey-Launches-200M-Energy-Resilience-Bank-For-Microgrids-and-Distrib].”

[4] “Department of Energy, Office of Electricity Delivery & Energy Reliability. (2016). Emergency Situation Reports, accessed on Aug. 26, 2016. [Online]. Available: http://www.oe.netl.doe.gov/emergency_sit_rpt.aspx.”

[5] J. He, D. W. Wanik, B. M. Hartman, E. N. Anagnostou, M. Astitha, and M. E. Frediani, “Nonparametric Tree-Based Predictive Modeling of Storm Outages on an Electric Distribution Network,” *Risk Anal.*, vol. 37, no. 3, pp. 441–458, 2017.

[6] L. M. Branscomb, “Sustainable cities: Safety and security,” *Technol. Soc.*, vol. 28, no. 1, pp. 225–234, 2006.

[7] C. S. Holling, “Resilience and stability of ecological systems,” *Annu. Rev. Ecol. Syst.*, pp. 1–23, 1973.

[8] E. Hollnagel, D. D. Woods, and N. Leveson, *Resilience Engineering: Concepts and Precepts*. Ashgate Publishing, Ltd., 2007.

[9] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, “Identifying, understanding, and analyzing critical infrastructure interdependencies,” *IEEE Control Syst.*, vol. 21, no. 6, pp. 11–25, Dec. 2001.

[10] N. O. Attoh-Okine, A. T. Cooper, and S. A. Mensah, “Formulation of resilience index of urban infrastructure using belief functions,” *Syst. J. IEEE*, vol. 3, no. 2, pp. 147–153, 2009.

[11] O. O. Aderinlewo and N. O. Attoh-Okine, “Assessment of a Transportation Infrastructure System using Graph Theory,” 2013.

[12] A. Arab *et al.*, “Stochastic Pre-hurricane Restoration Planning for Electric Power Systems Infrastructure,” *Smart Grid IEEE Trans. On*, vol. 6, no. 2, pp. 1046–1054, 2015.

[13] A. Arab, A. Khodaei, Z. Han, and S. K. Khator, “Proactive Recovery of Electric Power Assets for Resiliency Enhancement,” *Access IEEE*, vol. 3, pp. 99–109, 2015.

[14] A. Arab, A. Khodaei, S. K. Khator, K. Ding, and Z. Han, “Post-hurricane transmission network outage management,” in *Proc. IEEE Great Lakes Symp. Smart Grid New Energy Econ*, 2013, pp. 1–6.
[15] A. Arab, A. Khodaei, S. K. Khator, and Z. Han, “Transmission network restoration considering AC power flow constraints,” in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2015, pp. 816–821.
[16] A. Arab, A. Khodaei, S. Khator, and Z. Han, “Electric Power Grid Restoration Considering Disaster Economics,” 2016.
[17] L. Molyneaux, L. Wagner, C. Froome, and J. Foster, “Resilience and electricity systems: A comparative analysis,” *Energy Policy*, vol. 47, pp. 188–201, 2012.
[18] D. A. Reed, K. C. Kapur, and R. D. Christie, “Methodology for assessing the resilience of networked infrastructure,” *Syst. J. IEEE*, vol. 3, no. 2, pp. 174–180, 2009.
[19] M. Ouyang, L. Dueñas-Osorio, and X. Min, “A three-stage resilience analysis framework for urban infrastructure systems,” *Struct. Saf.*, vol. 36, pp. 23–31, 2012.
[20] M. Ouyang and L. Dueñas-Osorio, “Multi-dimensional hurricane resilience assessment of electric power systems,” *Struct. Saf.*, vol. 48, pp. 15–24, 2014.
[21] C. M. Bishop, “Pattern recognition,” *Mach. Learn.*, vol. 128, pp. 1–58, 2006.
[22] C. K. Pang, F. S. Prabhakara, A. H. El-abiad, and A. J. Koivo, “Security Evaluation in Power Systems Using Pattern Recognition,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-93, no. 3, pp. 969–976, May 1974.
[23] T. T. Teo, T. Logenthiran, and W. L. Woo, “Forecasting of photovoltaic power using extreme learning machine,” in *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, 2015, pp. 1–6.
[24] L. Wehenkel, “Machine learning approaches to power-system security assessment,” *IEEE Expert*, vol. 12, no. 5, pp. 60–72, 1997.
[25] S. D. Guikema, “Natural disaster risk analysis for critical infrastructure systems: An approach based on statistical learning theory,” *Reliab. Eng. Syst. Saf.*, vol. 94, no. 4, pp. 855–860, 2009.
[26] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, “Artificial neural network and support vector machine approach for locating faults in radial distribution systems,” *IEEE Trans. Power Deliv.*, vol. 20, no. 2, pp. 710–721, 2005.
[27] R. Nateghi, S. D. Guikema, and S. M. Quiring, “Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes,” *Risk Anal.*, vol. 31, no. 12, pp. 1897–1906, 2011.
[28] D. J. Sobajic and Y.-H. Pao, “Artificial neural-net based dynamic security assessment for electric power systems,” *IEEE Trans. Power Syst.*, vol. 4, no. 1, pp. 220–228, 1989.
[29] A. D. Papalexopoulos, S. Hao, and T.-M. Peng, “An implementation of a neural network based load forecasting model for the EMS,” *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 1956–1962, 1994.
[30] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, “Wind power forecasting using advanced neural networks models,” *IEEE Trans. Energy Convers.*, vol. 11, no. 4, pp. 762–767, 1996.

