A PROBABILISTIC MODELING FRAMEWORK FOR INTEGRATION OF DISTRIBUTED ENERGY RESOURCES

A Dissertation

by

AAQIB AHMAD PEERZADA

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Miroslav M. Begovic
Co-Chair of Committee,	Robert Balog
Committee Members,	Tie Liu
	Sergiy Butenko
Head of Department,	Miroslav Begovic

December 2022

Major Subject: Electrical Engineering

ABSTRACT

The emergence of distributed energy resources has led to new challenges in the operation and planning of power networks. Of particular significance is the introduction of a new layer of complexity that manifests in the form of new uncertainties that could severely limit the resiliency and reliability of modern power networks. Some of the new uncertainties that emerge as a direct consequence of the integration of distributed energy resources include generation uncertainties typical of solar and wind power, uncertain consumer demand patterns due to the increasing adoption of unconventional loads such as Plug-in Hybrid Electric Vehicles, topological uncertainties that include outages of one or many components of a power system. To facilitate the widespread adoption of distributed energy resources, it is thus essential to develop robust methodologies that would adequately capture and quantify the uncertainties associated with using them. Such decision-making problems that involve uncertainties embedded in the input data naturally lend themselves to be addressed by statistical decision-theoretic methods.

A natural advantage of using statistical decision theory in addressing power system uncertainties is using sample information to make inferences about the unknown quantities. Thus, using computational methods like Bayesian statistics and Markov Chain Monte Carlo is natural within the context of decision-making under uncertainty in energy systems. This research proposes the use of computational methods such as scenario generation techniques, probabilistic mixture models, Bayesian analysis, and Markov Chain Monte Carlo to model the complex stochastic processes such as solar generation, power system load, and various topological uncertainties like accelerated aging and the premature failing of components such as On-load Tap Changers and switched capacitor banks.

Furthermore, this research work is also concerned with investigating the impact of modern reactive power compensation in the form of a solid-state-based capacitor-less power quality compensator to further the integration of distributed resources, particularly distributed roof-top solar generation in low voltage distribution networks. To that end, a stochastic cost-benefit equation is

developed, considering the uncertainties associated with distributed energy resources. The cost-benefit study investigates the economic viability of deploying such power electronics-based reactive power compensation devices in low-voltage distribution networks.

DEDICATION

To my wife, son, mother, and father.

ACKNOWLEDGMENTS

I'd like to express my sincere gratitude for assistance and support provided to me in connection with this thesis to several people, including

First and foremost, my supervisor Prof. Miroslav Begovic for his invaluable guidance and support in the absence of which this thesis would not have come to fruition. His careful observation of this manuscript's development and the articles published based on this work have been of great importance and instrumental in sculpting the principal facade of this work. In addition to honing my technical skills, Prof. Begovic's more intimate musings on love and life and the importance of being earnest in pursuing scientific knowledge have been a source of great inspiration.

I'd also like to extend my sincere appreciation to my co-chair Prof. Robert S. Balog, for collaborating with me and allowing me to work on an exciting project at the bleeding edge of power systems and power electronics. This collaboration enabled me to gain a fundamental understanding of several new ideas and helped me develop new skills to combat the emerging issues in power engineering research.

I'd also like to extend my appreciation to Prof. Wesam Rohouma, with whom I have shared many insightful discussions on a wide variety of problems, which eventually helped me delve deeper and attain a better understanding of several research problems.

My parents, whose patience and unbridled love and support helped me navigate moments of ambiguity and never lose sight of my final goal.

My beautiful wife, who is my defense against nihilism and whose fierce love keeps many of my existential quandaries at bay. Whenever I am faced with doubt or bogged down by dogma, I turn to her in search of hope and purpose.

Finally, my son, whose arrival into this world redefined the meaning of love and life for me. In him, I find the strength to battle my inner demons. In him, I find the reason to live. .

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Miroslav M. Begovic [doctoral advisor], Professor Robert S. Balog [doctoral co-advisor], Professor Tie Liu of the Department of Electrical and Computer Engineering, and Professor Sergiy Butenko of the Department of Industrial and Systems Engineering.

Funding Sources

The graduate study was supported by NPRP grant # 9-204-2-103 from the Qatar National Research Foundation and a fellowship from the Texas Engineering Experimentation Station (TEES).

NOMENCLATURE

GGD	Generalized Gaussian Distribution
GGMM	Generalized Gaussian Mixture Model
DERs	Distributed Energy Resources
EVs	Electric Vehicles
$N_{EV}(t)$	Number of EV arrivals in $[0,t]$
Ω_N	Sample Space on which the Counting Process is defined
S_k	k^{th} Arrival Epoch of Homogeneous Poisson Process
$\lambda(t)$	Time Varying Intensity Function
$\wedge(t)$	Mean Value Function
X_k	k^{th} Inter-arrival time
$F(.)$	CDF Function
λ^+	Constant Rate Intensity Function
S_k^{NHPP}	k^{th} Arrival Epoch of Nonhomogeneous Poisson Process
$SoC_{Arrival}$	State of Charge of electric vehicle battery at Arrival
E_{req}	Energy Required to charge electric vehicle battery
T_{ch}	Time required to charge electric vehicle battery
P	L-2 Charging Rate of electric vehicle
$f_Y(.)$	Probability Density Function
π_j	Weight assigned to j^{th} component density function
Ψ	Vector of Model Parameters
$\mathbf{Z}$	Vector of Hidden Variables
$\mathbf{Y}$	Vector of Load Measurements

β	Shape Parameter of Generalized Gaussian Distribution
s	Scale Parameter of Generalized Gaussian Distribution
μ	Location Parameter of Generalized Gaussian Distribution
ε	Tolerance level
$\mathbf{S}$	Vector of arrival epochs
d_m	Daily Driven Miles
E_{cons}	Electricity consumption in kWh/100 miles
C_b	EV Battery Capacity in kWh
d	Number of dimensions
Σ	$d \times d$ Positive definite Symmetric Covariance Matrix
ζ	Multidimensional Shape Parameter Analogue
$\gamma_2(\overline{\mathbf{Y}})$	Multidimensional kurtosis coefficient of d- dimensional vector $\overline{\mathbf{Y}}$
k_t	Clearness Index
$ARMA(p,q)$	Autoregressive moving average model with orders p and q
G^i_{PV}	Stochastic Process representing power generated at bus i
$MBCA$	Marginal Benefit of Carbon Abatement
C_T	Total operational cost in $\$/year$
c_t	Carbon Tax rate in $\$/MT$
$LCOE$	Levelized cost of energy
$\{T_i \vert i = 1,2,...,r\}$	Random sample of observed failures
$\{S_j \vert j = r+1,...,N\}$	Random set of in service assets
$K_{\beta,r}$	Normalization constant of Shape Parameter
$K_{\theta,r}$	Normalization constant of Scale Parameter
$K_{\alpha,r}$	Normalization constant of Stress Parameter
π^*	Invariant Distribution

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND LITERATURE REVIEW

The electric grid is currently in the midst of a radical transformation. In particular, three ongoing trends are well posed to cause significant disruptions in how the electric grid functions. The first disruptive trend is the trend of electrification. Electrification refers to more things powered by electricity. The transportation sector is the biggest beneficiary of the rapid growth in electrification. One of the major reasons is the progressively declining battery costs from $1000 per kilowatt-hour (kWh) to less than $200 per kilowatt-hour (kWh) in 2021. This has enabled lower-cost models such as Nissan Leaf and Tesla Model 3 to increase production. The advances in the electric vehicle (EV) technology have further environmental benefits by shifting many end users of electricity away from fossil fuels and thus playing an important role in lowering the transportation sector's carbon footprint. In the United States, the transportation sector accounted for roughly 27% of the total greenhouse emissions in 2020 [6]. Light-duty vehicles accounted for more than half of the emissions from the transportation sector, thus making it a critical area for decarbonization.

The second trend with the potential to disrupt the power industry is the trend of decentralization. Decentralization is an umbrella term used to describe technologies that enable the shift from centralized to distributed generation. These include distributed generation from renewable sources such as solar and wind. Distributed storage helps alleviate the local peak demand by collecting the excess electrical energy from renewable sources and flattening the sharp peaks in electrical energy consumption. Demand response is another disruptive technology that enables energy control during peak hours and allows customers to respond to high pricing periods. The deployment of solar photovoltaic panels has been extremely effective in many parts of the world. In 2020, the global installed capacity of solar photovoltaics amounted to roughly 773 gigawatts. The average cost of a residential solar panel system in the United States is $2.94 per watt [1]. New technologies in this area, like solar photovoltaic tiles and buildings, integrated with solar photovoltaic, further broaden the potential of distributed generation.

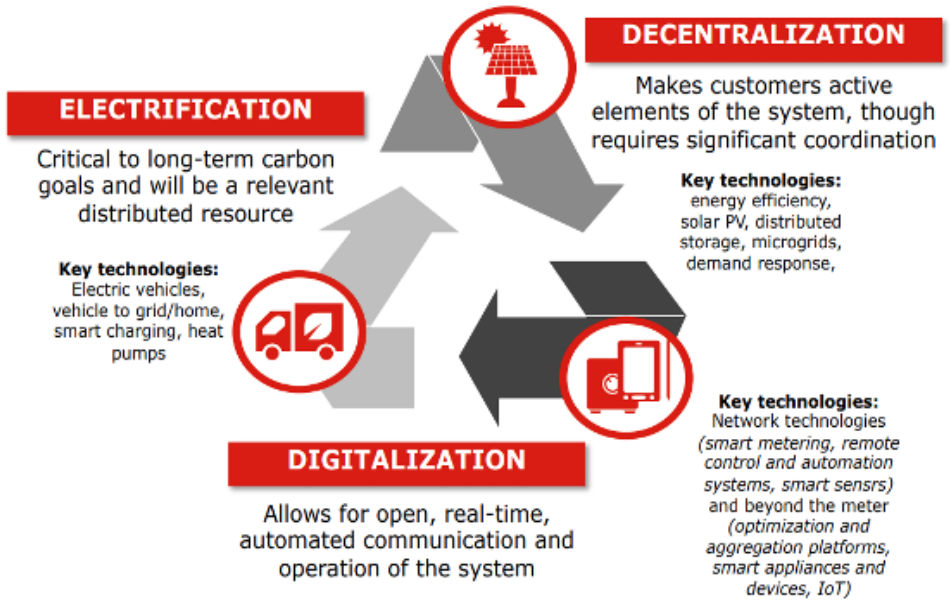

Figure 1.1: Trends in the grid-edge technologies[1]

The third important trend that is reshaping the power industry is digitization. Digitization refers to the increased use of smart devices that enable communication across the grid and provide crucial data for grid management. Smart meters, new IoT sensors, network remote control, and automated systems that focus heavily on grid optimization and aggregation facilitate the real-time operation of the grid while improving situational awareness and utility services. Such real-time data from smart devices is critical for integrating distributed generation. A good example that illustrates the importance of smart devices is the phasor measurement data captured by the phasor measurement units (PMUs). The real-time phasor data consists of voltage and current phasors at the loading location in the grid, which can be converted to give a quick estimate of the stability margin. This procedure bypasses the complex computation of the system dynamics and is a good alternative in situations that require the deployment of fast and emergency controls[7]. Though the market penetration of smart devices and IoT sensors is low, current projections predict a dramatic increase in the use of such devices.

The overarching objective of the grid-edge technologies is to create a decentralized grid with a renewed focus on reliability and environmental sustainability. By creating a more efficient and

resilient electrical grid, grid-edge technologies can unlock a significant economic value for the utilities and the customers. A recent survey by the world economic forum proposed a $2 trillion valuation of the transformation of the electric grid over the next decade [1]. Figure 1.2 shows the adoption rates of key technologies such as telephone and radio and compares them with the adoption rate of the grid-edge technologies. The adoption rate of the grid-edge technologies will likely follow the typical S-curve as seen with previous technologies.

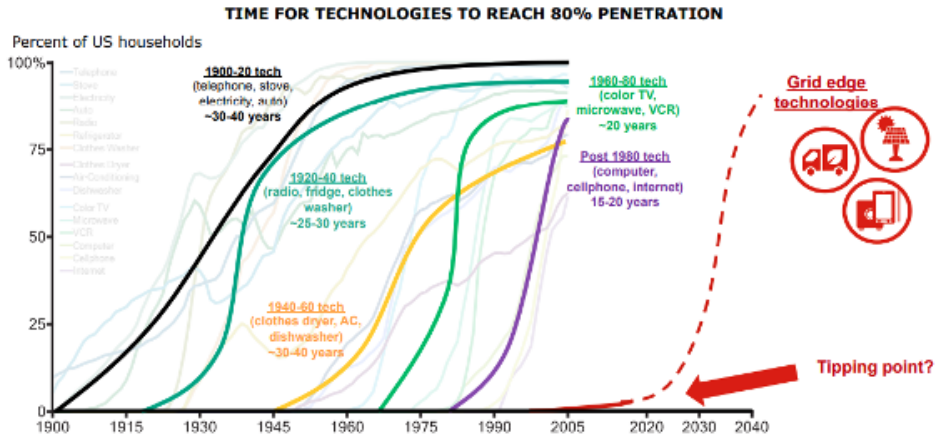

Figure 1.2: Adoption rate of grid-edge technologies [1]

The transition of the electric grid, however, faces major challenges that range from technical to regulatory issues. This dissertation addresses the technical issues associated with integrating distributed energy resources (DERs). A major technical challenge with the use of DERs is the problem of intermittency which adds a new layer of uncertainty to the operation and planning of the power grid. Most engineering problems are subject to uncertainty due to the inherent randomness of natural phenomena and/or the lack of perfect knowledge about a physical process. In any case, to achieve a cost-effective solution that improves the system's reliability, it is imperative to tackle such uncertainties via computational models that are immune or less sensitive to environmental influences or lack of expert knowledge. In energy systems, the ever-increasing penetration of distributed energy resources (DERs) has introduced new uncertainties that significantly impact the operation and planning of modern power systems. Uncertainty characterization in the context

of DER-rich power systems is crucial for the system's increased reliability and for facilitating the widespread adoption of such distributed energy resources. The power system uncertainties are usually grouped into two categories: a) Technical uncertainties and b)Economic uncertainties [8]. The technical uncertainties are further grouped into topological uncertainties related to network topologies like forced outages or failures of lines, generators, transformers, etc., and operational uncertainties tied to operational parameters like demand and generation. The economic uncertainties include the cost of fuel, pool prices of electricity, business taxes, cost of emissions etc.

1.1 Demand Uncertainty

One of the most noticeable operational parameters in a power system is the demand or load. The load on a power system is highly time-dependent, usually peaking in the evening and is usually lowest during the last quarter of the night. Traditionally, load uncertainty arises due to environmental factors, appliance variations, and consumer behavior. However, the embedded generation uncertainties associated with using DERs further accentuate the randomness in the system demand. For example, the intermittent nature of renewable energy introduces random fluctuations in the load, resulting in a voltage profile subject to higher variability. Similarly, the charging of Plug-In Hybrid Electric Vehicles (PHEVs) and fully Electric Vehicles (EVs) disrupts the normal load profiles, which, if left accounted for, could lead to undesirable peaks in the electric energy consumption. Even in the absence of DER-induced uncertainties, power system demand patterns show a lot of variabilities when measured at different buses across the network. This is especially true for distribution systems where different load types follow dissimilar patterns of use and, as such, cannot be represented by a single probability function.

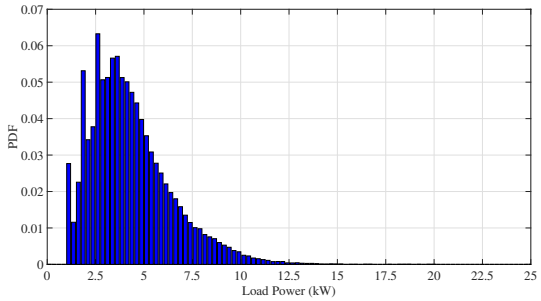

Figure 1.3: Residential Load Distribution

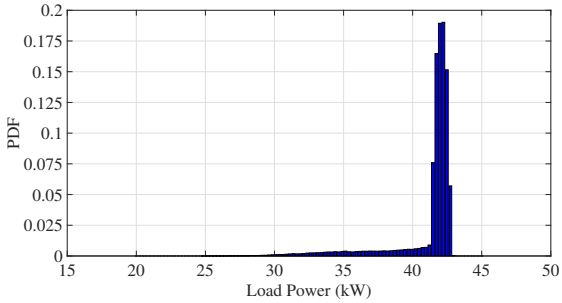

Figure 1.4: Commercial Load Distribution

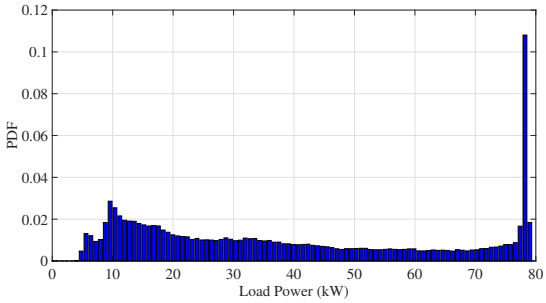

Figure 1.5: Large Industrial Load Distribution

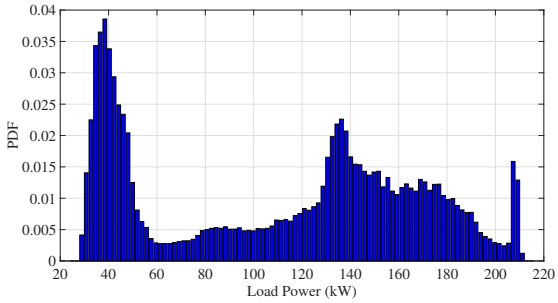

Figure 1.6: Small Industrial Load Distribution

An example of variability in the probability distribution of the different types of loads is presented in Figures 1.3-1.6. If uncertainty characterization is done by fitting a univariate or a multivariate probability distribution, then it is clear that no single density function can accurately quantify the stochasticity embedded in the different load types. In the classical literature, the Gaussian distribution has often been used to model loads. However, recent experimental and theoretical research has shown that Gaussian distribution is not justified for all the loads. This time-dependent variable behavior of the load is further complicated by the addition of renewable energy to the existing generation pool. An example of such complexity is the concept of net load. Traditionally, the system load profile has been extensively used for power system management applications. However, the accelerating expansion of renewables has rendered the system load profile less informative and thus inadequate for most operation and planning applications. Instead, system net load, defined as the demand that must be met by dispatchable (non-intermittent) sources, has gained popularity. Since the renewable output is subject to changes in the weather conditions, which can take place over a relatively faster time scale, the system net load as seen from a secondary of the

substation, also referred to as diversified demand, or at the point of common coupling experiences a greater degree of variability. The increased variability results in a load profile with a higher degree of "peakiness" than the traditional load. The increased peak behavior of system net load can be attributed to an overall increase in the uncertainty due to the addition of intermittent generation to the existing energy mix. In addition to an increase in the demand uncertainty, higher penetration of renewable output can also lead to negative demand or "back-feeding" to the grid. This happens when the energy produced by the distributed resources exceeds the demand and is a consequence of a mismatch between the peak solar output that usually happens sometime around noon and the peak demand that is frequently observed in the evening. The "back-feeding" to the grid disrupts the protection schemes usually designed to work in one direction downstream from the substation transformer.

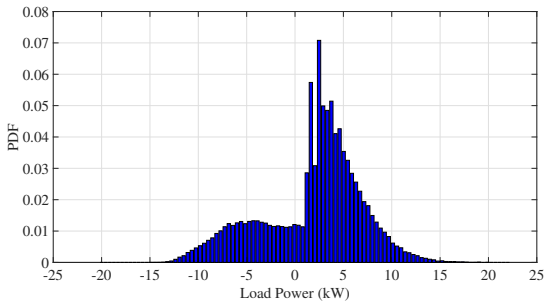

Figure 1.7: Net Residential Load Distribution

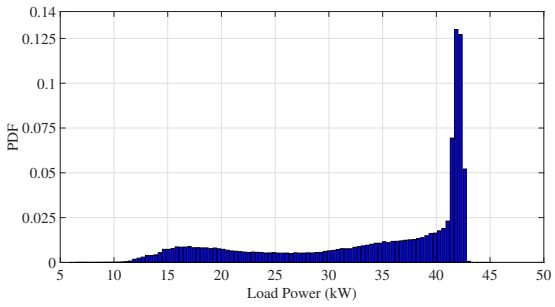

Figure 1.8: Net Commercial Load Distribution

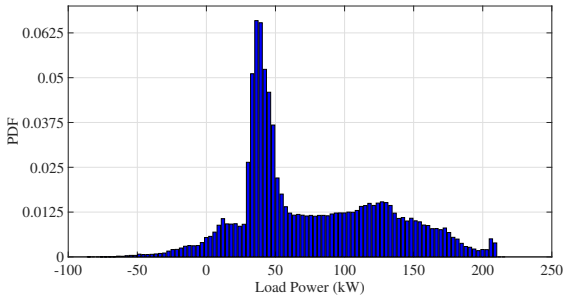

Figure 1.9: Net Large Industrial Load Distribution

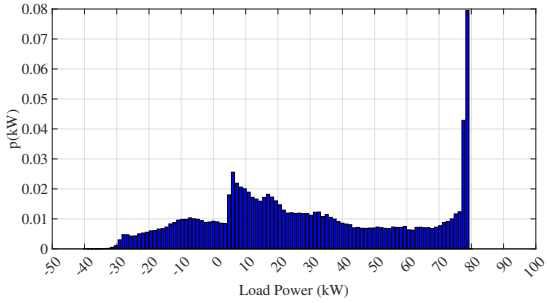

Figure 1.10: Net Small Industrial Load Distribution

Figures 1.7-1.9 show the variations in different load types when distributed solar generation is added to the existing generation mix. The net load at a load bus is obtained by subtracting the solar output from the load. As discussed, introducing solar generation results in a histogram plot with a pronounced "peaky" behavior. A quick visual comparison between Figures 1.3-1.6 and Figures 1.7-1.9 reveals the impact that distributed solar generation, in this example, has on the consumer demand patterns. Clearly, a Gaussian assumption is not justified for modeling uncertainty in consumer demand patterns.

The most common method for characterizing load uncertainty is the probabilistic approach. The probabilistic method is predicated on finding the best fit for the empirical load distribution. The most commonly used distribution in the literature is the Gaussian distribution [9]. However, as discussed, the Gaussian distribution is not a good model for characterizing the load uncertainty since the power system load, especially in distribution systems, does not follow any particular probability distribution [10]. Despite this critical observation, several attempts have been made to model the loads using unimodal distributions [11],[12],[13],[14],[15]. The limitations of the unimodal distributions to model multimodal load data prompted research into the applicability of probability mixture models for load modeling. Early work on using mixture models for the statistical load representation is presented in [16]. The study uses a Gaussian mixture model (GMM) to represent the probability density function of all the loads in a distribution system. The study uses the Expectation-Maximization algorithm to evaluate the parameters of the mixture model. The study, however, ignores the impact of distributed energy resources, for example, distributed generation and plug-in electric vehicles, on the density function of different load types in a distribution system. The study [17] uses a Gaussian Mixture Model (GMM) to approximate non-Gaussian density functions such as correlated wind power output and aggregated load in the presence of non-Gaussian correlated random input variables. The study [18] evaluates the performance of the GMM and Mixture of Factor Analyzers (MFA) method in modeling residential loads, and the results are compared with the existing load models. The study concludes that GMM and MFA offer superior performance characteristics compared to the existing British load model. The study [19]

uses probability density functions based on GMM to statistically quantify key charging metrics of electric vehicles. The study uses real data from 221 electric vehicles as part of the largest trial in the UK and Europe. A key constraint in the application of GMM is that the area under the curve of each component density must equal unity over the entire sample space. The study in [20] develops a generalized model based on GMM to fit the density functions of wind power ramping. The generalized model differs from the conventional GMM in that the integral of each component density over the entire sample space is not required to be unity. Also, the associated weights of each component density can be negative as opposed to a conventional GMM, where each component's weight must be nonnegative. However, the component density functions used in the model are Gaussian functions. The nonlinear least squares method with a trust-region algorithm is used to obtain the mixture components. The optimal number of components is obtained by minimizing the euclidean distance between model fit and actual histogram distribution. The study in [21] uses a GMM to fit the multimodal empirical probability density function of wind power generation. In [22], GMM is used to model the density function of the error in the wind power forecast. However, the model parameters are obtained by an optimization algorithm based on the Riemannian manifold. The proposed L-BFGS optimization method requires fewer iterations considering multivariate data. The study in [23] provides analytical solutions to the load flow problem for low voltage network planning, considering a GMM for the load as an input. The study in [24] uses a GMM to model the joint distribution of the measured wind power and the forecasts, and the results are compared with Gaussian, Beta, and t-Location distributions. The study uses Root Mean Square Error (RMSE) metric to assess the fitting performance. It is shown that the GMM has the lowest RMSE. The mixture models based on Gaussian distribution have also been applied to quantify the uncertainty in renewable generation for use in market clearing mechanisms. In [25], a peer-to-peer (P2P) joint energy and reserve market is proposed where agents can negotiate with neighboring agents. The uncertainty of all the renewable energy agents is modeled using a GMM.

1.2 Generation Uncertainty

Outside the atmosphere, the incident solar radiation on any arbitrarily-sloped surface can be predicted exactly. However, due to the irregular and random movement of the clouds in the atmosphere, such an exact prediction of the incident solar radiation on any arbitrarily-sloped plane on the earth's surface is impossible. The impact of the cloud movement on the net irradiance incident on the solar panel can be quantified in terms of the clearness index. The clearness index is defined as the ratio of the global horizontal irradiance on the earth's surface and the irradiance calculated for cloudless conditions, often referred to as clear sky irradiance. The clearness index cannot be predicted with complete confidence and hence must be treated as a random variable [26]. The probability model of the clearness index can be constructed by utilizing the statistical distributions of past occurrences. Such a probability model can be used to predict the value of the clearness index for some future event with some probability and within specified limits. The model can further synthesize changes in the clearness index to model the cloud movement at shorter time scales (sub-hourly). Such a synthesis, however, must be done by utilizing a large measured historical data set of the changes in the clearness index to offset the effect of one bad day or bad week [27]. In the past several researchers have studied the statistical distributions of the clearness index by measuring the irradiance data at hourly and even sub-hourly time scales [28],[29]. While the density function appears unimodal when measured at hourly time steps, the sub-hourly measurements (5-minutes) have revealed the density function of the clearness index to have a bi-modal structure. An example is provided in Figure 1.11 and Figure 1.12.

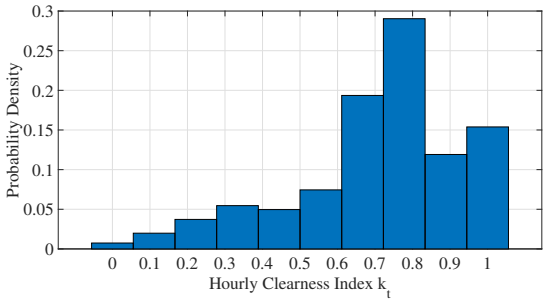

Figure 1.11: Hourly Clearness Index

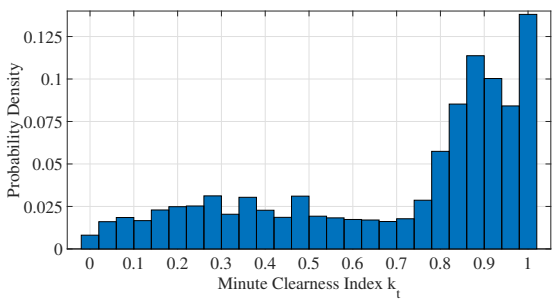

Figure 1.12: 1-Minute Clearness Index

Since the energy generated by a photovoltaic system is a function of the incident irradiance on the photovoltaic panel, the total power produced by a photovoltaic system can also be treated as a random variable. The resulting uncertainty in the photovoltaic generation can be quantified using scenario generation. An example of how the changes in the clearness index impact the electrical power generated by a photovoltaic system is shown in Figure 1.13. The dotted red curve refers to the power generation corresponding to clear sky conditions. The visualization is done based on one-minute values of the clearness index and the photovoltaic generation for a period of seven days. The impact of the cloud movement on the power generated is evident. When the clearness index is at or near unity, the power generation equals the generation for clear sky conditions. For all other values of the clearness index, the power generated (blue curve) is less than the clear sky conditions (red dotted curve). The random behavior of the clearness index results in an uncertain generation that needs to be rigorously modeled to study the impact of such variable generation on the electric grid.

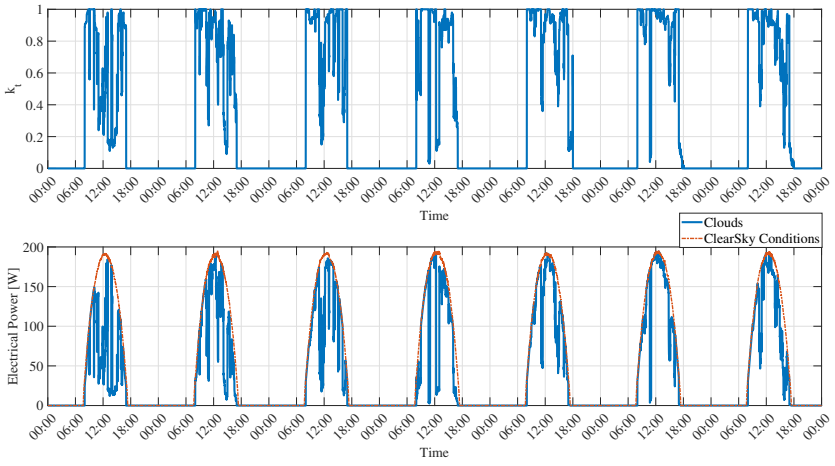

Figure 1.13: 1-Minute Clearness Index and PV Generation

1.3 Topological Uncertainties

Topological uncertainties refer to the network element outages or failures that can happen either due to external factors or gradual aging of the network infrastructure. Furthermore, the increased

adoption of distributed energy resources, particularly solar and wind generation, has led to the phenomenon of accelerated aging of certain mechanically-operated equipment like transformers fitted with on-load tap changers. In today's context, where a substantial amount of renewable generation is being added to the generation pool, topological uncertainties, particularly the accelerated aging and premature failure of equipment, play an important role in the planning and operation of future power grids and thus cannot be ignored.

The increasing penetration of DERs in the distribution networks particularly has led to new challenges not only in the area of operation and planning of distribution grids but also in the reliability of the grid infrastructure. Certain mechanically-operated voltage control devices such as transformers equipped with on-load tap changers and switched capacitor banks are particularly more vulnerable to the rapid changes in solar and/or wind generation. The tap-changing transformers also referred to as voltage regulators, use a line drop compensation circuit to measure the voltage at the load center. Only the forward power flow settings are provided and the voltage regulators are expected to provide regulation at the load center, which is usually the secondary of the transformer, in the range specified by the voltage set point and the bandwidth. Figure 1.14 shows the finer details of a line drop compensation circuit.

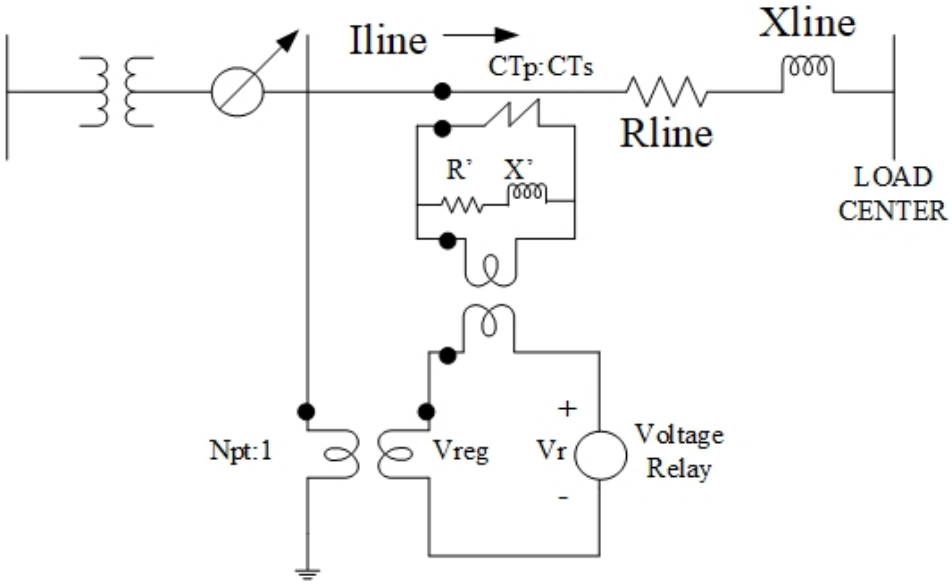

Figure 1.14: Line Drop Compensation Circuit

The parameters R' and X' represent the equivalent impedance from the regulator to the load center. The LDC settings include the potential and current transformer ratios and the parameters R' and X.' These parameters depend on the line impedance between the voltage regulator and load center, and the CT and PT ratio. The transformers with a line drop compensation are circuits that are usually modeled as autotransformers with a nominal voltage regulation capability of $\pm 10\%$. This allows 32 taps with a minimum tap ratio of 0.9 and a maximum tap ratio of 1.1.

The act of changing the transformer taps to maintain the secondary voltage within the specified bandwidth is purely a mechanical operation. The oil-type load tap changers generally require maintenance interventions between 50,000 and 100,000 operations. Due to the mechanical nature of such devices, a continuous control of the secondary voltage is not possible. While such devices can effectively mitigate the slow variations in the voltage (caused primarily by the change in system demand), the stochastic nature of the renewable generation could stretch these devices to the limits of their operation. The rapid excursions in the voltage due to the highly variable output of the solar and/or wind systems cannot be mitigated by legacy voltage regulators utilizing line drop compensation and switched capacitors. The mechanical design of these devices simply renders them inadequate for regulating voltages at shorter time scales of minutes and/or seconds.

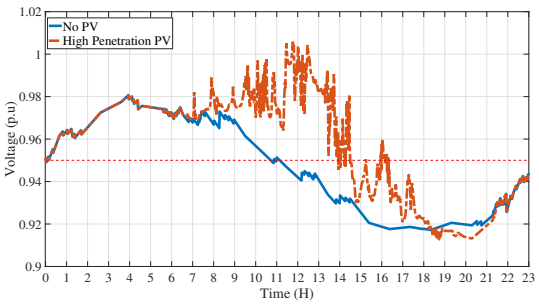

Figure 1.15: Bus Voltage

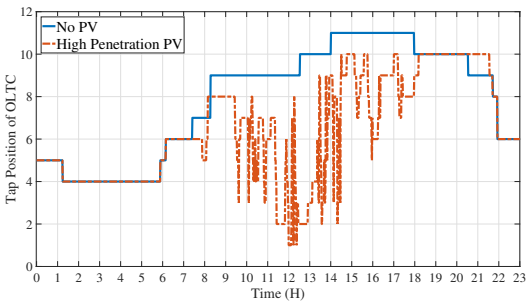

Figure 1.16: Tap Position of Nearest Regulator

An example of device degradation or accelerated aging is shown in Figure 1.15 and Figure 1.16. The red dotted curve in Figure 1.15 represents the voltage under high penetration of solar generation, and the corresponding dotted red curve in Figure 1.16 shows the movement of the tap of the nearest transformer fitted with a load tap changer. It is clear that the transformer changes

tap much more frequently (to control a highly variable voltage) when the distribution network has a high penetration of solar generation. This operational irregularity, a direct consequence of high-capacity renewable generation, can lead to premature failures of such voltage control devices. The tap failure can happen either due to the asynchronous operation of the switches, usually caused by a broken axis, or due to the carbon formation and oxidation of contacts. While a broken axis may be a sudden event, the carbon formation on the contacts represents gradual aging, one that is exacerbated by the intermittent non-scheduled generation [30]. In [31], the impact of PV penetration on a distribution system with a 20 MVA, 69 kV/12,47 kV, delta-wye connected transformer serving two residential feeders was demonstrated. It is shown that a 20% PV penetration amounting to 4 MWs can increase the number of tap-changing operations four times compared with 0% PV penetration. However, the study was performed on global horizontal irradiance data at 1-minute intervals for three days in the summer. The study ignores the effect of the solar PV panel's tilt and azimuth angles on the PV output. Furthermore, the study assumes a lumped load model at the transformer secondary bus, thus ignoring the voltage drop along the feeder. In [32], a yearlong PV data of 1-minute interval was used on a model of a residential network consisting of 60 houses. A 33 kV/11 kV transformer is used to step down the upstream voltage, further decreasing to 415V. The PV data used for modeling has a resolution of 1 minute and accounts for 1.22 MWp in a system with a 12.5 MW base load and 22 MW peak load with an assumed average nominal load of each household to be 3 kW, implying a 10% or less PV penetration. It is shown that with increased variability of the PV output, the number of tap changing operations also increases and the operations saturate at a PV penetration of more than 90%. However, the solar irradiance's dependence on the solar panel's tilt angle and the azimuth is not specified. In [33], the high-frequency solar variability is characterized based on the ramp rate distributions at ten locations across the United States. Weekly simulations with a time resolution of 1 s are carried out on a 12 kV agricultural feeder, and the impact on the voltage regulator operations is studied. Global horizontal irradiance is used to compute the power output of a 3 MW single-axis tracking PV power plant. However, the study is carried out every week and hence is inadequate for capturing the seasonal shifts of the PV output. Moreover,

the study ignores the impact of scattered PV installations across the feeder. In [34], the study tries to assess the variability in the load and solar irradiance from a frequency domain perspective. The study outlines the impact of solar irradiance on the feeder load and the modifications caused to the net load as seen by the utility. Several methods have also been proposed to mitigate the effect of solar variability on the operation of voltage-regulating equipment. The use of on-site battery storage is presented in [35]. In [36], an optimization problem is formulated to minimize the operation of voltage-regulating equipment by dispatching an optimal reactive power control strategy based on the load and the irradiance forecast. The method is tested on an 11kV network with 95 buses equipped with two PV plants with a combined output of 2 MW, operating at 0.95 lead/lag power factor.

In reliability literature, the degradation of equipment, particularly by the device's aging, is consistent with the Weibull distribution [37]. The density and the hazard function of the Weibull distribution have many interesting properties. In particular, the hazard function can assume a variety of shapes. For most aging-related failures, the hazard increases with time, thus increasing the probability of failure given that the device has survived until the present time instant. In a recent survey of the 6057 utility-owned power transformers operating in mainland Australia and Tasmania, the mean life of three transformer populations was determined by applying the Weibull distribution [38]. Previous research in this area has shown that the power transformers' aging strongly depends on the aging of the on-load tap changers. Several surveys over the years suggest that the mechanical failures due to damage in the on-load tap changers account for nearly 55% of the total failures, followed by the dielectric and thermal failures. This is especially true of distribution transformers installed in the substations where it has been shown by empirical data that the majority of outages were a direct result of the tap-changer initiated failures [39] [40]. Further, given the scarcity of the failure data available, especially in the case of devices such as power transformers, several research teams have justified the use of Weibull distribution to fit the relatively small sample sizes of the available data [41], [42],[43]. This is not surprising given that Weibull survival and hazard functions have a lot of interesting properties and offer a lot

of flexibility, unlike other probability distributions such as Lognormal, Gamma, and Exponential distributions. To overcome the challenges of small sample sizes, researchers in [38] analyzed the reliability data from 97% of the 6057-owned utility-owned power transformers in Australia and Tasmania. The researchers use the Weibull distribution to understand failure statistics and conclude that in distribution transformers, aging-related failures dominate after 20 years while the winding-related random failures are more likely to occur before 20 years of age. This is indicated by the shape parameter values of 1.6 before 20 years of age and 3.6 after 20 years of service. Other important works in this area make use of two-parameter Weibull [44], exponential [45] and Perks' distribution [5] to model equipment lifetimes. However, an important limitation of these failure models is their inability to account for failures resulting from the accelerated aging of the devices.

1.4 Contributions

The main contribution of this work lies in the formulation of a stochastic analytical framework that can be used as a tool to assess the profitability of investing in renewable generation at a utility-scale level and power electronics-based solid state devices at a distribution level. For electric utilities, the profitability estimation in renewable generation such as solar and wind farms is a hard problem that is further complicated by economic policies such as carbon tax based on carbon emissions. The methodology presented in this work considers the uncertainties in solar generation and the aggregated system load while accounting for the impact that a certain economic policy may have on the overall cost incurred to the electric utility. Similarly, the proposed stochastic framework can be used to address several operational problems in distribution systems. For example, the economic viability of power electronics-based solutions to challenges such as voltage and harmonic control can be assessed in greater detail and with much higher accuracy.

The other significant contribution of this work lies in developing several computational algorithms based on unsupervised learning and Bayesian statistics to answer some of the most challenging problems in distribution systems. Reliability challenges in power systems have been gaining much traction in recent years. This is mostly a result of the rapid growth of distributed energy resources, particularly solar generation. This work studies the reliability issues related to the me-

chanical devices in the distribution networks, such as voltage regulators and switched capacitor banks. A major contribution in this area is developing a predictive failure model that incorporates the impact of the distributed generation. The electric utilities can use the failure model as a tool to assess the reliability of the distribution system infrastructure and make predictions about either imminent or long-term failures.

1.5 Outline of the Dissertation

The dissertation is organized as follows

Chapter 2 This chapter deals with the problem of statistical load modeling. The applicability of several probability mixture models based on the generalized density functions is investigated. In addition, the chapter also contains a detailed description of the mathematical modeling of Plug-in Hybrid electric vehicle charging.

Chapter 3 This chapter focuses on the uncertainty characterization of distributed generation in the form of solar generation. Two different models for uncertainty characterization are presented. The first model is based on the Markov Chain Monte Carlo method, while the second is a time series auto-regressive moving average model. Both models can generate synthetic scenarios with the same statistical properties as the empirical data.

Chapter 4 This chapter leverages the computational models of the earlier chapters and presents a case study on the profitability of investing in utility-scale solar generation.

Chapter 5 This chapter contains a comprehensive description of topological uncertainties in power systems. A failure model based on Bayesian statistics is presented. Furthermore, a detailed mathematical analysis of the parameter estimation problem is provided.

Chapter 6 This chapter leverages the computational models of the earlier chapters to develop the cost-benefit model of a power electronics-based solid state device for use in distribution systems.

Chapter 7 The final chapter provides a summary of the work done and offers several conclusions as well as possible future directions for this work.

2. STATISTICAL MODELING OF LOAD

The emergence of distributed energy resources has led to new challenges in the operation and planning of power networks. Of particular significance is the introduction of a new layer of complexity that manifests in the form of new uncertainties that could severely limit the resiliency and reliability of modern power networks. For example, the increasing adoption of unconventional loads, such as Plug-in Hybrid Electric Vehicles, can result in uncertain consumer demand patterns, often characterized by random undesirable peaks in energy consumption. Traditionally, load uncertainty arises due to environmental factors, appliance variations, and consumer behavior. However, the embedded uncertainties associated with using EVs/PHEVs further accentuate the randomness in the electric demand profiles. Even in the absence of DER-induced uncertainties, it has been shown that power system demand patterns show a lot of variabilities when measured at different buses across the network. This is especially true for distribution systems where different load types follow dissimilar patterns of use and, as such, cannot be modeled by any specific probability distribution function.

The use of unconventional loads such as PHEVs and/or EVs impacts feeder load profiles. It presents considerable security and reliability challenges to the normal operation of a power distribution grid. The charging of EVs, for instance, disrupts the normal feeder load profiles, which, if left unaccounted for, could lead to undesirable peaks in electrical energy consumption. From the distribution system operator's point of view, some of the various factors of concern with the EV integration, in addition to demand patterns with a pronounced 'peaky' behavior, include feeder power losses, feeder voltage profile, voltage unbalance, injection of harmonics into the grid, violation of thermal line and transformer limits, load forecasting challenges due to the introduction of new customer demand patterns, impact on the tap changing under load transformers and other mechanical switched devices in distribution grids, etc. The system is most vulnerable to severe impacts, component overloading, and excessive voltage drops during charging periods that coincide with the normal peak electrical energy consumption. In several studies examining the most

damaging effects of EV charging, it has been concluded that existing distribution grids should accommodate substantial EV penetration, provided that a major chunk of EV charging is restricted to low charging rates at off-peak times. Furthermore, uncoordinated charging, particularly fast three-phase charging, can increase the instances of thermal violations in cables and transformers. The impact could exacerbate if it coincides with the peak energy consumption [46][47].

In this chapter, we will explore the applicability of generalized mixture models for statistical modeling of aggregated load in distribution systems enhanced with high capacity renewable energy systems in the form of distributed solar generation and unconventional load types such as uncoordinated charging of several plug-in electric vehicles. To my knowledge, the generalized mixture models have not been considered for statistical load modeling of aggregated load. The generalized mixture models considered in this work include the generalized Gaussian mixture model (GGMM), which is based on generalized Gaussian distribution (GGD), and the generalized Beta mixture model, which is based on generalized Beta Prime distribution (GB2). We also consider the multivariate version of the GGMM to study the correlation among adjacent loads. The GGMM has been used in the literature to address various issues in pattern recognition. For example, the study in [Image thresholding based on EM algorithm and GGD] proposes the use of a GGMM for global image thresholding based on the assumption that the statistical parameters of "object" and "background" can be modeled by a Generalized Gaussian distribution (GGD). The parameters of the resulting mixture model are obtained using the E-M algorithm. The study [48] considers the parameter estimation problem of a multivariate generalized Gaussian distribution using the fixed point (FP) method [49]. The paper establishes some properties related to the FP equation and an algorithm based on an iterative procedure like Newton-Raphson is used to compute the maximum likelihood estimate of the real scatter matrix of the multivariate GGD. The contributions of this chapter are as follows

- This paper explores the properties of generalized mixture models such as the generalized Gaussian mixture model (GGMM) and generalized Beta mixture model (GBMM) for statistical modeling of aggregated load in systems with high capacity distributed energy resources

(DERs) such as distributed solar generation and unconventional loads such electric vehicles (EVs).

- The electric vehicle charging load is estimated by modeling the EV connection with the electric grid as a stochastic counting process based on a nonhomogeneous Poisson process. The EV arrival times are simulated by implementing a version of the acceptance-rejection-based algorithm called "thinning".

- The mixture components of the generalized models are obtained using the Expectation-Maximization algorithm. The performance assessment of the generalized mixture models considered in this work is done through various goodness-of-fit measures such as Mean Square Error (MSE), Root Mean Square Error (RMSE), and Normalized Root Mean Square Error (NRMSE), and Kullback-Lielber divergence (K-L).

- A procedure for random sampling from the proposed mixture models is presented. The sampling is based on the idea of rejection sampling which involves the use of a proposal density function to calculate the acceptance probability of the sampled values.

2.1 Aggregate Load Behavior*

[1] The distribution system considered in this work is the IEEE-123 bus test feeder. Based on the peak loading data of the test system [50], the following consumer classes are identified for load profile calculations. These are a) Residential loads (R), b) commercial loads (C), and c) industrial loads (I). The base residential load profile data is obtained from [51], which is based on the 2009 residential energy consumption survey (RECS) data set. The hourly base load profiles for

[1]Part of the data reported in this chapter is reprinted with permission from "On Statistical Modeling of Load in Systems with High Capacity Distributed Energy Resources" by Aaqib Peerzada, Miroslav Begovic, Wesam Rohouma, Robert Balog,arXiv preprint arXiv:2207.11355,2022

commercial and industrial loads are obtained from [OpenEI] and normalized with respect to the peak values. The power factors considered for the three classes of consumers are 0.99, 0.98, and 0.9 lagging, respectively.

The presence of DERs, such as renewable energy in the form of solar generation and EV charging, significantly impacts the aggregate load. To observe the impact on the aggregate load, it is important to have reliable forecasts of EV energy consumption and renewable energy generation. The EV energy consumption further depends on the frequency with which the EVs connect to the grid at a charging location. In the literature on EV load modeling, the queuing theory has often been used to model the EV connection. The study in [] performed a null hypothesis test on vehicle travel patterns based on the National Household Travel Survey (NHTS) data. It is shown that the electric vehicle arrivals for charging can be modeled as a constant-rate Poisson process in short intervals of time (30 minutes). To the best of our knowledge, the methods proposed in the literature to forecast the EV demand do not provide any information on the generation of the nonhomogeneous Poisson process used to estimate the arrival times of the EVs. This chapter presents a procedure to generate a nonhomogeneous Poisson process with the desired intensity function. The simulated nonhomogeneous Poisson process is used to estimate the arrival times of the EVs.

2.1.1 EV Connection as an Arrival Process

We consider a stochastic counting process $\{N_{EV}(t); t \geq 0\}$ defined on a sample space Ω. The random variable $N_{EV}(t)$ is the realization of the number of EV arrivals in the interval $[0,t]$. For any $\tau \geq t$, the counting process $\{N_{EV}(t); t \geq 0\}$ has the property $N_{EV}(\tau) \geq N_{EV}(t)$. This means that $N_{EV}(\tau) - N_{EV}(t)$ is a nonnegative random variable. By definition, the stochastic counting process $\{N_{EV}(t); t > 0\}$ is integer-valued, non-decreasing, and right continuous. The counting process $\{N_{EV}(t); t > 0\}$ can be characterized either by the sequence of the interarrival times $X_1, X_2, X_3, ...$ or by the number of arrivals $N_{EV}(t)$ in the interval $[0,t]$. The k^{th} arrival epoch (time) can be specified in terms of the interarrival times

$$S_k = \sum_{i=1}^{k} X_i \tag{2.1}$$

In a similar way, the interarrival times can be specified by the arrival epochs (times) with $X_k = S_k - S_{k-1}$. For a given k^{th} arrival such that $k \geq 1$ and time $t > 0$, the k^{th} arrival epoch S_k is related to the counting random variable $N_{EV}(t)$

$$\{S_k \leq t\} = \{N_{EV}(t) \geq k\} \tag{2.2}$$

This relationship can be easily verified by considering that the number of arrivals by time t must be at least k if the k^{th} arrival epoch (time) S_k happens before or at time t. Conversely, it is also true that

$$\{S_k > t\} = \{N_{EV}(t) < k\} \tag{2.3}$$

Equation (4.7) holds because if the k^{th} arrival epoch occurs at $\tau > t$ then the number of arrivals up to and including t must be less than k.The counting process $\{N_{EV}(t); t > 0\}$ is a non-homogeneous Poisson process with a time-varying arrival rate of $\lambda(t)$. In addition, $\forall t \geq 0$ and $\delta > 0$, $\{N_{EV}(t); t > 0\}$ satisfies [52]

$$\begin{aligned} \Pr\{\tilde{N}_{EV}(t,t+\delta) = 0\} &= 1 - \lambda(t)\delta + o(\delta^2) \\ \Pr\{\tilde{N}_{EV}(t,t+\delta) = 1\} &= \lambda(t)\delta + o(\delta^2) \\ \Pr\{\tilde{N}_{EV}(t,t+\delta) \geq 2\} &= o(\delta^2) \end{aligned} \tag{2.4}$$

In (4.8), $\tilde{N}_{EV}(t,t+\delta) = N_{EV}(t+\delta) - N_{EV}(t)$. The non-homogeneous Poisson process (NHPP) defined in (4.8) has the independent increment property but does not have the stationary increment property. $\lambda(t)$ is also called the rate function of the NHPP. The NHPP (4.8) is characterized by the mean value function $\wedge(t) \equiv \mathbb{E}[N_{EV}(t)]$. The mean value function in terms of the intensity function $\lambda(t)$ is

$$\wedge(t) = \int_0^t \lambda(y)dy < \infty \tag{2.5}$$

For a NHPP, the probability of having k arrivals in the interval $[0,t]$ is given by [reference]

$$\Pr\{N_{EV}(t) - N_{EV}(0) = k\} = \frac{[\wedge(t) - \wedge(0)]^k}{k!} \exp(-[\wedge(t) - \wedge(0)]) \tag{2.6}$$

2.1.2 Simulating Arrival times of an NHPP

The interarrival times of an NHPP are not independent and are not exponentially distributed. The distribution of the interarrival times can be obtained by using the independent increment property of an NHPP. More specifically, the independent increment property states that the number of EV arrivals $N_{EV}(s)$ is independent of $N_{EV}(s+t)-N_{EV}(s)\forall s,t \geq 0$. In other words, we consider the random variables $\{N_{EV}(I_n); 1 \leq n \leq k\}$ to be independent where $\{I_n\}_{1\leq n\leq k}$ are piecewise disjoint intervals. The CDF of $X_k = S_{k+1} - S_k$ conditional on the first k arrival epochs $S_1 = s_1, S_2 = s_2, ..., S_k = s_k$ is

$$\begin{aligned} F_{X_k}(x) &= \Pr\{X_k \leq x | S_i = s_i, i = 1,2,...,k\} \\ F_{X_k}(x) &= \Pr\{N(S_k+x) - N(S_k) \geq 1 | S_k = \sum_{i=1}^{k} X_i\} \end{aligned} \tag{2.7}$$

The independent increment property of the NHPP implies that $N_{EV}(S_k+x)-N_{EV}(S_k)$ is independent of $N_{EV}(S_k)$. Using this property we get

$$F_{X_k}(x) = 1 - \Pr\{N(S_k+x) - N(S_k) = 0\} \tag{2.8}$$

The probability of having 0 arrivals in the interval S_k, S_k+x is $\exp -[\wedge(S_k+x) - \wedge(S_k)]$. The cdf of the interarrival times is

$$F_{X_k}(x) = 1 - \exp\left(-\wedge(S_k+x) + \wedge(S_k)\right) \tag{2.9}$$

Furthermore, the interarrival times $X_1, X_2, ..., X_k$ are conditionally increasing if and only if the intensity function of the NHPP is decreasing. The proof is given in [Subash C Kochar Some results on interarrival times of NHPP].