[31] M. Panteli and P. Mancarella, “The grid: Stronger, bigger, smarter?: Presenting a conceptual framework of power system resilience,” *IEEE Power Energy Mag.*, vol. 13, no. 3, pp. 58–66, 2015.
[32] D. R. Godschalk, “Urban hazard mitigation: creating resilient cities,” *Nat. Hazards Rev.*, vol. 4, no. 3, pp. 136–143, 2003.
[33] J. Winkler, L. Duenas-Osorio, R. Stein, and D. Subramanian, “Performance assessment of topologically diverse power systems subjected to hurricane events,” *Reliab. Eng. Syst. Saf.*, vol. 95, no. 4, pp. 323–336, 2010.
[34] N. Perrier *et al.*, “A survey of models and algorithms for emergency response logistics in electric distribution systems. Part I: Reliability planning with fault considerations,” *Comput. Oper. Res.*, vol. 40, no. 7, pp. 1895–1906, 2013.
[35] N. Perrier *et al.*, “A survey of models and algorithms for emergency response logistics in electric distribution systems. Part II: Contingency planning level,” *Comput. Oper. Res.*, vol. 40, no. 7, pp. 1907–1922, 2013.
[36] R. A. Francis, S. M. Falconi, R. Nateghi, and S. D. Guikema, “Probabilistic life cycle analysis model for evaluating electric power infrastructure risk mitigation investments,” *Clim. Change*, vol. 106, no. 1, pp. 31–55, 2011.
[37] R. E. Brown, “Hurricane hardening efforts in Florida,” in *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, 2008, pp. 1–7.
[38] R. Eskandarpour and A. Khodaei, “Machine Learning based Power Grid Outage Prediction in Response to Extreme Events,” *IEEE Trans. Power Syst.*, 2016.
[39] R. Eskandarpour and A. Khodaei, “Leveraging Accuracy-Uncertainty Tradeoff in SVM to Achieve Highly Accurate Outage Predictions,” *IEEE Trans. Power Syst.*, 2017.
[40] R. Eskandarpour, A. Khodaei, and A. Arab, “Improving power grid resilience through predictive outage estimation,” in *2017 North American Power Symposium (NAPS)*, 2017, pp. 1–5.
[41] R. Eskandarpour, A. Khodaei, and J. Lin, “Event-Driven Security-Constrained Unit Commitment,” in *Innovative Smart Grid Technologies*, 2016.
[42] A. Arab *et al.*, “Stochastic Pre-hurricane Restoration Planning for Electric Power Systems Infrastructure,” *Smart Grid IEEE Trans. On*, vol. 6, no. 2, pp. 1046–1054, 2015.
[43] A. Arab, A. Khodaei, Z. Han, and S. K. Khator, “Proactive Recovery of Electric Power Assets for Resiliency Enhancement,” *Access IEEE*, vol. 3, pp. 99–109, 2015.
[44] A. Arab, A. Khodaei, S. K. Khator, K. Ding, and Z. Han, “Post-hurricane transmission network outage management,” in *Proc. IEEE Great Lakes Symp. Smart Grid New Energy Econ*, 2013, pp. 1–6.
[45] A. Arab, A. Khodaei, S. K. Khator, and Z. Han, “Transmission network restoration considering AC power flow constraints,” in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2015, pp. 816–821.
[46] A. Arab, A. Khodaei, S. Khator, and Z. Han, “Electric Power Grid Restoration Considering Disaster Economics,” *IEEE Access*, vol. 4, pp. 639–649, 2016.