An NHPP can be generated from a homogeneous Poisson process by considering a constant intensity function λ^+ that dominates the time-varying intensity function $\lambda(t); t \geq 0$ of the desired NHPP such that $\lambda^+ \geq \lambda(t)\ \forall t \in [0,T]$. A variation of the acceptance-rejection called "thinning" is

used to sample from the generated events of a homogeneous Poisson process such that the desired intensity function $\lambda(t)$ is achieved [53]. The thinning algorithm is based on the following theorem

Theorem 1 (Lewis and Shedler, 1979) [54] Consider a non-homogeneous Poisson process with intensity function $\lambda_v(t), t \geq 0$. Suppose that $S_1^*, S_2^*, ..., S_k^*$ are random variables representing event times from the non-homogeneous Poisson process with intensity function $\lambda(t)$ and lying in the fixed interval $(0,t]$. Let $\lambda(t)$ be a intensity function such that $0 \leq \lambda(t) \leq \lambda_v(t) \forall t \in [0,t]$. If the i^{th} events is independently deleted with probability $1 - \lambda(t)/\lambda_v(t)$, the remaining event times form a nonhomogeneous Poisson process with intensity function $\lambda(t)$ in the interval $(0,t]$. The proof is given in [55]

The thinning algorithm used in the paper to simulate a nonhomogeneous Poisson process is implemented as follows. We consider $\lambda(t)$ to be the intensity function of the NHPP over a fixed interval $[0,T]$.

- Simulate a homogeneous Poisson process (HPP) with constant intensity function $\lambda^+ \geq \lambda(t) \forall t \in [0,T]$ by drawing uniform random numbers $\{u_k; k = 1,2,...,k^*\} \sim U(0,1)$. Since the inter-arrival times are exponentially distributed in a HPP, the inter-arrival times are obtained by setting $X_k = -\frac{1}{\lambda^+} \log u_k$.

- The arrival times of an HPP are obtained by setting $S_k = S_{k-1} - \frac{1}{\lambda^+} \log u_k$. The total number of uniform random numbers drawn is $k^* = \max\{k; \sum_{n=1}^{k} S_n < T\}$.

- Independently generate uniform random numbers $\{w_j; j = 1,2,...,k^*\} \sim U(0,1)$ and calculate the indicator function

$$I_j = \begin{cases} 1; w_j \leq \frac{\lambda(S_j)}{\lambda^+} \\ 0; w_j > \frac{\lambda(S_j)}{\lambda^+} \end{cases} \tag{2.10}$$

- Form the set $\mathbf{J} = \{I_j; j = 1\}$ and the arrival times of the NHPP are $S_j^{NHPP} = \{S_j; j \in \mathbf{J}\}$

In this work, we consider a piecewise constant intensity function for intervals lasting 30 minutes. The piecewise constant intensity function depends on customer convenience and is based on real-

Algorithm 1: Acceptance Rejection based Thinning of HPP

Input:NHPP Intensity function $\lambda(t)$, HPP Constant-rate Intensity function λ^+, Interval length $[0,T]$)
Set $\lambda^+ = \max\{\lambda(t)\}$
Set the counting process $N_{EV} = a\lambda^+ T, a > 1$
while $i \leq N_{EV}$ **do**
  Draw $u \sim U(0,1)$
  Set $S_k = -\frac{1}{\lambda^+}\log u_k$
end while
Set $\mathbf{S} = \mathbf{S}(\mathbf{S} < N_{EV})$
Set $k^* = \max\{k; \sum_{n=1}^{k} S_n < T\}$
for all i=1,..,k **do**
  Draw $w \sim U(0,1)$
  Calculate acceptance probability, $r(j) = \frac{\lambda(S_j)}{\lambda^+}$
  if $w_j \leq r(j)$ **then**
    $\mathbf{I}_j = 1$
  else
    $\mathbf{I}_j = 0$
  end if
end for
$\mathbf{J} = \{\mathbf{I}_j; j = 1\}$
$S_j^{NHPP} = \{S_j; j \in \mathbf{J}\}$
Output: Arrival Times of NHPP

world data given in [3]. For a typical day, the piecewise constant arrival rate is given in Table **??**

Table 2.1: Arrival Rate of EV Connection [3]

Time (Mins)	% of Vehicles on Road	EVs/hour
$0 \leq t < 360$	≤ 4	[1,2,3,4]
$360 \leq t < 480$	>4 and ≤ 7	[5,6,7,8,9,10,11]
$480 \leq t < 780$	≤ 4	[1,2,3,4]
$780 \leq t < 1080$	≥ 7	[12,13,14,15,16,17]
$1080 \leq t < 1200$	> 4 and ≤ 7	[5,6,7,8,9,10,11]
$1200 \leq t \leq 1440$	≤ 4	[1,2,3,4]

To generate an NHPP with the desired intensity function $\lambda(t)$ as given in Table 2.1 we consider a homogeneous Poisson process (HPP) with a constant rate intensity function λ^+ such that $\lambda^+ = \max \lambda(t) \forall t \in [0,T]$. The EV arrival times corresponding to the intensity function $\lambda(t)$ are then obtained by applying the thinning algorithm to the HPP with the intensity function λ^+. The number of arrivals $N_{EV}(t)$ as a function of the arrival epochs S_k and the piecewise constant intensity function of Table **??** are shown in Fig 2.1 and Fig 2.2 respectively.

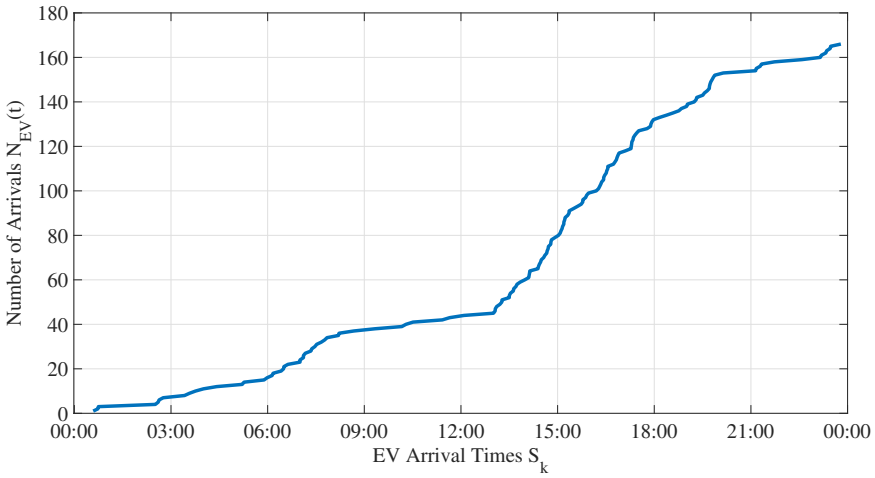

Figure 2.1: NHPP Count Trajectory

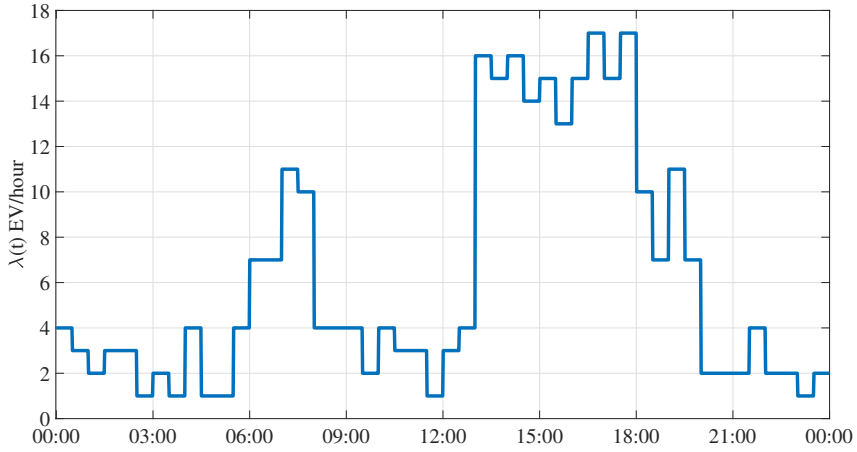

Figure 2.2: Time-varying Arrival Rate

2.1.3 EV Consumption Profile

The EV charging load can be calculated based on the expected number of arrivals in a fixed interval $[0, T]$. The expected number of arrivals can be estimated using the information about the time-varying intensity function of the counting process assumed to model the EV arrivals and the arrival times. As a first approximation, we consider a commercial building with an EV charging

facility. The EVs arrive at this charging facility with an intensity function consistent with that of an NHPP to receive charging service. We also consider homogeneous EVs and use the battery parameters of the Tesla Model 3 which is often regarded as the most popular EV in North America [56].

The daily driven miles can be modeled by a log-normal distribution with a mean of 3.37 and a standard deviation of 0.5 [46]. The battery state of charge (SoC) can be estimated from the daily driven miles (d), electricity consumption in kWh/100 miles (E_{cons}), and battery capacity (C_b) in kWh.

$$SoC_{arrival} = 1 - \frac{E_{cons} d}{C_b} \times 100 \tag{2.11}$$

The energy required to charge the battery to the desired SoC which in this case is 100% is given by

$$E_{req} = \frac{SoC_{final} - SoC_{arrival}}{\eta \times 100} C_b \tag{2.12}$$

η in (2.12) is the charging efficiency assumed to be 95%. The total charging time is given by

$$T_{ch} = E_{req}/P \tag{2.13}$$

P in (2.13) is the charging rate. In this case we assume 3-ϕ L-2 charging. Table 2.2 gives the battery-related parameters of Tesla Model 3 considered in this work.

Table 2.2: EV Battery Characteristics

Parameter	Value
C_b(kWhRated)	75 kWh
P(kWRated)	11.5 kW ; 48 Amps (3-ϕ)
Efficiency η	95%
% Charging	100% of kWRated
E_{cons}	27 kWh/100 miles

The total EV demand is calculated based on the EV arrival rate as shown in Fig 2.3, the expected

number of events of the EV counting process $N_{EV}(t)$ in the interval $(0,t)$, the energy required to charge the EVs (2.12) and the total charging time given by (2.13). Based on the arrival rate shown in Fig 2.2, the total EV electric demand for one day is shown in Fig 2.3

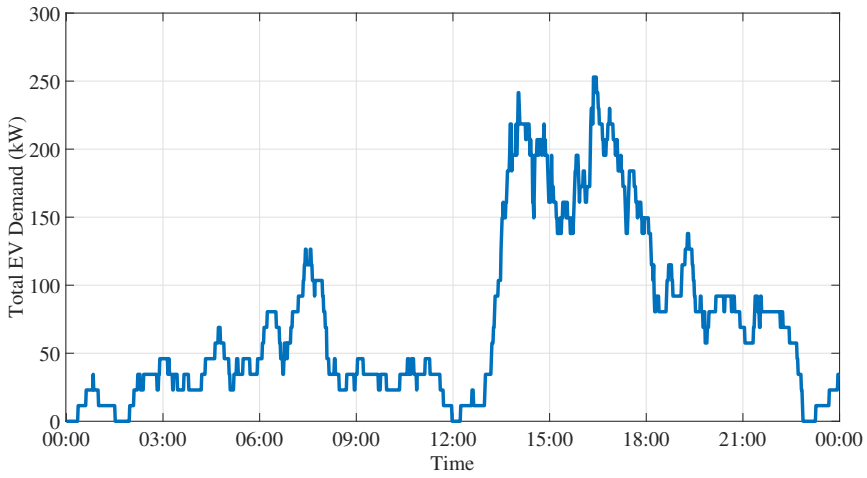

Figure 2.3: EV Aggregate Demand

2.1.4 Aggregate Load with EV charging and Distributed Generation

The stochastic nature of EV charging and distributed generation significantly impact the overall electric demand profile. In the case of EV charging, the demand profile of the charging facility will exhibit a pronounced peaky behavior with peaks in demand occurring at random times, which depend on the arrival times of the EVs. Combined charging could alter the diversified electric demand seen from the secondary distribution substation transformer if enough EV charging facilities are considered. Similarly, distributed generation in the form of distributed roof-top PV systems can lower the diversified demand when the PV systems are generating peak power, followed by the subsequent ramping up of the load as the output of the PV systems dwindles. However, the PV system output can experience significant variability depending on the weather conditions, leading to a highly variable demand profile at the point of common coupling. Suppose the PV systems are dispersed geographically over sufficiently long distances. In that case, the variability in the combined output and hence the variability in the diversified demand may be less than the outputs

of the individual PV systems. However, this mitigation of the PV variability will depend on the measure of similarity in the generation patterns, which further depends on the spatial spread of the distributed PV systems.

In this work, we assume that the distribution test feeder has enough capacity to support the EV charging and that all the physical power system constraints are satisfied. The histogram plot of the one-day aggregated load of a commercial building considering EV charging is shown in Fig 2.4 . Also, Fig 2.5 shows the histogram plot of the weekly diversified demand as seen from the secondary substation transformer in the IEEE 123 feeder with and without distributed generation.

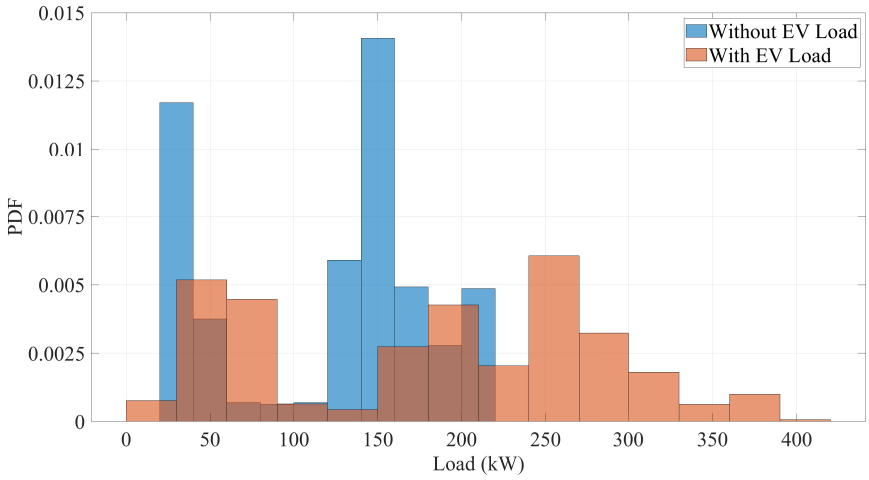

Figure 2.4: PDF of one day aggregated load with and without EV charging

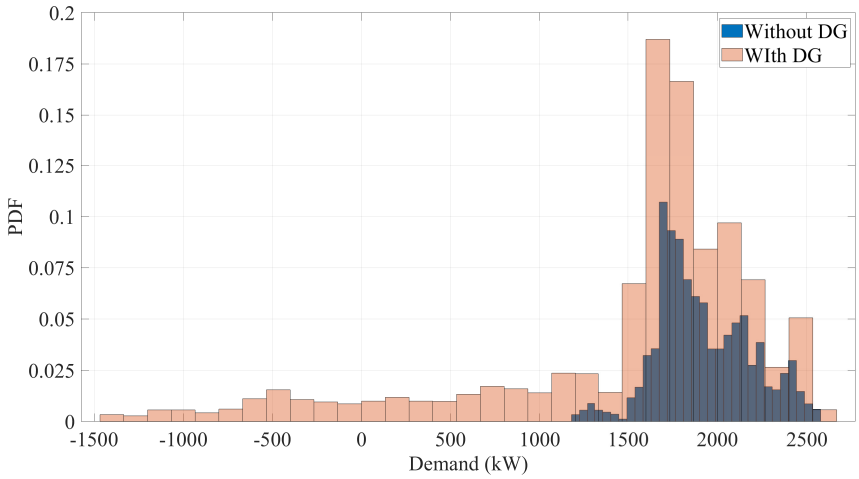

Figure 2.5: PDF of Diversified Demand with and without Distributed Generation

It is clear from Fig 2.4 and Fig 2.5 that EV charging and distributed generation significantly impact electric demand. While EV charging disrupts the pattern of energy consumption by introducing random peaks that correspond to the times when an EV either connects or disconnects with the grid, high-capacity distributed generators can lead to reverse power flow along the feeder, as indicated by the negative values of the electric demand in Fig 2.5. In this case, we consider a high penetration scenario of distributed solar generation that amounts to approximately 90% of the total peak demand of the IEEE 123 test feeder.

2.2 Generalized Mixture Models for Load Modeling*

[2] A generalized probability mixture model is based on the generalized version of any specific density function. For example, a generalized Gaussian mixture model (GGMM) comprises a weighted convex combination of a finite number of generalized Gaussian density functions. Generalized distributions often employ more parameters to achieve a greater degree of flexibility. For instance, a generalized Gaussian density requires estimating one additional parameter than a Gaussian distribution. This parameter, called the shape parameter, controls the shape of the distribution with larger values resulting in a distribution with smaller tails and vice versa. Similarly, the gen-

[2]Part of the data reported in this chapter is reprinted with permission from "On Statistical Modeling of Load in Systems with High Capacity Distributed Energy Resources" by Aaqib Peerzada, Miroslav Begovic, Wesam Rohouma, Robert Balog,arXiv preprint arXiv:2207.11355,2022

eralized beta prime distribution, also referred to as the generalized beta distribution of the second kind has two additional parameters that control the shape and the scale of the distribution. Considering the impact of DERs such as EVs and distributed generation in the context of load modeling, a generalized mixture model, owing to its greater flexibility, can result in a better fit of the empirical load data. In the next section, we will consider two generalized mixture models based on generalized Gaussian and generalized beta prime distributions, respectively. To account for the load correlation at buses close to each other, we also present a multivariate version of the GGMM.

2.2.1 Parameter Estimation of Mixture Models

A probability mixture model, in general, is a convex combination of a finite number of density functions. It is characterized by the total number of components considered in the mixture, the weight of each component, and a parameter vector that consists of parameters of the component densities. In general, a mixture model that is a sum of a finite number of density functions has the following form

$$f_Y(y_i;\Psi) = \sum_{j=1}^{M} \pi_j p_{j,Y}(y_i, C = j, \Theta_j) \tag{2.14}$$

In the context of load modeling, $Y = y_i$ is the measured load data, M is the number of mixture components and $\pi_j = p(y_i \in C_j)$ is the weight of the j^{th} component density. The weights assigned to the component densities are subject to following constraints.

$$\begin{aligned} &\pi_j > 0 \; \forall \; j \in \{1,2,...,M\} \\ &\sum_{j=1}^{M} \pi_j = 1 \end{aligned} \tag{2.15}$$

Each component density function $p_{j,Y}$ is characterized by the parameter vector Θ_j and $\Psi = \{\pi_j, \Theta_j; j = [1,2,...,M]\}$ is the overall parameter vector. If $\{y_i\}_{i=1}^{N}$ is the measured data vector, then the likelihood function of the model in (4.3) has the form

$$\mathbb{L}(\Psi|y_i) = \prod_{i=1}^{N} \sum_{j=1}^{M} \pi_j p_{j,Y}(y_i, C = j, \Theta_j) \tag{2.16}$$

If we take the logarithm on both the sides of (2.16) we will get the log-likelihood function of the mixture model.

$$\log L(\Psi|y_i) = \sum_{i=1}^{N} \log \sum_{j=1}^{M} \pi_j p_{j,Y}(y_i|C=j,\Theta_j) \tag{2.17}$$

The overall parameter vector Ψ that characterizes the mixture model (4.3) can be obtained by maximizing (2.17).

$$\Psi^* = \arg\max_{\Psi} \sum_{i=1}^{N} \log \sum_{j=1}^{M} \pi_j p_{j,Y}(y_i|C=j,\Theta_j) \tag{2.18}$$

The optimization problem (2.18) is ill-posed and a solution can be obtained by using Expectation-Maximization (E-M) algorithm. The E-M algorithm find a solution by reinterpreting the measured data $\mathbf{Y}$ as incomplete data and positing the existence of a $M-$ dimensional binary variable $\mathbf{Z} = \{z_{ji}\}_{i=j=1}^{i=N,j=M}$ for each sample y_i such that the j^{th} component z_{ji} is 1 if and only if y_i belongs to the j^{th} component. That is

$$z_{ji} = \begin{cases} 1, y_i \in C_j \\ 0; y_i \notin C_j \end{cases} \tag{2.19}$$

Let $\mathbf{X} = [\mathbf{Y}^T, \mathbf{Z}^T]^T$ be the complete data vector. The complete data log-likelihood is given by

$$\log L(\Psi|\mathbf{Y},\mathbf{Z}) = \log P(\mathbf{Y},\mathbf{Z}|\Psi) \tag{2.20}$$

Using the chain rule we get

$$P(\mathbf{Y},\mathbf{Z}|\Psi) = P(\mathbf{Y}|\Psi)P(\mathbf{Z}|\mathbf{Y},\Psi) \tag{2.21}$$

Substituting (2.21) in (2.20) and using (4.3), the complete data log-likelihood has the following form

$$\log L(\Psi|\mathbf{Y},\mathbf{Z}) = \sum_{i=1}^{N}\sum_{j=1}^{M} z_{ji}\{\log \pi_j + \log(p_{j,Y}(y_i|C=j,\Theta_j))\} \tag{2.22}$$

A challenge in estimating (2.22) is that the assumed binary variable $\mathbf{Z}$ is an unknown. Further progress can be made if $\mathbf{Z}$ is assumed to be a random vector. In that case, we can derive an expression for the distribution of the unobserved data $\mathbf{Z}$. Suppose $\Psi^{(g)}$ is the appropriate parameter

vector for the mixture density. Given $\Psi^{(g)}$, the component density functions $p_j(y_i|\Psi^{(g)})$ can be evaluated $\forall j,i$. The weights π_j's can be considered prior probabilities of each mixture component. The posterior distribution of the unobserved data $\mathbf{Z}$ conditional on the observed data $\mathbf{Y}$ and the current parameter estimates $\Psi^{(g)}$ can be obtained using Bayes' rule.

$$P(z_{ji}|y_i,\Psi^{(g)}) = \frac{P(y_i|z_{ji},\Psi^{(g)})\pi_j}{P(y_i|\Psi^{(g)})} \tag{2.23}$$

The term in the denominator of (2.23) is the total probability of observing the measured data and is equal to the density function of the mixture model i.e, $P(y_i|\Psi^{(g)}) = \sum_{j=1}^{M} p_j(y_i|\Psi^{(g)})$. The joint probability of $\mathbf{Z}$ is

$$P(\mathbf{Z}|\mathbf{Y},\Psi^{(g)}) = \prod_{i=1}^{N} p(z_{ji}|y_i,\Psi^{(g)}) \tag{2.24}$$

The E-M algorithm is an iterative procedure and consists of two steps. In the E-step (first step) of the algorithm, we determine the missing information ($\mathbf{Z}$) by computing the conditional expectation of $\mathbf{Z}$ based on the current parameter estimate $\Psi^{(g)}$ and the observed data $\mathbf{Y}$.

E-Step

The conditional expectation of the unobserved binary valued data $\mathbf{Z}$

$$\mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}] = P(z_{ji}|y_i,\Psi^{(g)}) = \frac{P(y_i|z_{ji},\Psi^{(g)})\pi_j}{\sum_{j=1}^{M} p_j(y_i|\Psi^{(g)})} \tag{2.25}$$

Equation (2.25) can be used to calculate the conditional expectation of the complete date log-likelihood function.

$$Q\left(\Psi^{(g+1)},\Psi^{(g)}\right) = \mathbb{E}\left[\log L(\Psi|\mathbf{Y},\mathbf{Z})|\mathbf{Y},\Psi^{(g)}\right] \tag{2.26}$$

To solve this further, we get the conditional expectation of the complete data log-likelihood function

$$Q\left(\Psi^{(g+1)},\Psi^{(g)}\right) = \sum_{i=1}^{N}\sum_{j=1}^{M} \mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}][\log\pi_j + \log(p_{j,Y}(y_i|C=j,\Theta_j))] \tag{2.27}$$

M-Step

In the M-step the conditional expectation obtained in the E-step is maximized by varying the value of Ψ. The sub-problem is

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{N}\sum_{j=1}^{M} \mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}][\log\pi_j + \log(p_{j,Y}(y_i|C=j,\Theta_j))] \\ \text{subject to} \quad & \sum_{g=1}^{M}\pi_g = 1 \\ & \pi_j \geq 0 \, j \in [1,2,...,M] \end{aligned} \tag{2.28}$$

The optimization problem in (4.4) is solved using the method of Lagrange multipliers.

$$\frac{\partial}{\partial\pi_k}\sum_{i=1}^{N}\sum_{j=1}^{M} \mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}][\log\pi_j + \log(p_{j,Y}(y_i|C=j,\Theta_j))] + \lambda\left(\sum_{j=1}^{M}\pi_j - 1\right) = 0 \tag{2.29}$$

This above expression yields

$$\pi_k = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}] \tag{2.30}$$

Maximizing with respect to Θ_j

$$0 = \sum_{i=1}^{N}\mathbb{E}[z_{ji}|\mathbf{Y},\Psi^{(g)}]\frac{\partial}{\partial\Theta_k}\log p_k(y_i|C=k,\Theta_k) \tag{2.31}$$

2.2.2 Univariate Generalized Mixture Models

We consider the univariate mixture models to model the load distribution at a single bus. However, this loses any information about the correlation of loads in close proximity. At the same time, univariate models require less computational effort and can be used when the information about load correlation is not sought or important. We consider two univariate mixture models based on generalized Gaussian and beta prime distributions.

2.2.2.1 Generalized Gaussian Mixture Model

A generalized Gaussian mixture density is a weighted sum of a finite number of generalized Gaussian density functions. Due to an additional parameter, the generalized Gaussian density offers more flexibility in modeling a large variety of statistical behaviors. Also, the generalized Gaussian density function can approximate a large class of continuous density functions. The generalized Gaussian density adopted in this paper has the following form

$$p_Y(y;\mu,s,p) = \frac{\beta}{2s\Gamma\left(\frac{1}{\beta}\right)} \exp\left[-\frac{|y-\mu|^{\beta}}{s^{\beta}}\right] \tag{2.32}$$

In 4.2, $\beta \in \mathbb{R}^+$ is the shape parameter,$s \in \mathbb{R}^+$ is the scale parameter, $\mu \in \mathbb{R}$ is the location parameter and $\Gamma(.)$ is the gamma function. It is worth noting that $\beta = 1$ gives a Laplace distribution, and $\beta = 2$ gives a Gaussian distribution. A generalized Gaussian mixture density is characterized by the total number of mixture components, weight, location, scale, and shape of each component density.

The conditional expectation of the complete data log-likelihood function for the GGMM can be obtained by substituting (4.2) in (2.27). This results in

$$Q\left(\Psi^{(g+1)},\Psi^{(g)}\right) = \sum_{i=1}^{N}\sum_{j=1}^{M}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{k}\right]\log\pi_j + \sum_{i=1}^{N}\sum_{j=1}^{M} \left(\log\beta_j - \log 2 - \log s_j - \log\Gamma\left(\frac{1}{\beta_j}\right) - s_j^{-\beta_j}|y_i-\mu_j|^{\beta_j}\right) \tag{2.33}$$

The application of (4.4)-(2.31) yields the following update equations for the model parameters $\Psi^{(g+1)} = \{\pi_j,\mu_j,s_j,\beta_j; j=[1,2,...,M]\}$ of a GGMM given the current estimate $\Psi^{(g)}$ and the measured load data $\mathbf{Y}$

$$\pi_j^{(g+1)} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right] \tag{2.34}$$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\beta_j^{(g)}|\mu_j^{(g+1)} - y_i|^{\beta_j^{(g)}} = 0 \tag{2.35}$$

$$s_j^{(g+1)} = \left[\frac{\sum_{i=1}^{N} \mathbb{E}\left[z_{ji} | \mathbf{Y}, \Psi^{(g)} \right]}{\sum_{i=1}^{N} \mathbb{E}\left[z_{ji} | \mathbf{Y}, \Psi^{(g)} \right] \beta_j^{(g)} |\mu_j^{(g)} - y_i|^{\beta_j^{(g)}}} \right]^{-\frac{1}{\beta_j^{(g)}}} \tag{2.36}$$

$$\sum_{i=1}^{N} \mathbb{E}\left[z_{ji} | \mathbf{Y}, \Psi^{(g)} \right] \kappa = 0 \tag{2.37}$$

In (5.34), κ equals

$$\kappa = \frac{1}{\beta_j^{(g+1)}} + \frac{\psi\left(1/\beta_j^{(g+1)}\right)}{\left(\beta_j^{(g+1)}\right)^2} - \left(\frac{|y_i - \mu_j^{(g)}|}{s_j^{(g)}} \right)^{\beta_j^{(g+1)}} \left(\log|y_i - \mu_j^{(g)}| - \log s_j^{(g)} \right) \tag{2.38}$$

In (5.38) $\psi(.)$ is the digamma function defined as $\Gamma'(g)/\Gamma(g)$. The update equations for the location and the shape parameter are nonlinear and we use an iterative solver like the Newton-Raphson to obtain a numerical solution. The update equation for the shape parameter (5.34) is highly nonlinear and computationally very expensive. The shape parameter can also be obtained using the variance and the scale parameter information. In terms of the variance and the scale parameter, the update equation of the shape parameter has the following form

$$\sum_{i=1}^{N} \mathbb{E}\left[z_{ji} | \mathbf{Y}, \Psi^{(g)} \right] \left(y_i - \mu_j^{(g)} \right)^2 \Gamma\left(\frac{1}{\beta_j^{(g+1)}} \right) = \sum_{i=1}^{N} \mathbb{E}\left[z_{ji} | \mathbf{Y}, \Psi^{(g)} \right] \left(s_j^{(g)} \right)^2 \Gamma\left(\frac{3}{\beta_j^{(g+1)}} \right) \tag{2.39}$$

2.2.2.2 Generalized Beta Prime Mixture Model

The generalized beta prime mixture model is based on the density function of the generalized Beta prime distribution. Also known as the generalized beta of the second kind (GB2), the generalized beta prime distribution includes many important distributions as limiting cases. These include generalized gamma, Burr types 3 and 12, lognormal, Weibull, gamma, Rayleigh and exponential distributions. The generalized beta prime density used in this paper has the following form

$$p_Y(y; a, b, \gamma, q) = \frac{\gamma \left(\frac{y}{q} \right)^{a\gamma - 1} \left[1 + \left(\frac{y}{q} \right)^{\gamma} \right]^{-(a+b)}}{q \mathbb{B}(a, b)} \tag{2.40}$$

The support of $p_Y(y;a,b,p,q)$ does not encompass $\mathbb{R}$ but a smaller interval of $(0,\infty)$. In (2.40) *gamma* $\in \mathbb{R}^+$ is the shape parameter, $q \in \mathbb{R}^+$ is the scale parameter and $\mathbb{B}(a,b)$ is the beta function. It is worth noting that if $p = q = 1$, the generalized beta prime yields the standard beta prime distribution.

The conditional expectation of the complete data log-likelihood function of the generalized beta prime mixture model can be obtained by substituting (2.40) in (2.27). This substitution yields

$$Q\left(\Psi^{(g+1)},\Psi^{(g)}\right) = \sum_{i=1}^{N}\sum_{j=1}^{M}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{k}\right]\log\pi_j + \sum_{i=1}^{N}\sum_{j=1}^{M} \left(\log\gamma_j + (a_j\gamma_j - 1)[\log y_i - \log q_j] - (a_j+b_j)\log\left[1+\left(\frac{y_i}{q_j}\right)^{\gamma_j}\right] - \log q_j - \log\mathbb{B}(a_j,b_j)\right) \tag{2.41}$$

The application of (4.4)-(2.31) yields the following update equations for the model parameters $\Psi^{(g+1)} = \{\pi_j, a_j, b_j, p_j, q_j; j = [1,2,...,M]\}$ of a generalized beta prime mixture model given the current estimate $\Psi^{(g)}$ and the measured load data $\mathbf{Y}$

$$\pi_j^{(g+1)} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right] \tag{2.42}$$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\left[\gamma_j^{(g)}(\log y_i - \log q_j^{(g)}) - \log\left(1+\frac{y_i}{q_j^{(g)}}\right)^{\gamma_j^{(g)}} - \psi\left(a_j^{(g+1)}\right) + \psi\left(a_j^{(g+1)}+b_j^{(g)}\right)\right] = 0 \tag{2.43}$$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\left[\psi\left(a_j^{(g)}+b_j^{(g+1)}\right) - \psi\left(b_j^{(g+1)}\right)\right] - \sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\log\left[1+\left(\frac{y_i}{q_j^{(g)}}\right)^{\gamma_j^{(g)}}\right] = 0 \tag{2.44}$$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\left[\frac{1}{\gamma_j^{(g+1)}}-\frac{\left(a_j^{(g)}+b_j^{(g)}\right)y_i^{\gamma_j^{(g+1)}}\left(q_j^{(g)}\right)^{-\gamma_j^{(g+1)}}\left(\log y_i-\log q_j^{(g)}\right)}{y_i^{\gamma_j^{(g+1)}}q_j^{-\gamma_j^{(g+1)}}+1}\right]+$$
$$a_j^{(g)}\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\log\frac{y_i}{q_j^{(g)}}=0 \quad (2.45)$$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\left(\frac{a_j^{(g)}\gamma_j^{(g)}}{q_j^{(g+1)}}-\frac{1}{q_j^{(g+1)}}\right)+\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\left(\frac{1}{q_j^{(g+1)}}+\frac{\left(a_j^{(g)}+b_j^{(g)}\right)y_i^{\gamma_j^{(g)}}\left(q_j^{(g+1)}\right)^{-\gamma_j^{(g)}}}{q_j^{(g+1)}\left(1+y_i^{\gamma_j^{(g)}}\left(q_j^{(g+1)}\right)^{-\gamma_j^{(g)}}\right)}\right)=0 \quad (2.46)$$

2.2.3 Multivariate Generalized Gaussian Mixture Model

Consider a $d-$ dimensional random vector $\overline{\mathbf{Y}}=[\mathbf{Y}_1,\mathbf{Y}_2,...,\mathbf{Y}_d]^T$ with $d\geq 1$. $\overline{\mathbf{Y}}$ has a multivariate generalized Gaussian distribution with parameters μ,Σ,ζ where $\mu\in\mathbb{R}^d,\Sigma$ is a $(d\times d)$ definite positive symmetric matrix and $\zeta\in(0,\infty)$ if it can modeled with a density function of the form [57]

$$p_{\overline{\mathbf{Y}}}(Y;\mu,\Sigma,\zeta)=\alpha|\Sigma|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\left[(Y-\mu)^T\Sigma^{-1}(Y-\mu)\right]^{\zeta}\right) \quad (2.47)$$

where

$$\alpha=\frac{d\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{d}{2}}\Gamma\left(1+\frac{d}{2\zeta}\right)2^{\left(1+\frac{d}{2\zeta}\right)}} \quad (2.48)$$

Equation (2.47) is a multivariate generalized Gaussian distribution with scale parameter replaced by the scatter matrix Σ. It is worth noting that $\zeta=1,d=2$ in (2.47) gives the bivariate Gaussian distribution. The notation is $PE_d(\mu,\Sigma,\zeta)$ [57]. The parameter ζ controls the sharpness of the distribution and the sharpness diminishes as ζ increases. In that sense ζ is the multivariate analogue of the shape parameter β in the univariate generalized Gaussian distribution.

2.2.3.1 Stochastic Representation

If $\overline{\mathbf{Y}}$ is a d- dimensional random vector and is distributed as $PE_d(\mu,\Sigma,\zeta)$ then Gómez et al [57] have shown that $\overline{\mathbf{Y}}$ admits the following equality in distribution.

$$\overline{\mathbf{Y}} \overset{d}{=} \mu + RA'u^{(d)} \tag{2.49}$$

In (2.49) $u^{(d)}$ is a random vector uniformly distributed on the unit sphere in $\mathbb{R}^d$, R is a continuous non negative random variable and A is a square matrix with d order such that $A'A = \Sigma$. The density function of R is

$$h_R(r) = \frac{d}{\Gamma\left(1+\frac{d}{2\zeta}\right)2^{\frac{d}{2\zeta}}} \exp\left[-\frac{1}{2}r^{2\zeta}\right] I_{(0,\infty)}(r) \tag{2.50}$$

2.2.3.2 Parameter Estimation

The multivariate generalized Gaussian mixture model is characterized by the weight, location parameter, scatter matrix, and shape parameter of each component. The conditional expectation of the log-likelihood function of the multivariate generalized Gaussian mixture model can be obtained by substituting (2.47) in (2.27.

$$\begin{aligned} Q\left(\Psi^{(g+1)},\Psi^{(g)}\right) = \sum_{i=1}^{N}\sum_{j=1}^{M}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^k\right]\left[\log w_j + \log d + \log\Gamma\left(\frac{d}{2}\right) - \frac{d}{2}\log\pi - \right. \\ \left. \log\Gamma\left(1+\frac{d}{2\zeta}\right) - \left(1+\frac{d}{2\zeta}\right)\log 2 - \frac{1}{2}\log|\Sigma_j| - \frac{1}{2}\left[(Y_i-\mu_j)^T\Sigma_j^{-1}(Y_i-\mu_j)\right]^{\zeta_j}\right] \end{aligned} \tag{2.51}$$

In (2.51) w_j is the weight of the j^{th} component multivariate density. The application of (4.4)-(2.31) yield the following update equations for the model parameters $\Psi^{(g+1)} = \{w_j,\mu_j,\Sigma_j,\zeta_j; j = [1,2,...,M]\}$ of the multivariate generalized Gaussian mixture model given the current estimate $\Psi^{(g)}$ and the measured load data $\overline{\mathbf{Y}}$

$$\sum_{i=1}^{N}\mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]\zeta_j^{(g)}\left[\left(Y_i-\mu_j^{(g+1)}\right)^T\left(\Sigma_j^{(g)}\right)^{-1}\left(Y_i-\mu_j\right)\right]^{\zeta_j^{(g)}} = 0 \tag{2.52}$$

$$\Sigma_j^{(g+1)} = \left(\frac{\zeta_j^{(g)} \sum_{i=1}^{N} \mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right] \left[\left(Y_i-\mu_j^{(g)}\right)\left(Y_i-\mu_j^{(g)}\right)^T\right]^{\zeta_j^{(g)}}}{\sum_{i=1}^{N} \mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right]} \right)^{\frac{1}{\zeta_j^{(g)}}} \tag{2.53}$$

$$\sum_{i=1}^{N} \mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right] \left(\frac{d\log 2}{2\left(\zeta_j^{(g+1)}\right)^2} + \frac{d}{2\left(\zeta_j^{(g+1)}\right)^2} \psi\left(1+\frac{d}{2\zeta_j^{(g+1)}}\right)\right) -$$
$$\frac{1}{2}\sum_{i=1}^{N} \mathbb{E}\left[z_{ji}|\mathbf{Y},\Psi^{(g)}\right] \left[(Y_i-\mu_j^{(g)})^T \left(\Sigma_j^{(g)}\right)^{-1} (Y_i-\mu_j^{(g)})\right]^{\zeta_j^{(g+1)}} \log\left[(Y_i-\mu_j^{(g)})^T \left(\Sigma_j^{(g)}\right)^{-1} (Y_i-\mu_j^{(g)})\right] = 0 \tag{2.54}$$

It is worth noting that for $\zeta = 1$ in (2.53), the update equation for the scatter matrix of the multivariate generalized Gaussian reduces to the update equation of the covariance matrix of the multivariate Gaussian distribution [ref]. An alternate expression for the update equation of ζ can be obtained from the relationship between the multidimensional kurtosis coefficient $\gamma_2(\overline{\mathbf{Y}})$ and ζ.

$$\left(\gamma_2^j\right)^{(g+1)}(\overline{\mathbf{Y}})\Gamma^2\left(\frac{d+2}{2\zeta_j^{(g+1)}}\right) + d(d+2) - d^2\Gamma\left(\frac{d}{2\zeta_j^{(g+1)}}\right)\Gamma\left(\frac{d+4}{2\zeta_j^{(g+1)}}\right) = 0 \tag{2.55}$$

The kurtosis coefficient $\left(\gamma_2^j\right)^{(g+1)}(\overline{\mathbf{Y}})$ of the j^{th} component density at the current iteration is estimated from the variance. The update equation for the j^{th} multidimensional kurtosis coefficient has the form

$$\left(\gamma_2^j\right)^{(g+1)}(\overline{\mathbf{Y}}) = \mathbb{E}\left[\left(\left(Y-\mu_j^{(g)}\right)^T Var(Y)^{-1}\left(Y-\mu_j^{(g)}\right)\right)^2\right] \tag{2.56}$$

2.3 Initialization Procedures

The E-M algorithm has the property that the expected value of the log-likelihood function increases at each iteration. This means that the EM sequence of observed log-likelihood $\{\log \hat{L}\}_{\Psi|\mathbf{Y},\mathbf{Z}}$ will converge to the stationary point of $\log L(\Gamma|\mathbf{Y},\mathbf{Z})$. This stationary point is guaranteed to be the

local maximizer of the log-likelihood function of the model. In some cases, the EM sequence converges to a saddle point [58], but such instances are rare. One disadvantage of the E-M algorithm is the sensitivity of the final solution to the initial parameter values. The first iteration of the E-M algorithm requires the initialization of the mixture model's overall parameter Γ. Since mixture models tend to have many parameters, the initialization problem is very difficult.

In the literature on E-M initialization, several methods have been proposed. One possible initialization approach involves doing short E-M runs starting with random initial guesses followed by a long E-M run. Also known as iteratively constrained E-M, the idea rests on the assumption that the initial guess that maximizes the observed log-likelihood function will cause the algorithm to reach convergence more quickly than random initialization that changes the log-likelihood less rapidly. Once the initial guess that maximizes the log-likelihood is identified through short E-M runs, the long E-M run commences using the final solution of the short E-M as the initial guess. A disadvantage with iteratively constrained E-M is the number of iterations for short E-M runs must be specified. A second major disadvantage is deciding on the number of random starts. A second possible initialization strategy is the k-means algorithm which is the initialization procedure adopted in this paper. k-means is a clustering algorithm that can partition the data into distinct non-overlapping subgroups based on a specific similarity measure. The objective function of the k-means algorithm is

$$J = \sum_{i=1}^{N} \sum_{k=1}^{K} w_{ik} \|x_i - \mu_{ik}\|^2 \tag{2.57}$$

In (2.57), $w_{ik} = 1$ if the data point x_i belong to cluster k and zero otherwise. The k-means algorithm works on the Expectation-Maximization principle but with hard assignments. The study [59] has shown that k-means algorithm provides reasonably good parameter estimates when used as initial values for the E-M algorithm.

2.3.1 Univariate generalized Gaussian Mixture Model

The k-means algorithm outputs the total number of clusters, centroid locations of each cluster, and a label for each data point which indicates the membership of the data point with respect to

the clusters. The univariate generalized Gaussian mixture parameters are initialized by setting up the number of clusters from the k-means equal to the number of the component densities M in the mixture model. The location parameter of each mixture component is set equal to the centroid locations of the respective clusters. The scale parameter of each mixture component is initialized as

$$s_j^{(0)} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \mu_j^{(0)}\right)\left(y_i - \mu_j^{(0)}\right)^T \tag{2.58}$$

The initial value of the shape parameter β_j of the j^{th} component density is estimated from the first and second statistical moments. With the location parameter $\mu_j^{(0)}$ of the j^{th} component density initially set to the location of the j^{th} centroid, β_j is estimated as

$$\beta_j^{(0)} = \frac{m_{1,j}}{\sqrt{m_{2,j}}} \tag{2.59}$$

In (2.59) $m_{1,j}$ and $m_{2,j}$ are the first and the second statistical moment of the j^{th} cluster respectively. The initialization of β provides an alternate initialization strategy for the scale parameter s. More specifically, s can be also be initialized as

$$s_j^{(0)} = \left(\frac{\beta_j^{(0)}}{N}\sum_{i=1}^{N}\left|y_i - \mu_j^{(0)}\right|^{\beta_j^{(0)}}\right)^{\frac{1}{\beta_j^{(0)}}} \tag{2.60}$$

2.3.2 Multivariate generalized Gaussian Mixture Model

The application of the k-means algorithm in the multivariate case results in a $M \times d$ matrix for the centroid locations. The centroid locations are used to initialize the location parameter vector for each component density of the mixture model. The initialization of the scatter matrix is based on the covariance of the d- dimensional measured data vector $\overline{\mathbf{Y}}$. The multivariate scale parameter ζ is initialized in the same way as the univariate scale parameter β.

2.4 Sampling from Mixture Distributions

The mixture models discussed in the previous sections can be used to fit the empirical multimodal load distribution with high accuracy. The fitting process involves using the empirical load data to estimate the parameters of the mixture model. A random number generator function can draw repeated samples from the fitted model. The new load profile thus created will be statistically similar to the measured data. For the generalized Gaussian mixture model, a random number generator function can be obtained by taking the inverse of the CDF of the model.

$$F_Y(y|\Psi) = \int_{-\infty}^{y} \sum_{j=1}^{M} \pi_j \frac{\beta_j}{2s_j\Gamma\left(\frac{1}{\beta_j}\right)} \exp\left[-\frac{|t-\mu_j|^{\beta_j}}{s_j}\right] dt \tag{2.61}$$

In the case of mixture models with component densities based on generalized probability distributions, a closed form solution to $\hat{y} = F_Y^{-1}(y|\Psi)$ however, does not exist, and hence numerical techniques such as Newton-Raphson must be used to generate random samples from the fitted model. For a univariate mixture model, rejection sampling offers an alternate solution to the random number generation problem. Rejection sampling can be used to draw samples from distribution when the inverse CDF transform method is either computationally too expensive or intractable. However, the efficiency of rejection sampling depends heavily on the choice of the proposal distribution. More specifically, if $\mathscr{X}_f$ is the support of the target distribution and $\mathscr{X}_g$ is the support of the proposal distribution, then $\mathscr{X}_f \subset \mathscr{X}_g$. In other words $g(x) > 0$ whenever $f(x) > 0$. The second important requirement concerns the scaling of the proposal distribution. The proposal distribution must be scaled by a factor C such that

$$C = \sup_{x \in \mathscr{X}_f} \frac{f(x)}{g(x)} < \infty \tag{2.62}$$

The scaling factor C is the appropriate bound on $\frac{f(x)}{g(x)}$ that satisfies $f(x) < Cg(x) \forall x \in \S_f$. Rejection sampling efficient in low dimensional spaces (<10 dimensions) due to inverse relationship between the acceptance probability and the scaling factor. The acceptance probability in rejection sampling

is

$$\Pr\left\{u \leq \frac{f(x)}{Cg(x)}\right\} = \frac{1}{C}; u \sim U(0,1) \tag{2.63}$$

2.5 Mixture model fitting of Load Data

This section will study the mixture models' applications in the power system load modeling. Two types of loads are considered. A residential load representative of a medium household with the infrastructure to support level- 1 (L-1) residential electric vehicle charging. The second load is a commercial building model augmented with high-level (L-2) electric vehicle charging infrastructure. In both cases, the base load data (load without EV charging) is obtained from the OpenEI website [60]. The web page was established by the Department of Energy (DoE) in 2009 and is a part of DoE mission to disseminate data in the public domain. The annual aggregate load distribution of a residential household with and without EV charging is shown in Figures 2.6 and 2.7

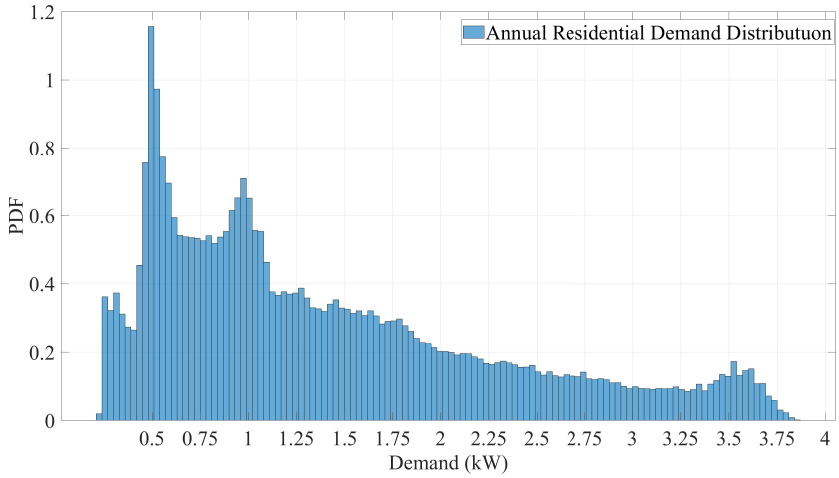

Figure 2.6: Annual Residential Load Distribution

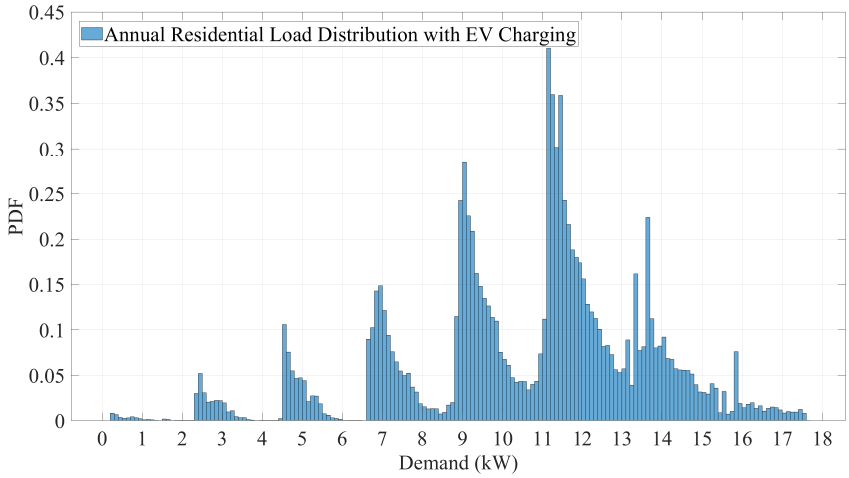

Figure 2.7: Annual Residential Load Distribution with EV charging

Residential EV charging has the overall effect of transforming a quasi-unimodal load distribution into a distribution characterized by multiple high-density regions. This is because the aggregate effect of smoothening of variations is absent in the case of residential EV charging, which mostly consists of just one vehicle connected to the grid. The annual aggregate load distribution of a commercial facility with and without commercial EV charging is shown in Figures 2.8 and 2.9

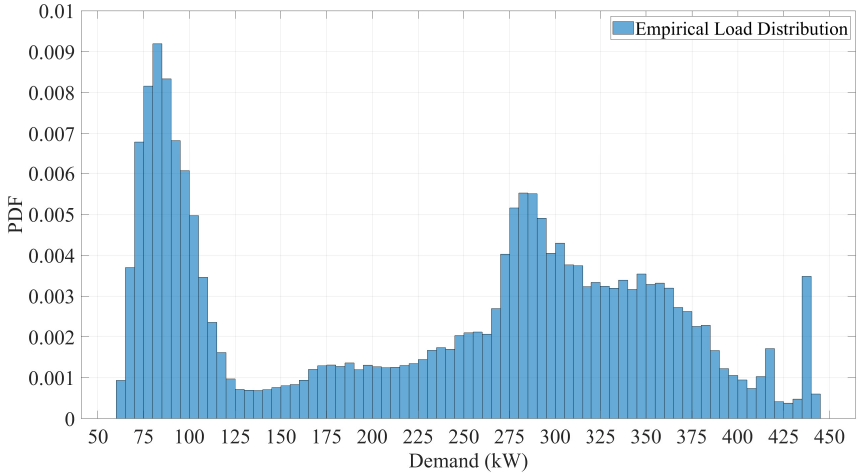

Figure 2.8: Annual Commercial Load Distribution

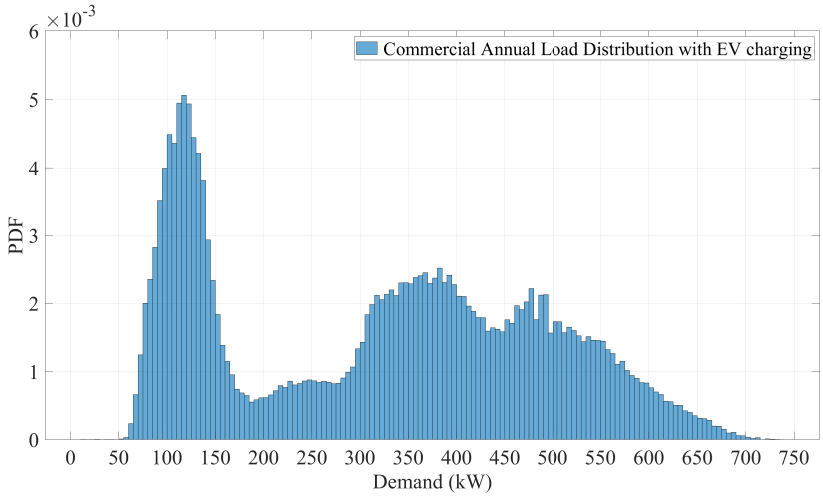

Figure 2.9: Annual Commercial Load Distribution with EV charging

The E-M algorithm used to obtain the parameters of the various mixture models as described in the previous sections was coded in MATLAB [61] and initialized using $\mathbf{K}-means$. The algorithm was run on an Intel Xeon processor with 32 GB of RAM. At each iteration of the algorithm, the log-likelihood function is calculated, and the difference between two consecutive values of the log-likelihood is compared with the tolerance. If the difference is less than tolerance, the algorithm is terminated.

$$\left| \frac{l^k - l^{k-1}}{l^{k-1}} \right| \leq \varepsilon \tag{2.64}$$

A threshold value of $\varepsilon = 1e-07$ was used to terminate the E-M algorithm. The GGMM fit of the residential and the commercial load distributions considering EV charging is shown in Figures 2.10 and 2.9. The GGMM model can capture the multi-modal characteristics of the overall load distribution. It is possible that a certain value of the load may be completely characterized by a single component of the GGMM in which that component may need to be identified and isolated from the rest. Most load values are generally best represented by some weighted combination of the mixture components.