[47] R. Eskandarpour, G. Edwards, and A. Khodaei, “Resilience-constrained unit commitment considering the impact of microgrids,” in *North American Power Symposium (NAPS), 2016*, 2016, pp. 1–5.
[48] F. E. Harrell, “Regression modeling strategies,” *Implemented R Package ‘rms’ Version*, vol. 3, no. 3, 2014.
[49] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[50] J. Platt, “Sequential minimal optimization: A fast algorithm for training support vector machines,” 1998.
[51] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines,” *ACM Trans. Intell. Syst. Technol. TIST*, vol. 2, no. 3, p. 27, 2011.
[52] M. Ouyang and L. Dueñas-Osorio, “An approach to design interface topologies across interdependent urban infrastructure systems,” *Reliab. Eng. Syst. Saf.*, vol. 96, no. 11, pp. 1462–1473, 2011.
[53] R. Kohavi and others, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Ijcai*, 1995, vol. 14, pp. 1137–1145.
[54] L. Economics, “Estimating the Value of Lost Load,” *Brief. Pap. Prep. ERCOT 17 June 2013*, 2013.
[55] T. Schott *et al.*, “The Saffir-Simpson hurricane wind scale,” *Natl. Weather Serv.*, 2012.
[56] R. Eskandarpour, A. Khodaei, and J. Lin, “Event-Driven Security-Constrained Unit Commitment with Component Outage Estimation Based on Machine Learning Method,” in *North American Power Symposium*, 2016.
[57] S. Ding, J. Yu, B. Qi, and H. Huang, “An overview on twin support vector machines,” *Artif. Intell. Rev.*, pp. 1–8, 2014.
[58] R. Khemchandani and S. Chandra, “Twin Support Vector Machines (TWSVM) for Classification,” in *Twin Support Vector Machines*, Springer, 2017, pp. 43–62.
[59] Y.-H. Shao, N.-Y. Deng, Z.-M. Yang, W.-J. Chen, and Z. Wang, “Probabilistic outputs for twin support vector machines,” *Knowl.-Based Syst.*, vol. 33, pp. 145–151, 2012.
[60] R. Batuwita and V. Palade, “Class imbalance learning methods for support vector machines,” *Imbalanced Learn. Found. Algorithms Appl.*, vol. 83, 2013.
[61] K. Veropoulos, C. Campbell, N. Cristianini, and others, “Controlling the sensitivity of support vector machines,” in *Proceedings of the international joint conference on AI*, 1999, vol. 55, p. 60.
[62] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: synthetic minority over-sampling technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
[63] Z. Lin, Z. Hao, X. Yang, and X. Liu, “Several SVM ensemble methods integrated with under-sampling for imbalanced data learning,” in *International Conference on Advanced Data Mining and Applications*, 2009, pp. 536–544.
[64] T. Imam, K. M. Ting, and J. Kamruzzaman, “z-SVM: an SVM for improved classification of imbalanced data,” in *Australasian Joint Conference on Artificial Intelligence*, 2006, pp. 264–273.