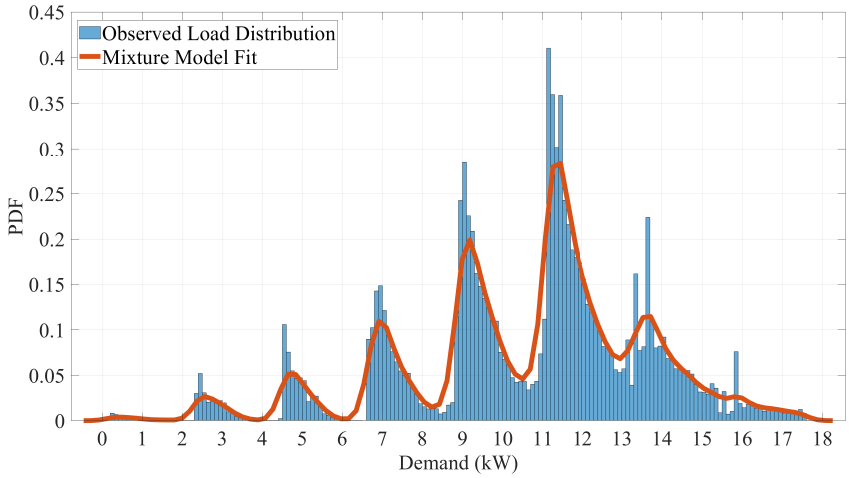

Figure 2.10: GGMM Fitting of Residential Load Distribution with EV Charging

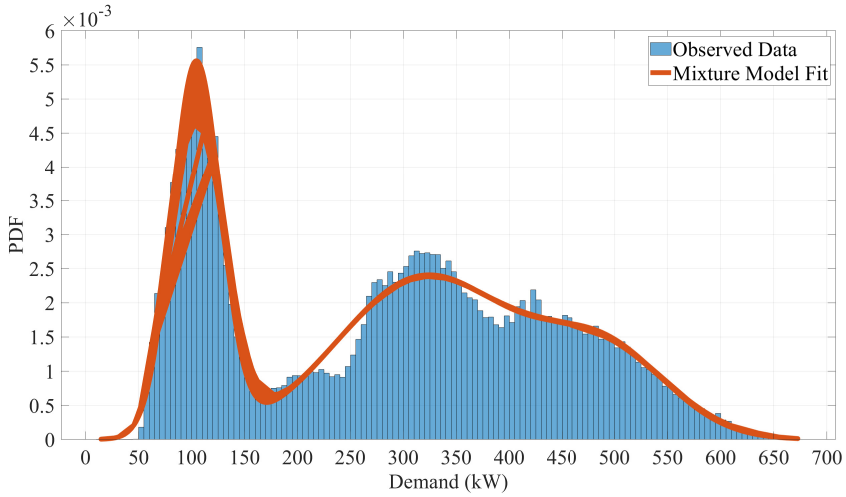

Figure 2.11: GGMM Fitting of Commercial Load Distribution with EV Charging

Figure 2.12 shows the log-likelihood function (2.21) of the GGMM fitted to the commercial load with EV charging shown in figure 2.10 plotted for a different number of mixture components as a function of the iteration number. It is clear that the log-likelihood is nondecreasing at each iteration step, and it can be seen from Fig 2.13 that as the number of mixture components of the GGMM increases, the computational time to solve the model also increases. This can be inferred from Fig 2.12 since the log-likelihood takes more time to converge for a higher number of mixture components.

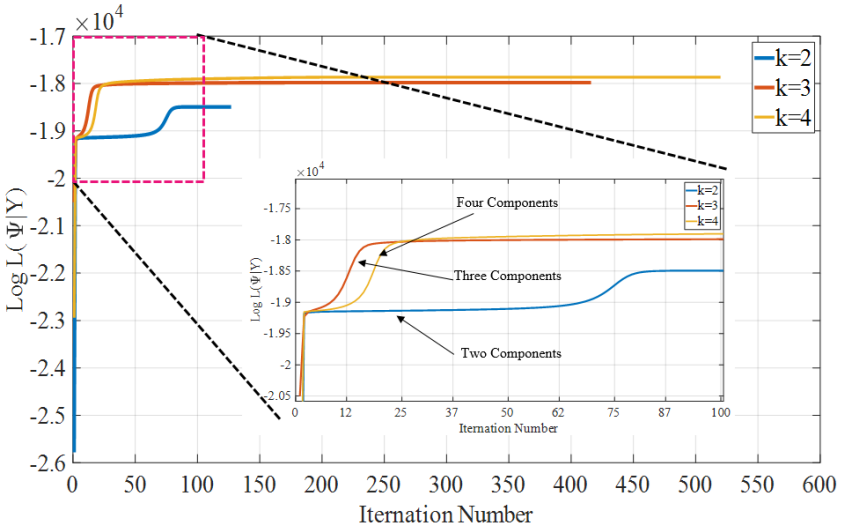

Figure 2.12: Log-likelihood function

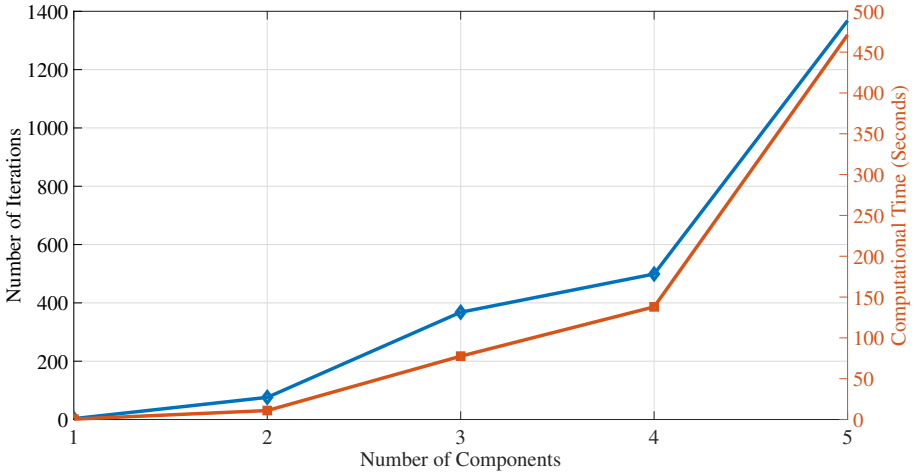

Figure 2.13: Performance characteristics of GGMM Fitting

The number of iterations required to estimate the parameters of the GGMM and the computational time in seconds is plotted as a function of the number of components. The results are shown in Fig 2.13 . On closer look, the results shown in Fig 2.13 corroborate the changes in the log-likelihood function as shown in Fig 2.12 The iteration count and the computational time share a nonlinear relationship with the number of mixture components. If we increase the component densities for more accuracy, the algorithm requires more iterations and hence more computational

time to achieve convergence. There is no general rule for determining the optimal number of components in mixture models. Our application uses the mean square error metric to determine the optimal number of component densities. The mean square error as a function of several components is plotted in Fig 2.14. It is clear that there is no significant improvement in the fitting accuracy for $M > 4$. For this reason, we chose $M = 4$ to generate the GGMM fit for the load distribution of Figure 2.14.

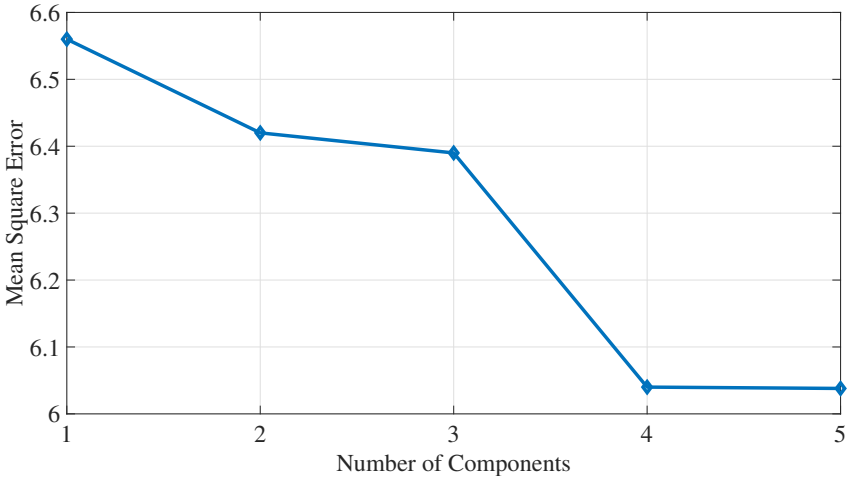

Figure 2.14: Mean Square Error of GGMM Fits

2.6 Conclusions

The work presented in this chapter explores the applications of mixture models for load modeling, considering the impact of distributed energy resources. In particular, the impact of EV charging, residential and commercial, is considered. The EV charging profile is obtained by modeling the EV arrival process at a charging facility as a nonhomogeneous Poisson process. The NHPP model is used to estimate the arrival times of EVs. The charging time for each arriving EV is calculated from the daily driven miles and the state of charge of the EV at the time of arrival. The daily driven miles are assumed to be log-normally distributed, and the battery SoC at arrival is estimated considering the battery performance and parameters of Tesla Model 3. The EV charging profile thus obtained is added to the baseload profile of the commercial charging facility to obtain

the overall demand pattern. The planning horizon considered is one day (24 hours).

The statistical properties of the total electric demand with EV charging are modeled using mixture modals with generalized component densities. The parameters of the mixture models are obtained using the E-M algorithm. The results shown in the paper demonstrate the applicability of a GGMM in representing load distribution with a pronounced peaky behavior. Since the proposed mixture model is parametric and hence "generative". The model can generate synthetic load data with similar statistical properties as the measured data. Since the proposed model considers the EV charging load, the GGMM proposed has far-reaching applications such as probabilistic load flow with EV charging and distribution system state estimation (DSSE), where many pseudo measurements are used to run the state estimation algorithms. The GGMM can model the non-gaussian distributed measurements, especially scenarios involving heavy EV penetration and distributed generation.

Another important application of the proposed GGMM model is designing Monte Carlo simulations where the inputs are sampled from some underlying distributions. The GGMM model can be used to create different realizations of the electric demand curve and used as an input to run stochastic optimization algorithms.

3. UNCERTAINTY CHARACTERIZATION OF DISTRIBUTED GENERATION

The output of a solar photovoltaic system appears to be more stable when viewed at hourly intervals. However, due to the transient cloudy conditions and random weather disturbances, the generation output of a typical photovoltaic system usually suffers rapid variations. To estimate the impact of adding solar energy to the existing generation pool, it is thus crucial that the intermittency of the incident radiation, when viewed in sub-hourly time scales, be appropriately accounted for. Since the availability of high-resolution solar data adequately captures the sub-hourly variations in the solar insolation incident on the solar panel, it becomes imperative to use stochastic techniques to synthesize data with a very high temporal resolution. In this study, we utilize the hourly averaged Typical Meteorological Data (TMY3) [62] to generate the high resolution minute-by-minute solar irradiance profile. The Markov weather model used to generate changes in the solar insolation with a temporal resolution of one minute from the given averaged hourly values is described in Figure 1. The model is initialized by calculating the average hourly clearness index k_t. The clearness index k_t is defined as the ratio of measured irradiance E_m at the earth's surface and the irradiance corresponding to cloudless conditions at the same location. This is referred to as clear sky irradiance E_{clear}

$$k_t = \frac{E_m}{E_{clear}} \tag{3.1}$$

3.1 Markov Chain Monte Carlo Simulation Method*

[1] The accurate estimation of the clear sky irradiance E_clear is very important and significantly influences the clearness index. The clear sky irradiance is estimated based on the procedure given in [9]. The model works by using the hourly k_t values as input. The sub-hourly transitions of the clearness index are determined by utilizing the transition probabilities extracted from the clearness index, although at a higher temporal resolution as opposed to the hourly k_t values. This information

[1] Part of the data reported in this chapter is reprinted with permission from "On the Environmental and Economic Impact of Utility-Scale Solar Deployment" by Aaqib Peerzada, Miroslav Begovic, Dejan Ostojic, Proceedings of the 55th Hawaiian International Conference on System Sciences|2022

is passed on to a first-order Markov process which estimates the next state of the clearness index (sub-hourly) based on the current state (hourly). Assuming that the clearness index can assume a total of n states, the transition probability of the first order Markov process can be described by the following equation

$$\Pr\{k_t(n+1) = j | k_t(n) = i\} = p_{ij} \tag{3.2}$$

These transition probabilities are determined based on changes in the k_t values that are generated at a higher temporal resolution which is the same as the desired resolution of the model output. The transition probabilities thus calculated are grouped together in a transition probability matrix and $n \times n$ matrix with n representing the total number of states of the form.

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$$

The transition probability matrix is a stochastic matrix since the cumulative probabilities of each row sum to one.

$$\sum_{j=1}^{n} p_{ij} = 1; i = 1, 2, ..., n \tag{3.3}$$

The final step of the model is executed by running a Monte Carlo simulation of the Markov process. This is accomplished by sampling a uniformly random $u \sim U(0,1)$ and comparing it with the cumulative probability of each row in the transition probability matrix. For instance, if the current state of the clearness index is i, the cumulative probability of state i can be determined by summing the transition probabilities of the i^{th} row in the transition probability matrix. If j and m are two consecutive states and $k_t(n) = i$, the clearness index at the next time step $k_t(n+1) = m$ if

$$F_{ij} \leq u \leq F_{im} \tag{3.4}$$

The cumulative probability F_{im} of state i is

$$F_{im} = \sum_{k=1}^{m} p_{ik} \tag{3.5}$$

An example of a transition probability matrix extracted from the sub-hourly clearness index values is shown in Figure 3.1. The sub-hourly clearness index values are calculated based on the measured global irradiance and the estimated clear sky irradiance of the Milford area in Utah. The raw data used to extract the transition probabilities has a temporal resolution of one minute and was retrieved from the NREL solar database [62]. The diagonal dominance of the transition probability matrix is evident with some outliers that indicate rapid changes in the clearness index. The vertical axis represents the current state of the clearness index, and the horizontal axis is the next state. The clearness index takes values in the interval $[0, 1]$. The colored boxes in Figure 3.1 represent the corresponding transition probabilities which are color-coded and can be read off the color bar. From Figure 3.1 it can be inferred that the highest transition probabilities correspond to the clear sky index to stay the same or change very slightly. The same is true for a cloudy day with $k_t = 0$, thus giving the transition matrix a diagonal structure where the largest probabilities occur on the diagonal.

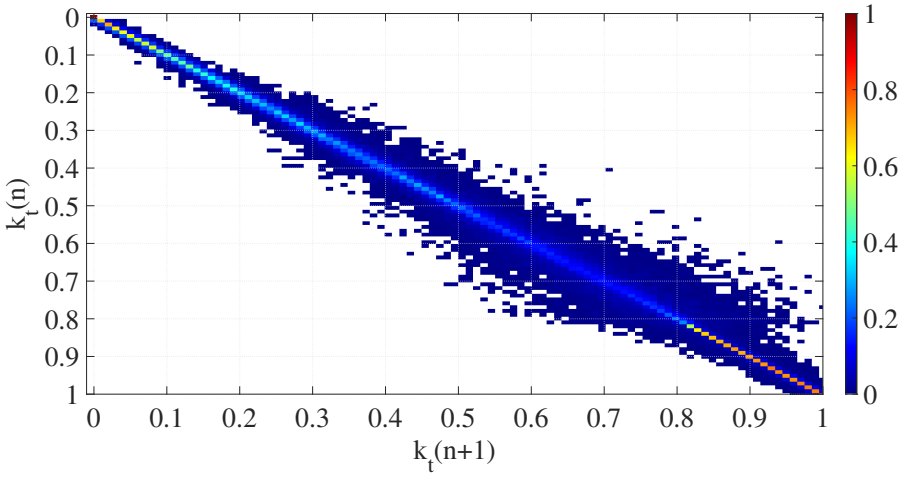

Figure 3.1: Representation of TPM indicating changes in k_t

The transition probability matrices thus obtained are used as an input to initiate the Markov Chain Monte Carlo simulations. The MCMC simulation algorithm combines the transition prob-

abilities generated at sub-hourly time scales with the hourly average TMY3 [62] meteorological weather data to generate high-resolution irradiance profiles. We use the hourly averaged TMY3 weather data of the state of Arizona. The roof-top photovoltaic (PV) systems are oriented at an azimuth of 180° (south-facing) with a panel tilt of 30°. The capacity of each PV system is chosen to be proportional to the peak load. Figure 3.2 shows the total power generated from all the roof-top PV systems in the IEEE-123 bus system for one day with transient cloud conditions.

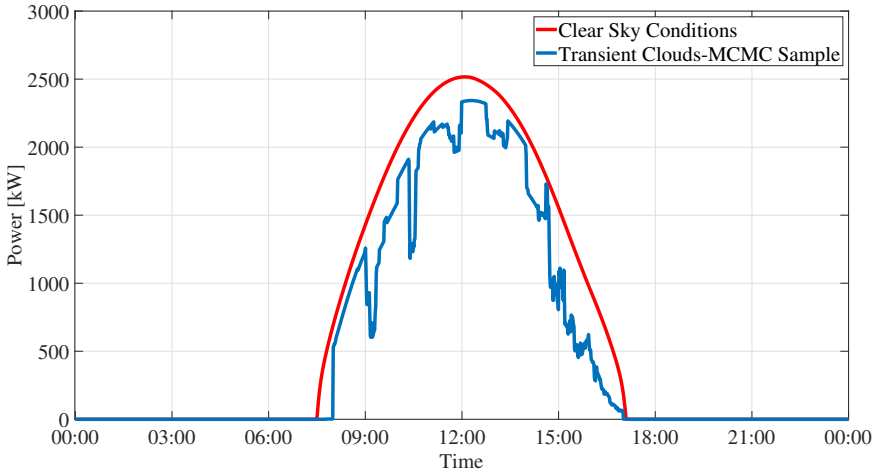

Figure 3.2: Combined output of all PV systems in IEEE-123 Bus System

3.2 Time series modeling of Solar Data

The Markov weather model discussed in the previous section can generate high-resolution solar irradiance data with high enough fidelity to reproduce the rapid excursions in the output of a solitary PV system. The estimation of the clearness index, which is defined as the ratio of measured irradiance at the earth's surface at a location and the cloudless sky irradiance at the same location, is central to the Markov weather model. The Markov weather model works by combining the low-resolution clearness index based on the hourly averaged TMY3 [62] data with transition probabilities generated at sub-hourly time scales from measured high-resolution data. This results in an output with a high temporal resolution, 1-minute in this case, and seasonal variation.

The solar output of a PV system can be modeled as a continuous stochastic process. Since optimization of continuous stochastic processes is complicated and even impossible in many cases, a

discrete-time approximation is often used in formulating optimization problems. The continuous stochastic process G^i_{PV} representing power generated at bus i in a network can be well approximated by a discrete process $\hat{G}^i_{PV}$ such that

$$\hat{G}^i_{PV} = \{\hat{G}^i_{PV}(t,\omega) = [g^i_{PV}(1,\omega), g^i_{PV}(2,\omega), ..., g^i_{PV}(T,\omega)], \omega = 1,2,...,H\} \tag{3.6}$$

In (3.6) T is the length of the time horizon, ω is the scenario index, and H is the number of possible scenarios, which is also equal to the number of Monte Carlo runs. The stochastic process $\hat{G}^i_{PV}$ is completely determined by the joint distribution of the random variables $\{g^i_{PV}(t,\omega); t = 1,2,...,T, \omega = 1,2,...,H\}$. The joint distribution can also be used to evaluate the marginal distribution of the random variable $g_{PV}(t,\omega)$ and the statistical dependencies that exist among these random variables. The estimation of the joint distribution, however, is challenging and is simplified by assuming that the joint distribution is a multivariate Gaussian and the stochastic process $\hat{G}^i_{PV}$ is stationary. The assumption of a multivariate Gaussian implies that the marginal distributions are all univariate Gaussian. The assumption of stationarity implies that the mean, variance, and covariance of the stochastic process $\hat{G}^i_{PV}$ are time-invariant. .

With these assumptions the joint distribution of the stochastic process $\hat{G}^i_{PV}$ can be determined by ARMA modeling of the time series data. The time series data used to train the ARMA model is obtained from the Markov weather model. Mathematically, an ARMA (p,q) model with p auto-regressive parameters and q moving average parameters has the form

$$g^i_{PV}(t,\omega) = \sum_{i=1}^{p} \phi_i g_{PV}(t-i,\omega) + \varepsilon(t) + \sum_{j=1}^{q} \theta_j \varepsilon(t-j) \tag{3.7}$$

In (3.7) the term $\varepsilon(t)$ is an uncorrelated stochastic process with zero mean and variance σ^2_ε. The stochastic process $\varepsilon(t)$ is called white noise or innovation term. Before realizing the ARMA (p,q) model, i.e., estimating the model orders and coefficients, it is important to test whether the stationarity assumption holds for the time series data used to train the ARMA model in (3.7). The annual 1-minute time series data from the Markov weather model is tested for stationarity using the Augmented Dickey-Fuller (ADF) test [63]. The ADF test contains the null hypothesis that the

time series is non-stationary. The test result rejects the null hypothesis, implying that the series may be stationary and that the use of the ARMA model may be justified.

The ARMA model orders p, q and the model coefficients $\phi_1, ..., \phi_p$ and $\theta_1, ..., \theta_q$ are estimated by implementing the model in the System Identification Toolbox (SIM) in MATLAB [64]. Using the high-resolution time series solar generation data from the Markov weather model, the SIM constructs mathematical models with different combinations of model orders. The order combination with the least Bayesian Information Criterion (BIC) value is selected. Conversely, the order combination with the largest log-likelihood value is chosen since large log-likelihood values represent better fits.

The scenarios for solar generation are generated using the Monte Carlo sampling of the fitted time series ARMA model. We assume a three-phase PV system in the IEEE 34 bus test system with a peak power rating of 450 kW proportional to the peak active load. The voltage rating of the PV system is 4.16 kV. The PV system is oriented at an azimuth of 180 °(south-facing) and a tilt angle of 30 °. Since the IEEE-34 bus test feeder is based on an existing distribution system located in the state of Arizona (AZ), the hourly average TMY3[62] data of AZ is used to synthesize a high resolution (1-minute) irradiance profile. The solar output in 1-minute intervals is estimated using the PV system model in OpenDSS[65]. The high-resolution solar output data is used to fit the ARMA model. The fully realized ARMA model is used to generate solar forecasts. Figure 3.3 shows one hundred 1-minute ahead solar output scenarios for a typical spring day in AZ along with the mean forecast and 95% confidence interval. The parameters of the realized ARMA model are given in Table 3.1.

An important challenge in time series forecasting is the error of the realized ARMA model as the size of the forecasting window increases. A smaller forecasting window (usually 1 minute) is characterized by a small forecasting error but, at the same time, is computationally expensive. A larger forecasting window reduces the computational burden while the associated forecasting error increases. The optimal choice of the forecasting window ultimately depends on the application and the computational resource available. Figure 3.4 plots the ahead solar forecasts for different choices of forecasting window. Clearly, as the model attempts to forecast much ahead into the future, the forecasting error increases.

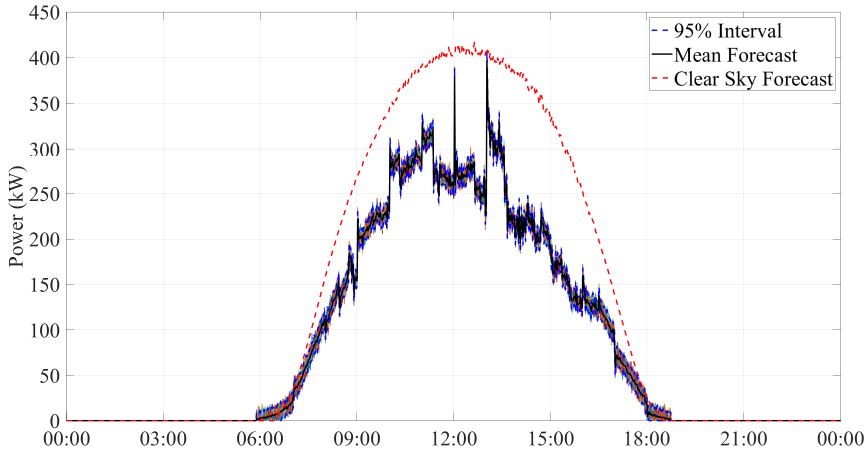

Figure 3.3: 1-minute ahead solar forecasts from ARMA model from Table 3.1

Table 3.1: The Realized ARMA (p,q) Model

p	q	ϕ_i	θ_j
3	3	$\phi_1 = 0.376$ $\phi_2 = -0.3722$ $\phi_3 = 0.9848$	$\theta_1 = 0.6151$ $\theta_2 = 0.9845$ $\theta_3 = -0.005$

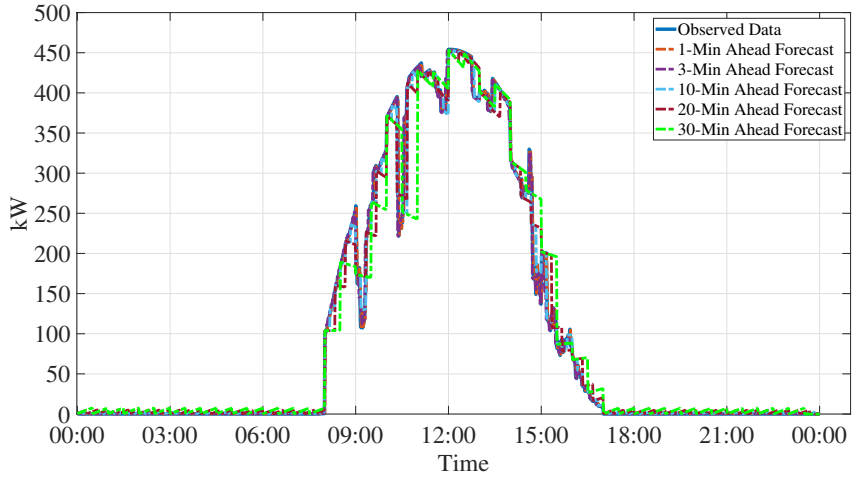

Figure 3.4: Ahead forecasts of Solar Generation

Figure 3.5 shows the root mean square error and the mean square error for different choices of the forecasting window. Both the metrics report a signiificant increase with increasing forecasting window.

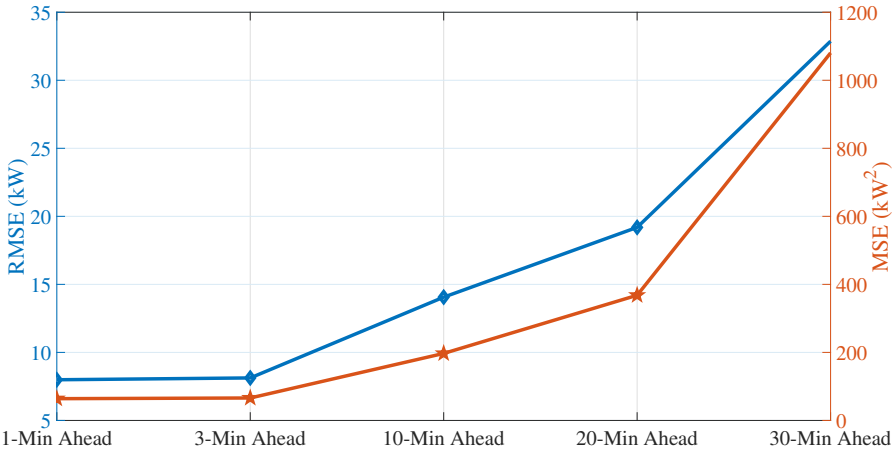

Figure 3.5: Ahead forecasts of Solar Generation

4. UTILITY SCALE SOLAR DEPLOYMENT: A CASE STUDY

The latest inventory report of the U.S Energy Information Administration states that renewable energy, most notably solar and wind, will account for nearly 70% of approximately 40 gigawatts of new generating capacity in 2021. Utility-scale solar is set to make significant strides in 2021, with 15.4 gigawatts of new capacity addition surpassing the 12 gigawatts increase in 2020. The rapid increase in the use of renewable energy is expected to have a significant impact on greenhouse emissions while at the same time altering the conventional generation resource pool. This chapter presents a case study investigating the impact of utility-scale solar capacity addition and the implementation of carbon tax policy on the overall profitability of adding renewable generation to the existing energy mix. The investigation is carried out on the generation resource pool of the southeast region of the U.S augmented by a substantial amount of utility-scale solar generation.

4.1 Background

Over the past decade, the deployment of renewable energy, especially solar and wind, has dominated the addition of new generation capacity in the United States [66]. Even though renewable energy appears to be the most popular new addition to the existing energy mix across many states in the U.S, the intermittent nature of such generation adds a new layer of complexity to the operation and planning of the power grid. One example of such complexity is the concept of system net load. Traditionally the system load profile has been extensively used for power system management applications. However, the accelerating expansion of renewables has rendered the system load profile less informative and thus inadequate for most operation and planning applications. Instead, system net load has gained popularity, defined as the demand that dispatchable (non-intermittent) sources must meet. An interesting recent work to estimate the system net load is presented in [67]. The system net load exhibits much faster changes than the traditional load profile when viewed in sub-hourly time scales. The changes are more pronounced when viewed at a very high temporal resolution. This is because renewables like solar and wind output are intermittent, thus imposing

a variable generation pattern on the system. This has an important consequence on the resource allocation of the dispatchable sources, the output of which needs to be regulated at much shorter timescales to meet the rapid changes in the system net load. Responding to such rapid changes may even incentivize the utilization of generation ramping rates beyond traditional elastic limits [68]. To meet the changes in the system net load, the system operator can re-dispatch fast response units like gas turbines while maximizing using cheap base load units like nuclear and coal. An example of such an economic dispatch model while considering ramping rates in the fuel cost function is given in [69]. While adding renewable generation to the resource pool will significantly affect the total generation cost, as has been reported in previous works [70], it also offers opportunities for utilities to lower their carbon footprint. This directly translates into monetary savings that could be achieved by policies controlling carbon pricing, such as carbon tax. However, the frequent ramping up or down of the dispatchable thermal power plants to meet the system net load may also lead to increased emissions, thus exposing electric utilities to additional losses due to carbon emissions. Some early work investigating the impact of gas turbine ramping on carbon emissions is presented in [71], [72]. These studies utilize wind or solar photovoltaic at one-minute and five-minute resolutions, respectively, and heat data from natural gas generators to assess the impact on emission reduction. Both studies have found evidence for an overall reduction of carbon emissions due to the addition of solar generation. The previous work on this topic does not consider the marginal cost of carbon and the changes in the total generation cost with different penetration levels of utility-scale renewable energy deployment. Whether or not to invest in renewable energy to offset the cost incurred due to the increased emissions is a techno-economic decision and should be addressed by considering various technical and economic aspects of the power system operation and planning. In this work, we aim to investigate the impact of utility-scale solar energy deployment on the total generation costs and the emissions profile and to propose a framework for assessment of the profitability of investing in renewable energy, especially utility-scale solar generation, to meet carbon footprint targets.

4.2 Problem Description*

[1] The objectives of this chapter are a) to study the impact of geographically dispersed utility-scale solar generation on resource allocation and b) to investigate the changes in the carbon footprint of the existing resource pool by adding solar generation to the mix, and c) to investigate the profitability of investing in geographically dispersed utility-scale solar generation to offset the increased emission costs incurred due to the use of fossil-fuel based energy resources like coal and gas. This investigation hinges on calculating a metric referred to as the annualized marginal benefit of carbon abatement (MBCA), defined as the net percentage change in the overall operational cost. More specifically

$$MBCA = \frac{C_T - C_T(PV, c_t)}{C_T} \times 100 \tag{4.1}$$

In (4.1) C_T is the total annual cost incurred when the overall system demand is met only with the non-intermittent sources such as coal-fired plants and natural gas plants, assuming no carbon tax policies exist. $C_T(PV, c_t)$ is the overall cost considering renewable energy such as utility-scale solar generation in the energy mix and assuming a carbon tax policy. The overall cost would then be the sum of annual generation and emission costs. The units for both are \$/year. The costs associated with the carbon emissions and the electric energy generation are estimated by running optimal dispatch on the coal-fired plants and gas turbines while considering gas ramping costs in the optimization process. The output of the solar generation plants is estimated by running Markov Chain Monte Carlo Simulations (MCMC) since the output of a photovoltaic system (PV) has been shown to have a Markovian dependence [73]. The emissions are estimated based on the generation profile of the non-renewable sources. This is accomplished by running optimal dispatch on the coal-fired plants and gas turbines while considering the ramping costs in the optimization process. The optimal dispatch is run to meet the system net demand, which is obtained by subtracting the net solar generation output of the geographically dispersed utility-scale plants from the actual system

[1] Part of the data reported in this chapter is reprinted with permission from "On the Environmental and Economic Impact of Utility-Scale Solar Deployment" by Aaqib Peerzada, Miroslav Begovic, Dejan Ostojic, Proceedings of the 55th Hawaiian International Conference on System Sciences|2022

load.

4.3 Modeling of High-Resolution Solar Data*

[2] The generation of transition probability matrices (TPMs) is central to synthesizing high-resolution solar data (irradiance). The procedure to generate the TPMs is outlined in chapter 3. The MCMC simulation algorithm combines the transition probabilities generated at sub-hourly time scales with the hourly average TMY3 meteorological weather data to generate high-resolution irradiance profiles. We use the hourly averaged TMY3 weather data of the seven representative sites in the southeast region of the United States. The PV systems at each representative site are oriented at an azimuth of 180°(south-facing) with a panel tilt of 30°. The capacity of each solar farm is chosen to be proportional to the population of each site in 2019. The solar generation output corresponding to high-resolution solar irradiance of the seven representative sites in the southeast region of the United States is shown in Figure 4.1 The solar output of each representative site

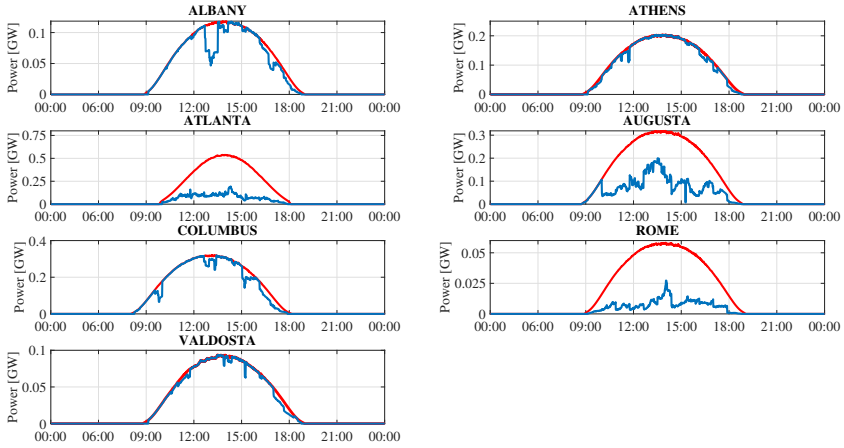

Figure 4.1: Generation output of seven sites in the Southeast region of the U.S

shows significant variability. When the generation outputs are averaged over several solar farms that are dispersed geographically, the variability in the production will decrease. The reduced variability in the aggregated output, however, will depend on the measure of similarity in the

[2]Part of the data reported in this chapter is reprinted with permission from "On the Environmental and Economic Impact of Utility-Scale Solar Deployment" by Aaqib Peerzada, Miroslav Begovic, Dejan Ostojic, Proceedings of the 55th Hawaiian International Conference on System Sciences|2022

production patterns, which in turn depends on the spatial spread of the cluster. For the seven representative sites shown in 4.1 the aggregated output on the same day in shown in Figure 4.2. The red curve represents the aggregated output for clear sky conditions, while the blue curve is the aggregated generation output that accounts for the cloud movement. Figure 4.2 shows a significantly smaller variation in the generation output than any of the individual solar farms in Figure 4.1 on a transient cloudy day.

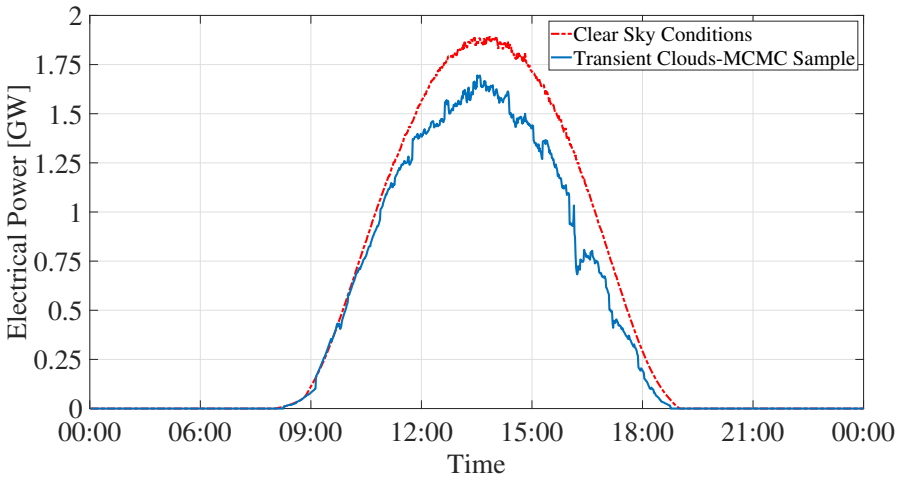

Figure 4.2: Generation output of seven sites in the Southeast region of the U.S

4.4 Framework for Optimal Dispatch*

[3] The optimal dispatch is run for coal-fired and gas turbines to meet the system net load. To account for the ramping of gas turbines the fuel cost model is modified by adding a ramping cost term to the quadratic cost function. The following cost functions for coal-fired and gas turbines are used

$$C_{g,Coal} = \alpha + \beta P_{g,Coal} + \gamma P_{g,Coal}^2 \tag{4.2}$$

$$C_{g,Gas} = \alpha + \beta P_{g,Gas} + \gamma P_{g,Gas}^2 + \tau \frac{dP_{g,Gas}}{dt} \tag{4.3}$$

[3]Part of the data reported in this chapter is reprinted with permission from "On the Environmental and Economic Impact of Utility-Scale Solar Deployment" by Aaqib Peerzada, Miroslav Begovic, Dejan Ostojic, Proceedings of the 55th Hawaiian International Conference on System Sciences|2022

Load following is achieved by regulating the output of gas turbines due to their fast ramping capability. The cost coefficients in (4.2) and (4.3) are estimated by least squares method applied to real data. The cost coefficients are given in [69]. The value of the ramp rate coefficient τ depends on the change in the gas turbine output. The formulation is such that a heavy penalty is levied for an increasing ramp rate as opposed to a decreasing ramp rate. The optimal dispatch problem for resource allocation is formulated as

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{\kappa}\sum_{s=1}^{S} C_s(P_{gs}) + \sum_{k=1}^{\kappa}\sum_{r=1}^{R} C_r(P_{gr}) \\ \text{subject to} \quad & \sum_{s=1}^{S} P_{gs} + \sum_{r=1}^{R} P_{gr} - P_D = 0 \\ & P_{gs}^{min} \leq P_{gs} \leq P_{gs}^{max} \\ & P_{gr}^{min} \leq P_{gr} \leq P_{gr}^{max} \end{aligned} \tag{4.4}$$

In 4.4 S and R refer to the total number of coal and gas plants, κ is the optimization period, P_D is the net load,P_{gs} and P_{gr} are the outputs of s^{th} and r^{th} coal and gas plant respectively.

4.5 Impact on conventional dispatch and carbon emission*

[4] To assess the impact of utility-scale solar deployment on the generation costs and emissions, the study assumes that solar plants are installed across the southeast region of the U.S. The study uses the energy mix of the southeast region as an input for running the economic dispatch. Table 4.1 lists the energy mix of the SE region as reported by EIA for the year 2020. The data pertains to the fuel type, percentage of the mix, and generation operating costs specific to fuel type in $/MWh. The study uses the unsubsidized average Levelized cost of energy of solar generation.

[4]Part of the data reported in this chapter is reprinted with permission from "On the Environmental and Economic Impact of Utility-Scale Solar Deployment" by Aaqib Peerzada, Miroslav Begovic, Dejan Ostojic, Proceedings of the 55th Hawaiian International Conference on System Sciences|2022

Table 4.1: Energy Mix of SE Region, 2020 [4]

Type	Generation (TWh)	Fuel	% Mix	$/MWh
Base	47.719	Nuclear	18.89	10.63
Intermediate	57.225	Coal	22.64	90
	12.982	Hydro	5.14	6.86
Peak	129.54	Gas	51.03	59
		Solar	2	40.96

The sites chosen for installing the solar plants are given in [70]. It is assumed that all solar plants are geographically dispersed across the SE region. The solar generation is simulated in capacities ranging from 5% to 30% of the annual peak demand. This translates to a total solar capacity of 2.25 GW, corresponding to 5% penetration and 13.5 GW for 30% penetration. The individual PV modules are oriented at an azimuth of 180 degrees (south facing) and a tilt angle of 30 degrees. To simulate the solar generation, the study collects the hourly TMY3 solar insolation data of the representative sites and estimates the hourly clearness index values. The National Solar Radiation Database [62] contains high-resolution data for the state of Georgia and Alabama. Hence the study uses high-resolution data from seven test sites from 2010-2012 [73]. The transition probability matrix from the Milford area in Utah is chosen to synthesize the high-resolution clearness index for the representative sites across the SE region since the annual output of Utah most closely resembles that of Georgia. The solar output in minute intervals is estimated using PV_LIB][74]. The study assumes Canadian Solar CS5P-220M solar modules and Siemens SINVERT PVS 1401 UL inverters. The study utilizes annual simulations of the net generation and generation by source and net system demand of the SE region for 2020. The results of the generation allocation when flexible sources like gas generation are dispatched to meet the net load are shown in Figures 4.3, 4.4 and 4.5. The generation profiles shown represent the day with some transient cloud activity corresponding to the aggregate generation output shown in Figure 4.2.

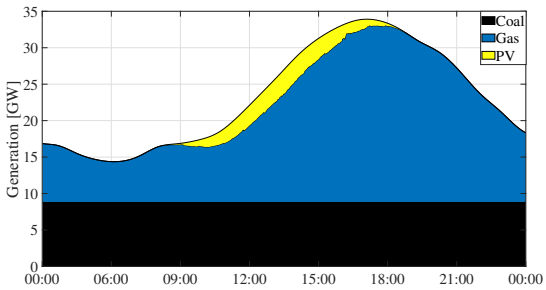

Figure 4.3: Generation Profile with 10% Solar

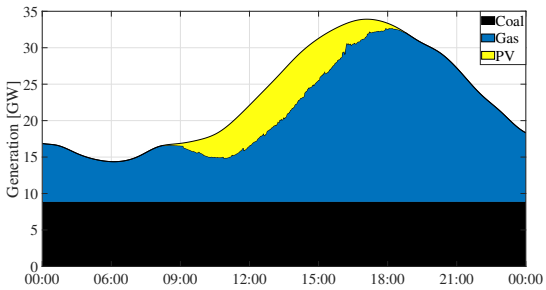

Figure 4.4: Generation Profile with 20% Solar

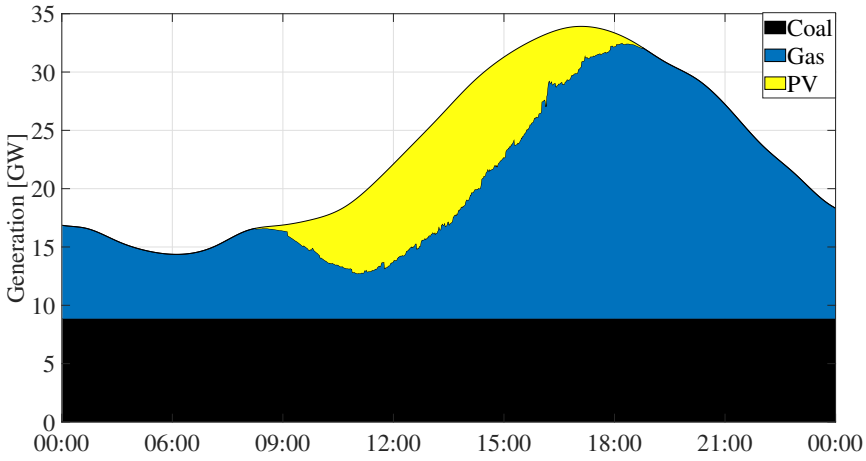

Figure 4.5: Generation Profile with 30% Solar

Since nuclear is the base unit with little or no impact on the emissions, resource allocation results are shown only for coal, gas, and solar generation. When the generation pool is assumed to consist of solar energy, the output of the gas turbine exhibits ramping characteristics, and the magnitude of the ramping increases with the increase in the penetration of solar generation. The results shown in Figures 4.3, 4.4 and 4.5 do not consider ramping constraints on coal since the load-following by coal is very expensive and large units may undergo significant damage due to fast ramping. The gas turbine follows the load and is assumed to have a fast ramping capability. The gas output increases when the solar production drops to compensate for the deficit. Similarly,

when the solar generation increases, the gas output ramps down to balance the load. The frequent ramping up and down of the more expensive gas along with the cost of solar generation has the impact of increasing the total cost of generation. As the share of solar continues to grow, at one point, solar will start pushing nuclear out. At that point, the system will need more coal and gas to balance solar, and the total CO2 emissions will start rising. Such a "renewable paradox" (i.e., increasing emissions as more RE is added to the system) is already noticed in Ukraine, where nuclear accounts for about 50% of power generation.

The annual cost of coal, gas and PV generation when no constraints are applied to the coal output is shown in Figure 4.6. The generation cost increases with penetration level since utility-scale photovoltaic is still the most expensive generation source compared to coal and gas. However, since the ramping constraints on coal are ignored, and coal is the least expensive resource in terms of the Levelized cost of energy, the presence of additional solar generation leads to an overall decrease in the more expensive gas generation to balance the load. Overall, the cost of generation

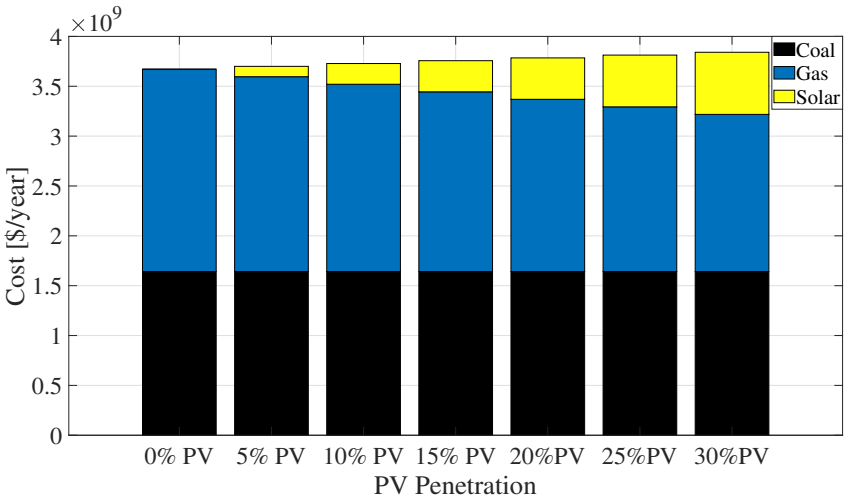

Figure 4.6: Generation cost without coal retirement

increases with the addition of solar to the mix. Figure 10 shows the annual generation costs when coal output is constrained and the amount of coal retired is equal to the peak solar generation at each penetration level [75]. The generation costs, in this case, are observed to increase faster than in the case where no ramping constraints were applied to the coal generation. Since coal is progressively retired, the amount of gas generation required to support the solar increases at each penetration level. As reported earlier, the overall generation costs do not exhibit an exorbitant rise. This can be attributed to the significant decline in the generation costs of utility-scale solar plants. As a

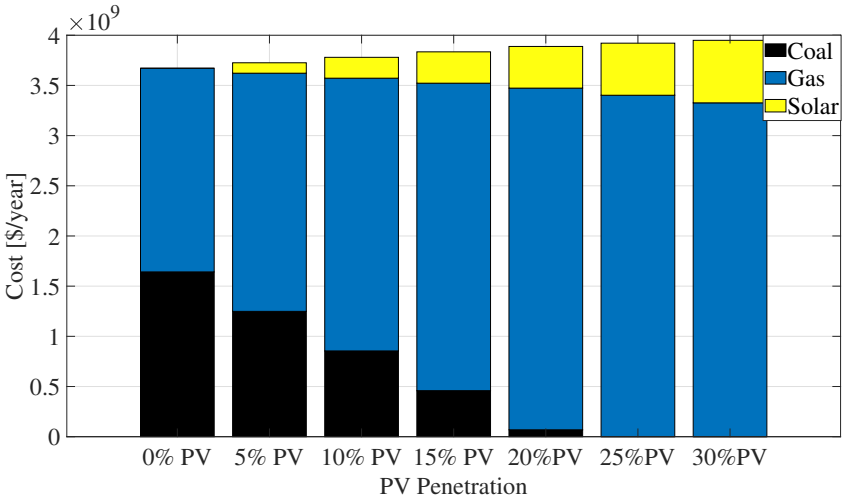

Figure 4.7: Generation cost with coal retirement

result, the relevance of the solar generation costs to act as a deterrent to the widespread adoption of solar generation is rapidly wearing down. What is concerning, however, is the economic impact of frequent ramping up and/or down of the gas turbines to compensate for the variability of the solar generation.