[65] P. Li, K. L. Chan, and W. Fang, “Hybrid kernel machine ensemble for imbalanced data sets,” in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 2006, vol. 1, pp. 1108–1111.
[66] R. Khemchandani and S. Chandra, “Twin support vector machines for pattern classification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 5, pp. 905–910, 2007.
[67] S. P. Arjunan, D. K. Kumar, and G. R. Naik, “A machine learning based method for classification of fractal features of forearm sEMG using twin support vector machines,” in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual international conference of the IEEE*, 2010, pp. 4821–4824.
[68] G. R. Naik and D. K. Kumar, “Twin SVM for gesture classification using the surface electromyogram,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 301–308, 2010.
[69] H. Cong, C. Yang, and X. Pu, “Efficient speaker recognition based on multi-class twin support vector machines and GMMs,” in *Robotics, Automation and Mechatronics, 2008 IEEE Conference on*, 2008, pp. 348–352.
[70] X. Zhang, X. Gao, and Y. Wang, “Twin support tensor machines for MCS detection,” *J. Electron. China*, vol. 26, no. 3, pp. 318–325, 2009.
[71] J. Platt and others, “Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods,” *Adv. Large Margin Classif.*, vol. 10, no. 3, pp. 61–74, 1999.
[72] B. Zadrozny and C. Elkan, “Transforming classifier scores into accurate multiclass probability estimates,” in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2002, pp. 694–699.
[73] “IEEE 118-Bus System - Illinois Center for a Smarter Electric Grid (ICSEG).” [Online]. Available: http://publish.illinois.edu/smartergrid/ieee-118- bus-system/.
[74] A. Arab, E. Tekin, A. Khodaei, S. K. Khator, and Z. Han, “Dynamic maintenance scheduling for power systems incorporating hurricane effects,” in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, 2014, pp. 85–90.
[75] “CPLEX 12.” [Online]. Available: https://www.gams.com/latest/docs/S_CPLEX.html.
[76] G. Wu and E. Y. Chang, “Class-boundary alignment for imbalanced dataset learning,” in *ICML 2003 workshop on learning from imbalanced data sets II, Washington, DC*, 2003, pp. 49–56.

APPENDIX – PUBLICATIONS

1. Eskandarpour, Rozhin, and Amin Khodaei. "Probabilistic load curtailment estimation using posterior probability model and twin support vector machine." Journal of Modern Power Systems and Clean Energy (2019): 1-11.

2. Eskandarpour, Rozhin, and Amin Khodaei. "Machine learning based power grid outage prediction in response to extreme events." IEEE Transactions on Power Systems 32, no. 4 (2017): 3315-3316.

3. Eskandarpour, Rozhin, and Amin Khodaei. "Leveraging Accuracy-Uncertainty Tradeoff in SVM to Achieve Highly Accurate Outage Predictions." IEEE Transactions on Power Systems 33, no. 1 (2018): 1139-1141.

4. Eskandarpour, Rozhin, and Amin Khodaei. "Component outage estimation based on Support Vector Machine." In Power & Energy Society General Meeting, 2017 IEEE, pp. 1-5. IEEE, 2017.

5. Eskandarpour, Rozhin, Amin Khodaei, and Ali Arab. "Load Curtailment Estimation in Response to Extreme Events." CIGRE Grid of the Future Symposium, Cleveland, Ohio, 2017.

6. Eskandarpour, Rozhin, Amin Khodaei, and Ali Arab. "Improving power grid resilience through predictive outage estimation." In Power Symposium (NAPS), 2017 North American, pp. 1-5. IEEE, 2017.

7. Eskandarpour, Rozhin, Amin Khodaei, A. Paaso, and N. M. Abdullah. "Artificial Intelligence Assisted Power Grid Hardening in Response to Extreme Weather Events." CIGRE Grid of the Future Symposium, Reston, VA, 2018.

8. Eskandarpour, Rozhin, Hossein Lotfi, and Amin Khodaei. "Optimal microgrid placement for enhancing power system resilience in response to weather events." In 2016 North American Power Symposium (NAPS), pp. 1-6. IEEE, 2016.

9. Eskandarpour, Rozhin, Amin Khodaei, and Jeremy Lin. "Event-driven security-constrained unit commitment." In 2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), pp. 1-5. IEEE, 2016.

10. Eskandarpour, Rozhin, George Edwards, and Amin Khodaei. "Resilience-constrained unit commitment considering the impact of microgrids." In 2016 North American Power Symposium (NAPS), pp. 1-5. IEEE, 2016.

11. R. Eskandarpour, A. Khodaei, and J. Lin, "Event-Driven Security-Constrained Unit Commitment with Machine Learning based Outage Estimation," North American Power Symposium (NAPS), 2016.

12. Eskandarpour, Rozhin, Amin Khodaei, " Predicting Power Grid Outage Response to Extreme Events from Historical data." CIGRE Grid of the Future Symposium, Philadelphia, Pennsylvania, 2016.