To make up for the rapid changes in the solar output, the gas spinning reserve is set equal to the maximum change in the solar generation for a given day. The spinning reserve is thus scheduled every minute to offset the changes in the solar output. This, however, results in increased emissions and ramping costs of the gas turbine. The relative ramping costs of gas at different penetration levels of solar can be visualized in the bar graph in Figure 11. Although the cost of gas ramping is small compared to the total fuel operating cost, it increases with additional solar deployment. This is expected since higher capacity solar will result in higher changes in the system load. From the emitted in the form of carbon dioxide in metric tons, this study used the emission data as reported by EIA [73]. The amount of emission is estimated based on the emission coefficients of each fuel type. Figures 4.9 and 4.10 show the amount of carbon emitted in metric tons for one day with and without coal retirement. When the ramping constraint on coal is ignored, the gas generation's carbon footprint decreases with the peak solar capacity increase (Figure 4.9). On the other hand, resulting range of emission costs as a function of solar penetration is shown in Figure 4.11 and

emissions perspective, deploying additional solar capacity reduces the number of pollutants and greenhouse gases by reducing the dependency on fossil fuels. To estimate the amount of carbon when coal is retired, as additional solar capacity is added to the system, the generation and carbon emissions of gas-fired plants increase with the increasing solar capacity (Figure 4.10). To estimate the costs related to carbon emissions and investigate the impact of the carbon tax policies, we estimate the emission costs for a range of carbon tax values in the interval [1, 100]$/MT. The

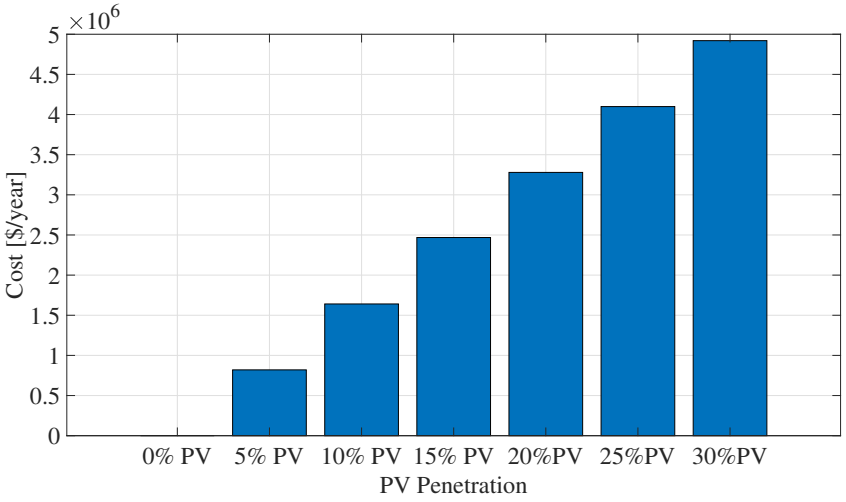

Figure 4.8: Annual Ramping cost of Gas Turbines

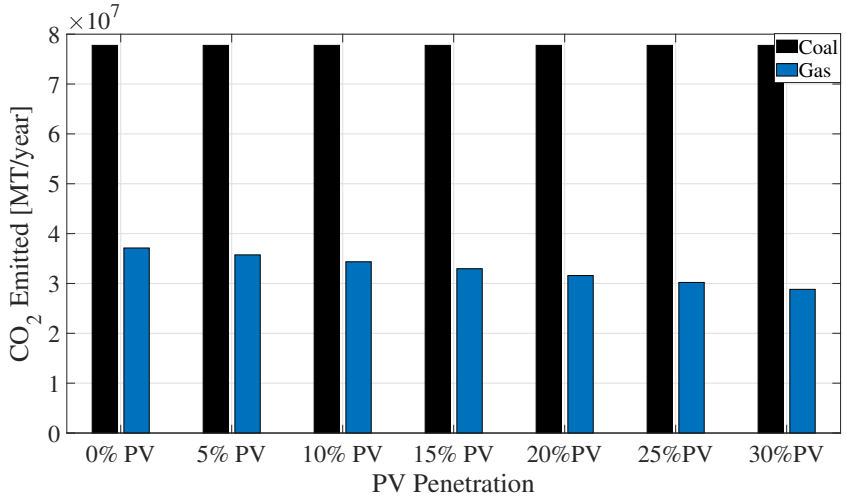

Figure 4.9: Carbon dioxide emitted in metric tonnes/year without coal retirement

4.12. The minimum value in the box plots of Figure 4.11 and Figure 4.12 for each penetration level corresponds to the carbon tax of $1/MT, and the maximum value corresponds to the carbon tax of $100/MT. When the output of coal-fired plants is held constant as the penetration of solar is progressively increased, the range of cost emissions narrows as the solar penetration increases. However, the constriction of the range of cost emissions is much more pronounced when the coal is progressively retired as opposed to not retiring the coat. This is mainly because the emission

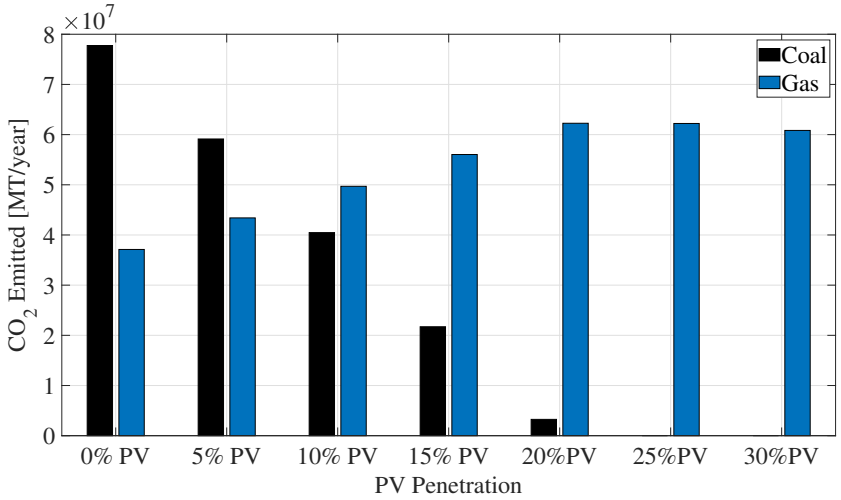

Figure 4.10: Carbon dioxide emitted in metric tonnes/year with coal retirement

coefficient of coal-fired generation is 2.21 pounds per kWh coal which compares to 0.91 pounds per kWh from gas-fired power generation. With subsequent retirement of coal (Figure 4.12) the

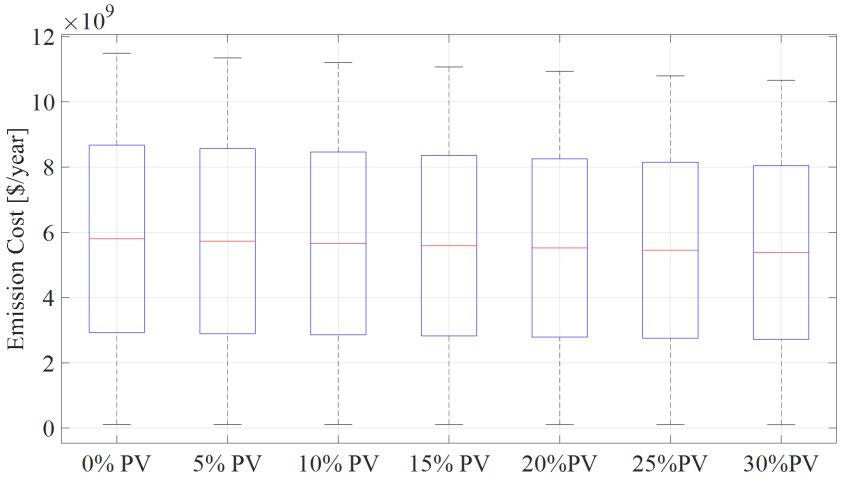

Figure 4.11: Range of cost emissions ($/yr) without coal retirement

cost of carbon emissions incurred to the utility decreases by more than 50% as the solar penetration level increases from 0% to 30% of the annual peak demand.

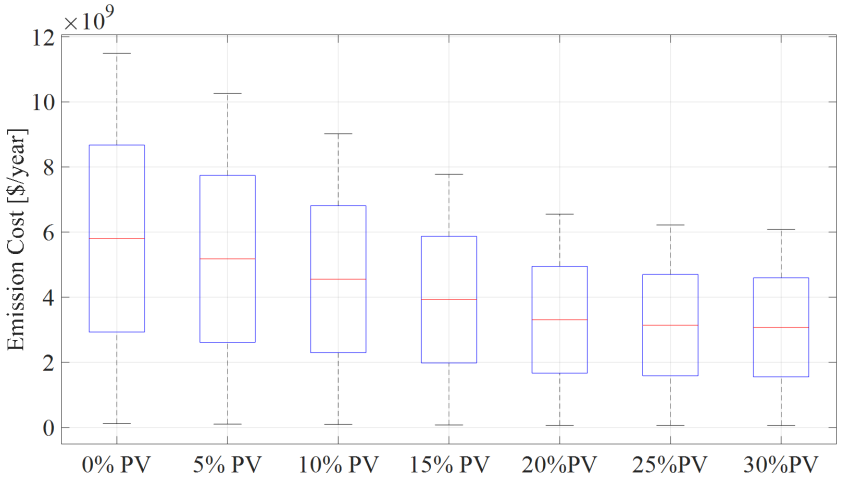

Figure 4.12: Range of cost emissions ($/yr) without coal retirement

4.6 Marginal Benefit of Carbon Abatement*

[5] This section presents a stochastic framework to estimate the profitability of investing in geographically dispersed utility-scale solar generation to offset the increased emission costs incurred due to the use of fossil-fuel-based energy resources like coal and gas. This investigation hinges on calculating a metric referred to as the annualized marginal benefit of carbon abatement (MBCA), defined as the net percentage change in the overall operational cost. More specifically

$$MBCA = \frac{C_T - C_T(PV, c_t)}{C_T} \times 100 \tag{4.5}$$

In (**??**) C_T is the total annual cost incurred when the overall system demand is met only with the non-intermittent sources such as coal-fired plants and natural gas plants and assuming no carbon tax policies exist. $C_T(PV, c_t)$ is the overall cost considering renewable energy such as utility-scale solar generation in the energy mix and assuming a carbon tax policy in effect. The overall cost would then be the sum of annual generation and annual emission cost. The units for both are $/year. The costs associated with the carbon emissions and the electric energy generation are

[5] Reprinted with permission from A. Peerzada, M. Begovic and D. Ostojic, "Carbon Tax and Utility-scale Solar Deployment," 2022 IEEE Power & Energy Society General Meeting (PESGM), 2022, pp. 01-05, doi: 10.1109/PESGM48719.2022.9917021

estimated by running optimal dispatch on the coal-fired plants and gas turbines while considering gas ramping costs in the optimization process.

4.6.1 Design of Uncertainty

To study the relationship defined in (4.5), we propose the use of Monte Carlo sampling of the input variables. For a fixed percentage of solar generation added to the existing energy pool, the input variables of the proposed computational model are the annual system demand profile and the carbon tax rate. Including a carbon tax rate is important to estimate the cost of carbon emissions and the overall generation cost since the Levelized cost of energy (LCOE) for coal-fired plants and gas turbines has a strong dependence on the carbon tax rate [76] . The use of Monte Carlo sampling entails that the output MBCA as defined in (4.1) be a function of the random variable $C_T(PV, c_t)$ that can assume a range of values. The first step involves identifying the underlying distributions of the input variables.

4.6.1.1 Annualized Load Profiles

To model the uncertainty in consumer demand, we propose using a probability mixture model in the form of a convex combination of a finite number of Generalized Gaussian distributions. The advantages of using a probability mixture model are twofold. First, the measured load data density functions usually have multi-modal characteristics, which can be better fitted with a mixture model rather than any specific distribution. Also, each demand category, whether residential, commercial, or industrial, has a unique parametrization that cannot be represented fairly by a single probability density function [16]. In the light of these observations, a general probability mixture model takes the following form.

$$f_Y(y_i;\Psi) = \sum_{j=1}^{M} \pi_j f_{j,Y}(y_i, C = j, \Theta_j) \tag{4.6}$$

In (4.6), $Y = y_i$ is the measured load data, M is the number of mixture components, π_j is the weight of the j^{th} mixture component subject to $\pi_j > 0 \ \forall j \in \{1,2,...,M\}$ and $\sum_{j=1}^{M} \pi_j = 1$ and each $f_{j,Y}$ is a density function parameterized by Θ_j. $\Psi = \{\pi_j, \Theta_j; j = [1,2,...,M]\}$ is the overall parameter vector. For a univariate Generalized Gaussian mixture model (GGMM), the parameter

vector Ψ consists of weight, mean, scale and shape parameters of each mixture component. The log-likelihood function of the density $f_Y(y_i;\Psi)$ in the mixture model in (4.6) is

$$\log L(\Psi|Y) = \sum_{i=1}^{N} \log \sum_{j=1}^{M} \pi_j f_{j,Y}(y_i|C=j,\Theta_j) \tag{4.7}$$

The overall parameter vector Ψ can be estimated by maximizing the log-likelihood function and is given by the solution of

$$\Psi^* = \arg\max_{\Psi} \sum_{i=1}^{N} \log \sum_{j=1}^{M} \pi_j f_{j,Y}(y_i|C=j,\Theta_j) \tag{4.8}$$

Unfortunately, the optimization of (4.8) is difficult because of the log $(\sum)$ term and the usual method of maximum likelihood estimation (MLE) cannot be used. The parameter vector can however be estimated via Expectation-Maximization (EM) algorithm by considering the log-likelihood function (4.7) to be a random variable. This can be achieved by introducing binary-valued hidden variables $Z = \{z_{ji}\}_{i=j=1}^{i=N,j=M}$ that carry the information about which component density "generated" each data item $Y = \{y_i\}_{i=1}^{N}$. Thus we assume that $z_{ji} = 1$ if the data sample y_i was generated by the j^{th} mixture component and $z_{ji} = 0$, otherwise. More details on the EM algorithm can be found in [16].

We propose using a Generalized Gaussian distribution to model the measured load data instead of a Gaussian distribution. A random variable Y is said to have a Generalized Gaussian distribution (GGD) if the density function of Y has the form

$$f_Y(y;\mu,s,p) = \frac{p}{2s\Gamma\left(\frac{1}{p}\right)} \exp\left[-\frac{|y-\mu|^p}{s^p}\right] \tag{4.9}$$

In (5.23) $p \in \mathbb{R}^+$ is the shape parameter and controls the shape of the distribution, $s \in \mathbb{R}^+$ is the scale parameter, $\mu \in \mathbb{R}$ is the location parameter and $\Gamma(.)$ is the gamma function. The GGD has an advantage over the Gaussian distribution (GD) in representing the statistical uncertainty of electric demand because of the extra shape parameter. The presence of the shape parameter makes a GGD

more flexible which in the context of load modelling means that it can approximate the peak load behavior better than a GD. For a GGMM the application of E-M algorithm to the measured load data yields the following update equations for the model parameters $\Psi^{(n+1)} = \{\pi_j, \mu_j, s_j, p_j; j = [1,2,...,M]\}$ given the current estimate Ψ^n and the measured load data Y

$$\pi_j^{(n+1)} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{E}\left[z_{ji}|Y,\Psi^{(n)}\right] \tag{4.10}$$

$$\sum_{i=1}^{N} \mathbf{E}\left[z_{ji}|Y,\Psi^{(n)}\right] p_j^{(n)}|\mu_j^{(n+1)} - y_i|^{p_j^{(n)}} = 0 \tag{4.11}$$

$$s_j^{(n+1)} = \left[\frac{\sum_{i=1}^{N} \mathbf{E}\left[z_{ji}|Y,\Psi^{(n)}\right]}{\sum_{i=1}^{N} \mathbf{E}\left[z_{ji}|Y,\Psi^{(n)}\right] p_j^{(n)}|\mu_j^{(n+1)} - y_i|^{p_j^{(n)}}}\right]^{-\frac{1}{p_j^{(n)}}} \tag{4.12}$$

$$\sum_{i=1}^{N} \mathbf{E}\left[z_{ji}|Y,\Psi^{(n)}\right] \kappa = 0 \tag{4.13}$$

In (5.34), κ equals

$$\kappa = \frac{1}{p_j^{(n+1)}} + \frac{\psi\left(1/p_j^{(n+1)}\right)}{\left(p_j^{(n+1)}\right)^2} - \left(\frac{|y_i - \mu_j^{(n)}|}{s_j^{(n)}}\right)^{p_j^{(n+1)}} \left(\log|y_i - \mu_j^{(n)}| - \log s_j^{(n)}\right) \tag{4.14}$$

In (5.38) $\psi(.)$ is the digamma function defined as $\Gamma'(g)/\Gamma(g)$. The update equations for the μ and p are nonlinear and we use the Newton-Raphson method to obtain a numerical solution.

4.6.1.2 Carbon Tax Rates

The LCOE for coal-fired plants and gas turbines depends on the carbon tax rate among other factors. If $'t'$ is the year in which the production takes place then the LCOE can be calculated using the following formula [76]

$$LCOE = \frac{\sum_t (I_t + M_t + c_t + D_t + F_t)(1+r)^{-t}}{\sum_t E_t (1+r)^{-}t} \tag{4.15}$$

In (A.9) I_t refers to the capital cost in year t, M_t is the operation and maintenance cost in year t, c_t is the carbon cost in year t, D_t is the decommissioning and waste management cost in year t, F_t is the fuel cost in year t, $(1+r)^{-t}$ is the real discount rate corresponding to the cost of capital and E_t is the amount of electric energy produced annually in MWh. In this work, we compute the LCOE of coal and gas plants as a function of the carbon tax rates while making standard assumptions for other factors. The carbon tax rates are sampled from a uniform distribution $c_t \sim [1-100]$ ($ per Metric Ton), which ensures that each value of the carbon tax rate has the same probability of occurrence. The uniform distribution is a good approximation in situations where the prior information about a parameter is unavailable. The LCOE of coal-fired and gas plants as a function of carbon tax is shown in Figure 4.13

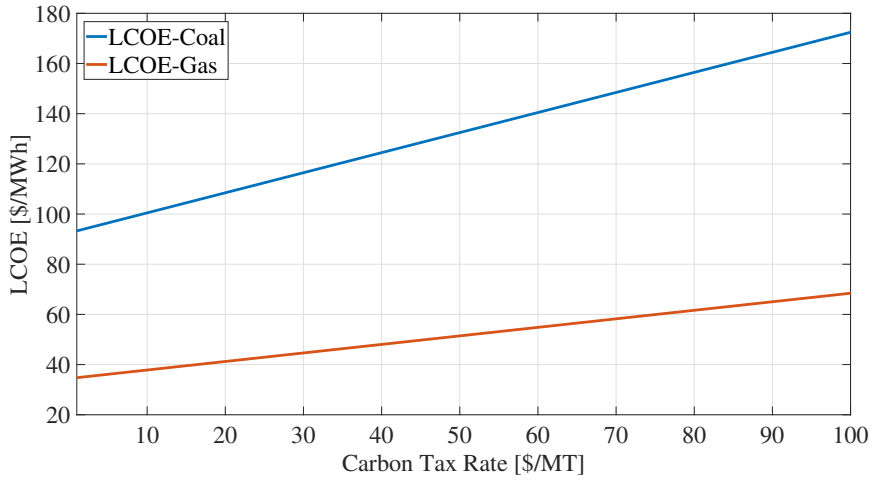

Figure 4.13: LCOE of coal-fired and gas turbines as a function of Carbon Tax

4.6.1.3 Solar Generation Modeling

The output of a solar photovoltaic system is subject to rapid changes when viewed at sub-hourly or sub-minute time scales. Thus, the solar generation model must be able to faithfully reproduce such rapid excursions in the power output. In this work, we use the Markov weather model, as described in chapter 3, to generate synthetic weather profiles with a temporal resolution of one minute. A key component of the model is the estimation of the clearness index, which is defined

as the ratio of measured irradiance at the earth's surface and the irradiance that corresponds to cloudless conditions. The hourly-averaged Typical Meteorological Data (TMY3)[62] is used as an input to the Markov weather model, which, when combined with transition probabilities that are generated at sub-hourly time scales, results in high-resolution solar irradiance profiles that show seasonal variation.

4.7 Results of Monte Carlo Simulations*

[6] The southeast region (SE) of the United States, which comprises the states Georgia, Alabama, and parts of Mississippi, is chosen as a case study to investigate the impact of carbon tax rates and installed solar capacity on the marginal benefit derived from adding renewables to the existing generation pool. We consider two operational scenarios. In the first scenario, the output of the coal-fired plants is assumed constant for all possible carbon tax rates and installed solar capacity. In the second scenario, the output of the coal-fired plants is adjusted based on the peak solar generation. This effectively entails retiring certain coal plants as the percentage of solar generation in the overall energy mix increases. The coal retirement strategy is based on the findings presented in [75]. The study uses the unsubsidized average LCOE for coal, gas, and solar generation. The energy consumption and generation profile data for the SE region was obtained from the EIA grid monitor website [4]. In 2020, the annual energy consumption in the SE region was 243,140 GWh with a peak load of 46 GW, while the combined annual generation from all the sources listed in Table 2.2 was 252,490 GWh. The solar plants in the SE region are geographically dispersed across seven representative sites as given in [77]. A range of installed solar capacity from 5% of the peak annual load to 30% in steps of 5% is considered. More specifically, a 5% penetration represents approximately 2.25 GW of solar capacity, whereas a 30% equals approximately 13.5 GW of solar capacity. The transition probabilities necessary for the Markov weather model are extracted from high-resolution irradiance data from National Solar Radiation Database (NSRD) [62]. The solar output at sub-hourly time scales is calculated using PV_LIB [74] assuming Canadian Solar CS5P-

[6]Reprinted with permission from A. Peerzada, M. Begovic and D. Ostojic, "Carbon Tax and Utility-scale Solar Deployment," 2022 IEEE Power & Energy Society General Meeting (PESGM), 2022, pp. 01-05, doi: 10.1109/PESGM48719.2022.9917021

220M solar modules and Siemens SINVERT PVS 1401 UL inverters.

The Monte Carlo simulation is initiated by repeatedly sampling the input parameters from the underlying distributions. To generate a synthetic annualized random load profile, the GGMM is fitted to the base-load profile data of the SE region as described in chapter 2. The generation output of coal-fired and gas plants is calculated by feeding the random samples of the input parameters to the optimal dispatch model. The overall cost of generation is calculated using the data listed in Table 4.1. The LCOE of coal-fired and gas plants is calculated as a function of a carbon tax, assuming a real discount rate of 7%, a standard capacity factor of 85%, a 30-year lifetime for gas plants, and a 40-year lifetime for coal-fired plants. A complete description of the individual costs used in (A.9) is given in [76].The emission costs are calculated by assuming an emission cost-coefficient of 2.21 pounds per kWh for coal-fired plants and 0.91 pounds per kWh for gas turbines [66]. Since the time horizon for the Monte Carlo simulation is one year, the overall operational costs are expressed in $/year.

4.7.1 MBCA manifolds

The marginal benefit of carbon abatement represents the net profit or loss incurred by the utility due to adding renewable energy, such as geographically dispersed utility-scale solar as considered in this work, to the existing generation pool. To calculate the percentage change in the overall cost of operation, we define a base case where the generation resource pool consists only of non-intermittent sources such as nuclear, coal, gas, etc., and a carbon tax rate of $\$0/MT$. The Monte Carlo simulations are run for six levels of solar penetration, and for each penetration level, carbon tax values in the range $[1-100]$\$/MT are considered. A total of 100 Monte Carlo runs are simulated, generating 10000 output samples of *MBCA* for each solar penetration level. For a single MC run, the dependence of *MBCA* on the installed solar capacity and the carbon tax rate for both scenarios is shown in Figure 4.14 and Figure 4.15 When viewed as a function of carbon tax alone the *MBCA* decreases with the increase in carbon tax. This is true for every solar penetration level. The important difference between the two scenarios is that while in the case of no coal retirement, the rate of decrease in the *MBCA* is the same for each solar penetration level, it is different for

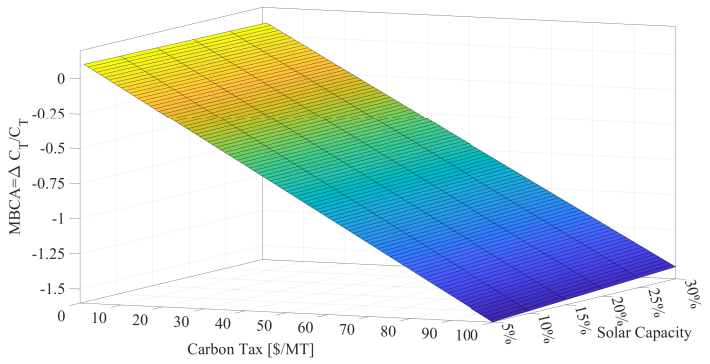

Figure 4.14: MBCA manifold without Coal Retirement

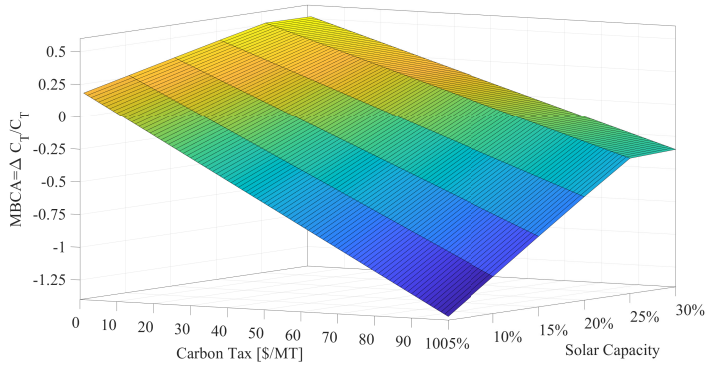

Figure 4.15: MBCA manifold with Coal Retirement

different solar penetration levels in the case when coal plants are retired. The manifold's shape in Figure 2.2 suggests that the *MBCA* decreases faster for low solar penetration levels than it does for higher solar penetration levels. As a function of solar penetration level alone, the *MBCA* decreases as solar penetration increases for a low carbon tax and exhibits a reverse trend for higher carbon tax values in the first scenario. However, in the second scenario, the *MBCA* increases with the increase in solar penetration for all carbon tax values reaching a maximum at about 25% solar penetration and decreasing slightly afterward. Furthermore, the rate of increase as a function of solar penetration is faster at high carbon tax values than at lower carbon tax values. It is important to note that an increase in the *MBCA* when $MBCA > 0$ means increasing profit, and a decrease in *MBCA* when $MBCA < 0$ means increasing loss.

4.7.2 MBCA distributions

The distributions of *MBCA* for each solar penetration level is shown in Figure 4.16 for the scenario where coal is not retired and Figure 4.17 when coal is subsequently retired.

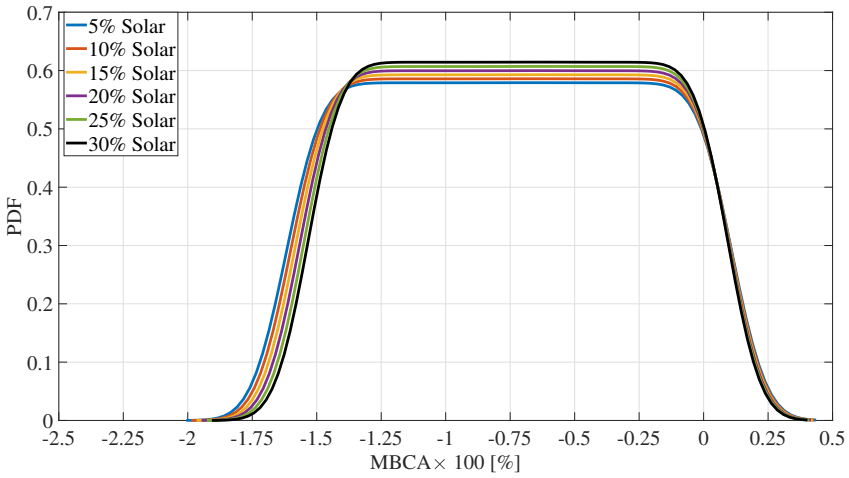

Figure 4.16: MBCA Distributions without Coal Retirement for 100 Monte Carlo Runs

Each *MBCA* distribution as shown in Figure 4.16 and Figure 4.17 corresponds to a certain solar penetration level and carbon tax rates in the range [1-100]$/MT. For both the scenarios, the distributions appear to be shifting to the right as the installed solar capacity changes from 5% to 30% of the peak annual demand. This suggests that increasing the solar generation capacity is more likely to increase the profits (or decrease the losses depending on whether $MBCA > 0$ or $MBCA < 0$ respectively) and lower the overall operational cost regardless of the carbon tax value. A positive increasing value of the *MBCA* means increasing profits and decreasing operational costs, while a negative decreasing value means increasing losses with increasing operational costs. It is important to note that adding utility-scale solar to the energy mix does not guarantee an overall increase in profits, which also depends on the carbon tax value.

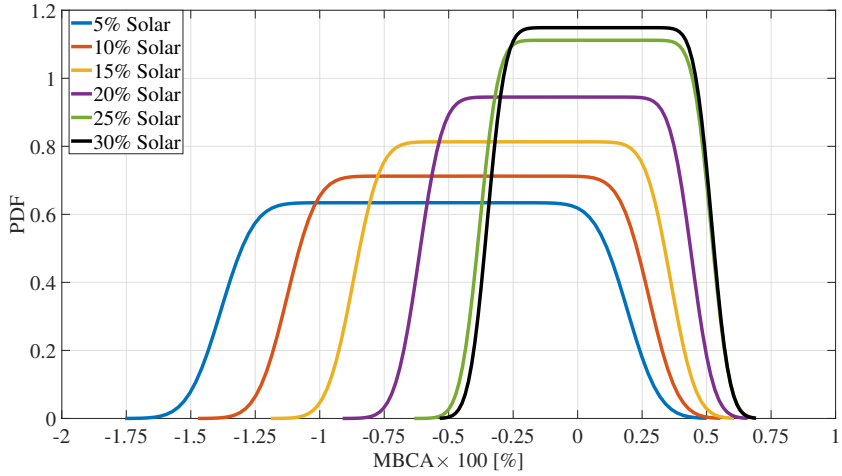

Figure 4.17: MBCA Distributions with Coal Retirement for 100 Monte Carlo Runs

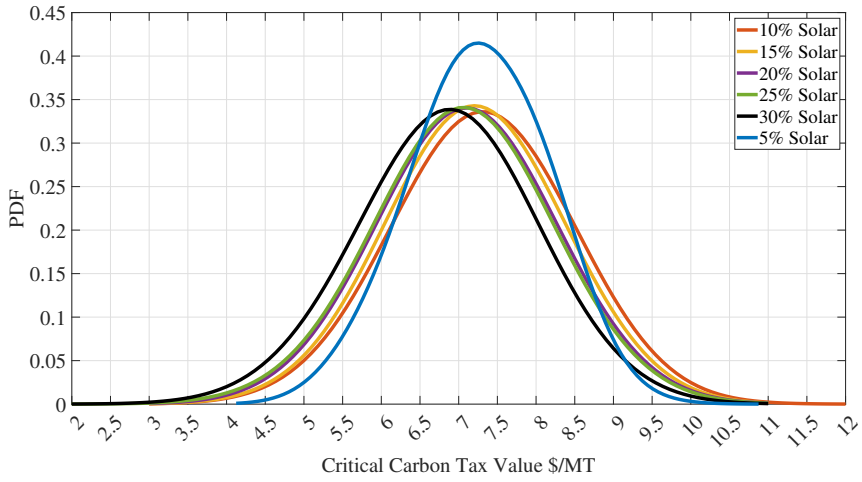

Figure 4.18: Distribution of Critical Carbon Tax with No Coal Retirement

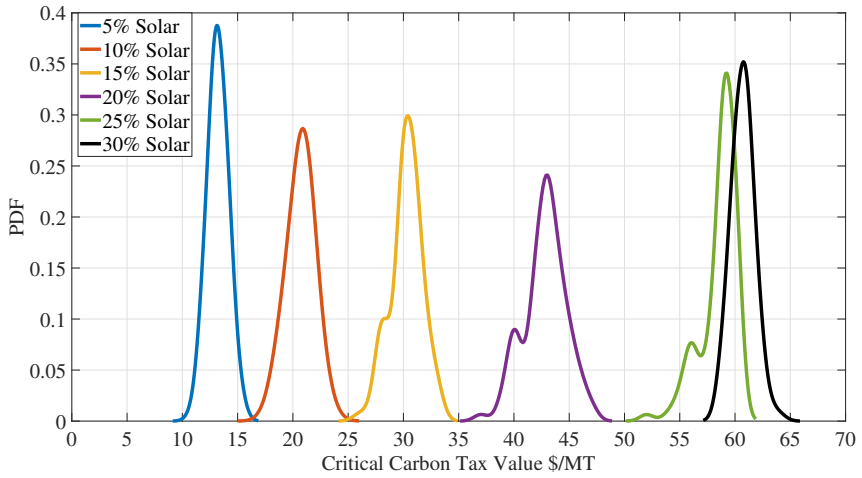

Figure 4.19: Distribution of Critical Carbon Tax Rate with Coal Retirement

Figure 4.18 and Figure 4.19 show the distribution of critical carbon tax rates for different solar penetration levels with and without coal retirement. A critical carbon tax rate is defined as the carbon tax rate at which the $MBCA <= 0$. This means that a carbon tax rate greater than the critical carbon tax rate will increase the overall cost of operation, leading to a utility facing higher economic losses. The value of the critical carbon tax rate is vital from the policy-making point of view. Interestingly, the expected value of the critical carbon tax changes with the solar penetration level in both scenarios. This suggests using a dynamic policy that adjusts the carbon tax policy depending on the installed solar capacity in the network. More specifically, the expected value of the critical carbon tax decreases as the installed solar capacity increases in the scenario when coal is not retired and vice versa when coal is retired. Table 4.2 lists the data pertaining to the Figures 4.16,4.17,4.18 and 4.19.

The acronyms "CNR" and "CR" in Table 4.2 refer to the two scenarios of no coal retirement and coal retirement respectively. The $\mathbf{E}[MBCA]$ is the expected value of the marginal benefit for each solar penetration level. Table 4.2 shows that coal retirement is essential for avoiding losses due to carbon tax, as E(MBCA) changes from -76.5% to -73.1% in the CNR case, irrespective of a significant increase of solar penetration. On the other hand, in the case of coal retirement (CR), it is necessary to increase solar penetration above [20%] to avoid losses and make E(MBCA) positive. In both scenarios, the critical carbon tax value plays an important role in deciding the profitability of investing in utility-scale solar generation.

Table 4.2: Profit and Loss Probability

Solar	$P\{MBCA \leq 0\}$		$P\{MBCA > 0\}$		$\mathbf{E}[MBCA]$	
	CNR	CR	CNR	CR	CNR	CR
5%	0.9125	0.8764	0.0875	0.1236	-76.54%	-59.53%
10%	0.9126	0.7890	0.0874	0.211	-75.86%	-42.56%
15%	0.9249	0.6987	0.0751	0.3013	-75.17%	-25.50%
20%	0.9250	0.5741	0.075	0.4259	-74.50%	-8.72%
25%	0.9250	0.4175	0.075	0.5825	-73.82%	+6.77%
30%	0.9250	0.4006	0.075	0.5994	-73.14%	+8.47%

4.8 Conclusions

The results presented here demonstrate a complex interplay between several factors and thus call for a more nuanced approach toward policy-making and regulation. A major implication of this work is that the targets set for a specific technology to decarbonize the power grid cannot be applied as a "one size fits all" across different adoption levels of the same technology and different power utilities. The economic equation that governs the profitability and hence the adoption of renewable energy such as utility-scale solar generation is heavily impacted by regional specific factors such as the energy mix, LCOE, system load, and proposed carbon tax value, and hence each utility is faced with a specific task of finding the optimum penetration of the renewable energy that would maximize the benefits with no adverse impact on system reliability.

Interesting future work in this direction involves using energy storage devices to facilitate the integration of utility-scale solar generation instead of natural gas turbines. Notably, carbon emissions are associated with energy storage devices, especially battery energy storage systems (BESS). It would be an interesting study to compare the performance characteristics of BESS versus gas turbines in terms of total generation cost and emission savings.

5. TOPOLOGICAL UNCERTAINTIES IN POWER SYSTEMS

The emergence of distributed generation in the low voltage distribution networks has led to new challenges in regulating feeder voltages. Of particular significance is the variability associated with the photovoltaic power and its impact on the operation of some mechanically switching voltage regulating equipment such as On-load Tap Changers (OLTCs) and switchable capacitor banks. This chapter describes the effects of increased penetration of distributed generation, in particular, photovoltaic power and a sustained load buildup on the operational activity of legacy voltage control frameworks such as on-load tap changers (OLTCs) and switched capacitor banks. Actual feeder load profile and high-frequency solar irradiance data have been used with varying levels of PV penetration during the time period, which spans an entire year. With the inclusion of some justifiable assumptions, it is concluded that the increased penetration of photo voltaic power adversely affects the operational lifetime of voltage-regulating equipment. In particular, there is a great need for a thorough analysis on the mechanism of accelerated wear and tear of devices like on-load tap changers and capacitor banks under conditions of rapid voltage fluctuations. Such an analysis necessitates the development of lifetime models of such devices to study the impact of increased stress, whether electrical or mechanical, on the shortening of device lifetime.

5.1 Background*

[1] The recent expansion of distributed energy resources, particularly photovoltaic power, in the distribution networks has resulted in several voltage regulation issues. With penetration projected to only increase in the future, network operators will have to contend with an increasing surge in the range of problems concerned with regulating voltage and power on distribution feeders. For example, in San Diego Gas and Electric distribution territory, PV installations accounted for 617 MW of peak load from 93000 installations at the end of June 2016. In August 2013, these

[1]Part of the data reported in this chapter is reprinted with permission from Peerzada, Aaqib, et al. "Impact of large distributed solar PV generation on distribution voltage control ." Proceedings of the 52nd Hawaiian International Conference on System Sciences | 2019 .

numbers were 175 MW from 24000 PV installations. This represents a substantial increase in the PV penetration over the course of roughly three years. The high penetration of PV can be attributed in large measure to the policies pursued by the federal as well as many of state governments with an increasing concern for climate change and declining PV system costs, among other factors.

The fundamental concern about photovoltaic generation is the intermittency involved which, when combined with other factors such as a fluctuating load profile, can induce irregularities not only in the voltage on the distribution feeder but also in the operational activity of several electro-mechanical devices which are designed and placed on the feeder for the purpose of voltage regulation. This places pressure on the distribution utilities, which are obligated under the ANSI standards to provide voltage within a dead band of ± 5V from the nominal distribution voltage. While the impact of such distribution generation in general and photovoltaic penetration, in particular on the voltage profile in the low voltage networks, has been documented in several studies, the impact of such solar intermittency on the operational life cycle of voltage regulating equipment has not been thoroughly investigated [78],[79]. These devices are mechanical in nature and rely on local signals to perform the necessary action. For example, the control setting of voltage regulators includes the time delay (TD), voltage set-point, and bandwidth. Similarly, for capacitor banks, discrete control of the reactive power output is implemented to improve the substation power factor. Furthermore, the voltage regulators in most practical implementations make use of line drop compensation to regulate the secondary bus voltage. In any case, the operation of these devices involves the movement of a mechanical switch to output the desired voltage value and the power factor.

5.2 Failure Statistics of On-Load Tap Changers*

[2] The transformer reliability working group, founded in 1975, launched a survey in 1978 aimed at studying the lifetimes of transformers and reactors. The survey is representative of the countries in CIGRE SC 12 [80]. The survey compiled data from more than 1000 failures between 1968-1978.

[2]Part of the data reported in this chapter is reprinted with permission from Peerzada, Aaqib, et al. "On Accelerated Aging of Mechanical Assets in Distribution Systems with Renewable Generation." Proceedings of the 53rd Hawaiian International Conference on System Sciences | 2020 .

Various distinctions were made to arrive at the failure rates of transformer populations. Some of them include the operating voltage, transformer type, age, and the presence of OLTC. Also, a distinction was made between forced outages and scheduled outages. Forced outages necessitate the transformer disconnection, while the required maintenance could be planned later for scheduled outages. The survey concluded that the failure rate of a transformer generally increased with the voltage due to the reduced reliability of the winding paper insulation. When the failure rates were estimated in terms of device components, it was found that OLTCs contributed to the transformer failure more than any other component for distribution transformers. In fact, out of the 702 failures in the substation transformers, 691 occurred in transformers equipped with OLTCs. The data in [80] suggests that more than 40 % of substation transformers failed due to failure of OLTC, 19% failed due to windings, and roughly about 12% failures resulted from the tank and dielectric fluid. In terms of the original failure mechanisms, mechanical failures account for roughly 55% of the total failures, followed by dielectric and thermal failures. Mechanical failures also resulted in forced outages much more than thermal and dielectric-based failures. The survey concluded that the failure in substation transformers occurred mostly due to the failure in the OLTCs. [81] presents the failure statistic of the 11 commonwealth independent nations. The data set includes failure modes of 5000 large power transformers with a power rating of 100 MVA and above. The primary cause of transformer failure for large power transformers was weak construction. Inadequate maintenance and low-quality repair were also determined to be the leading causes contributing to the failure rate of power transformers. In [82], the failure statistics of the Eskom network in South Africa are presented. The voltages from 88 kV to 765 kV and the transformers with power ratings between 20 MVA and 800 MVA were considered. These voltage and power ranges include distribution substation transformers and transmission power transformers. In the Eskom network, the study in [82] identified six failure modes of transformers. These include lightning, core problems, and tap-changer failures. General aging, short circuit problems, and others. The study concludes that most distribution substation transformers' failures resulted from tap changer-initiated failures. In large power transformers, lightning and insulation problems at higher voltages contributed to

most failures. The survey results in [82] are given in Figure 5.1 . As of now, it seems that tap

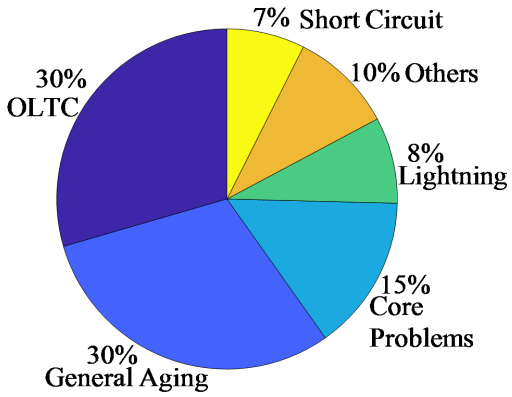

Figure 5.1: Eskom Network Failure Transformer Statistics

changer-enabled transformers will continue to play a major role in voltage regulation for long radial distribution feeders. As such, it is essential to study the reliability of a load tap-changer under adverse operating conditions. In recent years, many advances have been made in vacuum-type load tap-changers. However, most of the substation transformers are equipped with oil-type load tap changers. Oil-type load tap-changers are negatively impacted by a higher frequency of operation due to the formation of a carbon layer on the contacts of a tap-changer. As the frequency of operation increases, it is more likely to encounter asynchronous operation between the diverter switch and the tap selector. The oil-type load tap changers generally require maintenance interventions between 50,000 and 100,000 operations. Given a transformer lifespan of 40 years, roughly five maintenance intervals are required for oil-type OLTCs. The vacuum-type OLTCs on the other hand, require fewer maintenance interventions. In [30], a statistical model for the tap-changer degradation is presented. The Weibull parameters $\alpha$and β for the tap-changer are 109 years and 2.4. The mean time to failure (MTTF) is reported as 97 years.

5.3 Impact of High Capacity Solar Generation on OLTCs*

[3] To quantify the effects of high PV penetration on distribution feeders, IEEE 34 node feeder was chosen in this study owing to its large radial length and high voltage imbalance. It represents a typical rural distribution network with single and three-phase laterals which can incorporate distributed generation. Voltage compensation is provided by three-phase online capacitor banks and two voltage regulators. The total feeder length is 94 kilometers, and a 69/24.9 kV transformer is used to step down the primary voltage. Actual phase impedance values are used to model the three-phase and the single-phase overhead lines. The substation is rated at 2500 kVA and the system base active load is 1769 kW. This includes a combination of the spot loads and the distributed loads. The spot loads make up 1077 kW of active load and 677 kVAr of reactive load on the system, while the distributed loads comprise of 722 kW of base active load and 367 kVAr of base reactive load. In addition to the main substation transformer, there is an in-line step-down transformer connected between buses 832 and 888. The secondary voltage of this transformer is 4.16 kV. The test system comprises two three-phase capacitor banks connected to buses 844 and 848. The reactive injection from capacitors is used for power factor correction of the source. The one-line diagram of the test system is given in Figure 5.2 Voltage Regular 1 is located between buses 814 and 850 and provides voltage regulation on the secondary side with a dead band of 2.0 V. In contrast, the second regulator connects buses 832 and 852 with the exact specifications for the dead band. Three-phase capacitor banks rated at 300 kVAr and 450 kVAr are connected at buses 844 and 848 and operate at a line-to-line voltage of 24.9 kV. The ratings of the capacitors are decided by the amount of reactive power delivered by the substation transformer. The lines in the system are three-phase overhead and single-phase overhead with varying degrees of imbalance. Loads on the feeder are modeled as three-phase (balanced or unbalanced), spot or distributed (single phase or three phases). Three-phase loads can be connected in wye or delta, while single-phase loads are connected line to ground or line to line. Various models are used for modeling loads,

[3] Part of the data reported in this chapter is reprinted with permission from Peerzada, Aaqib, et al. "Impact of large distributed solar PV generation on distribution voltage control ." Proceedings of the 52nd Hawaiian International Conference on System Sciences | 2019 .

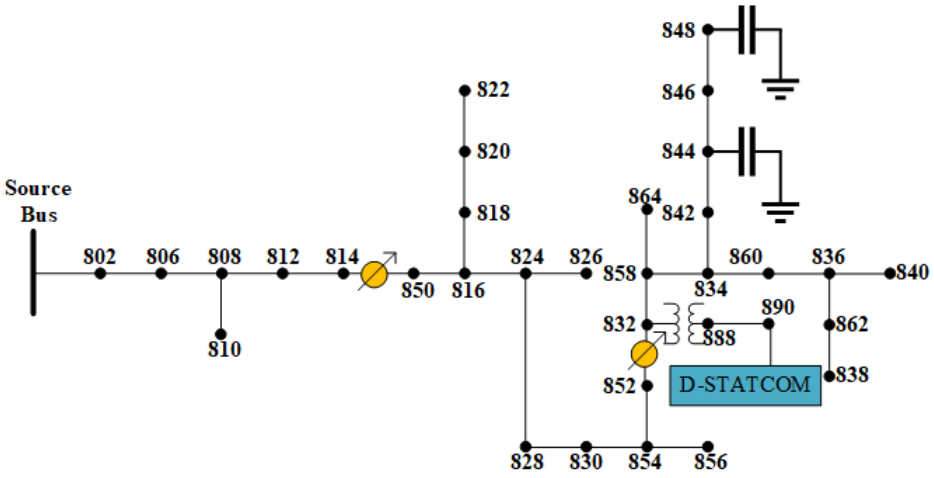

Figure 5.2: IEEE 34 Bus Test Feeder with Shunt Connected D-STATCOM

including constant P, constant PQ, constant Z, constant I, and exponential loads. The various components of the test system are modeled in OpenDSS, and the data processing is done in MATLAB. OpenDSS models an electrical component of a circuit by creating a primitive admittance matrix of the element, and the system admittance is obtained by combining all the elemental Y matrices.

The substation transformer is modeled as a three-phase, two-winding transformer with the operating primary voltage of 69 kV. There are 32 tap positions with uniform leakage impedance on the LV and HV sides. The tap positions need to take care of the daily fluctuations in the voltage. They allow the ratio to vary in the range of ±10 % with each tap change resulting in a 0.00625 p.u change in voltage. The two voltage regulators were modeled as single-phase autotransformers. One possible restriction in OpenDSS is that all conductors at a circuit element terminal must be connected to the same bus. This restriction, however, does not apply to nodes in a bus. OpenDSS allows a regulator control device that monitors the tap positions on the winding of the autotransformer. The two regulators are connected in wye configuration with the tap position of winding two controlled by the regulator control element. The PT ratio is set as 120, and the CT ratio is set as 100. The desired voltage is set as 122 V, and a band of 2V is selected. OpenDSS utilizes a combined model for the PV array and the inverter. The model is valid for a step size of greater than 1 second. The model consists of a PV array, the output of which is fed to an inverter with

a pre-defined efficiency curve. The inverter's efficiency will depend on theV_{MPP}, which refers to the voltage at the maximum power point. However, OpenDSS accepts only one efficiency curve among a family of possible curves. As seen from the rest of the circuit, a PV system appears as a power conversion element, similar to a generator or a load, which generates or consumes power according to some function. The PV active power output depends on the value of irradiance specified, the temperature and rated peak power P_{MPP}, at the maximum power point. P_{MPP}, specified by the user, is defined at standard test conditions: an irradiance of 1kW/m2 and a temperature of 25 C. The reactive power is specified separately, either as a fixed power factor or fixed kVAr values. Unity power factor control is implemented in the simulation. The panel output is calculated using the following equation

$$P_{kW}(t) = P_{MPP}(STC)Irr_{base}Irr(t)T_f(t) \tag{5.1}$$

Where $P_{kW}(t)$ is the output from the PV panel at a given time, $P_{MPP}(STC)$ is the power at the maximum power point defined under standard test conditions ($1kW/m^2$ irradiance and 25 C temperature), Irr_{base} is the base value of irradiance usually taken to be $1kw/m^2$, Irr(t) is the current irradiance, and $T_f(t)$is the current temperature factor which is interpolated from the temperature curve defined by the user. The panel output is multiplied by the inverter's efficiency to yield the PV output power, which is fed to the feeder.

$$P_{PV}(t) = \eta_{inv}P_{kW}(t) \tag{5.2}$$

5.3.1 Snap Shot Simulation Results

The test feeder was simulated using a snapshot simulation and a quasi-time series simulation. The snapshot simulations are important to assess some feeder properties and look for voltage violations along the length of the feeder. For the snapshot simulation, three different system conditions were studied. In the first case, the feeder was stripped of any voltage compensation. In the second case, capacitor compensation was provided to improve the voltage profile and the source power factor. It was seen that the reactive power injection by the capacitor banks corrected the voltage at

some of the buses on the feeder, and the source power factor was also improved. In the third case of snapshot simulation, voltage regulators were connected at locations that exhibited depressed voltages even with capacitor compensation. In the final case of snapshot simulations, an annual load growth of 3% was assumed on the feeder, and the feeder voltage and the voltage regular tap positions were simulated at the end of the tenth year. For the lack of space, the results for the final two cases are presented here since the main focus of this study is the operational irregularities observed in the tap changing devices due to fluctuations in the system load and the variable generation induced by PV penetration. To correct the low voltage problems, voltage regulators modeled as single-phase autotransformers are installed along the feeder. Since the minimum acceptable voltage on the feeder is 0.95 p.u, the voltage profile of the buses is scanned, and the first bus, as measured from the substation, where the voltage violation is observed, is chosen for the placement of the voltage regulator. This correction is be made by installing a voltage regulator at Bus 814. With the regulator installed and the tap settings determined by the Regulator Control device, the voltage profile is greatly improved. The voltages at Bus 814r seem to satisfy the voltage constraint. However, the voltage at bus 890 still violates the first constraint. To raise the voltage in the 888-890 lateral, it is necessary to boost the primary voltage of the inline transformer 24.9/4.16 kV. To accomplish this, the compensator settings are made $R = 2.5$ and $X = 1.5$, and the desired voltage is set at 122 V on a 120 V scale. This installation should boost the primary side voltage of the inline transformer by causing the regulator to adjust taps, thereby increasing the depressed voltages along the downstream feeder. With capacitors and voltage regulators on the feeder, the overall system losses, including the line and transformer losses, have been substantially reduced from the initial value of 20.15% in the uncompensated case and 16.8% in the case where only capacitors were deployed on the feeder. As a final case for the snapshot simulation study, the IEEE-34 bus feeder is simulated with loads projected to grow at a 3% annual rate for the next ten years. It was found that given no additional compensation for the next ten years, the feeder experiences severely depressed voltages at several locations. It is also found that most of the regulators hit their maximum positions at the end of year 10. This suggests that the feeder will need more voltage correction in

Table 5.1: IEEE-34 Bus System Case Summary at Peak Load

Active Source Power	2.039 MW
Reactive Source Power	281 kVAr
Source Power Factor	0.9906
Line Losses	262.2 kW
Transformer Losses	10 kW
Total Losses	272.3 kW
Total Load Power	1767.6 kW
Percentage Losses	15.4%

Table 5.2: IEEE-34 Bus System Regulator Tap Positions with Peak Load

Name	Tap	Minimum	Maximum	Position
Regulator 1A	1.0875	0.9	1.1	14
Regulator 1B	1.025	0.9	1.1	4
Regulator 1C	1.03125	0.9	1.1	5
Regulator 2A	1.08125	0.9	1.1	13
Regulator 2B	1.08125	0.9	1.1	13
Regulator 2C	1.075	0.9	1.1	12

the future. Some of it could be overcome by adding new reactive support along the feeder length as the loads continue to grow, while some could be mitigated by updating the control mechanism and settings of the voltage regulators. Table 5.3 lists the regulator tap position at the year 10 with a projected 3% annual load growth on the feeder load. Table 5.4 lists additional feeder results at the end of 10 years with a sustained annual load buildup of 3%. It can be concluded that the feeder performance deteriorates under increased load and no additional compensation. The line losses at the end of year 10 equal 574.3 kW. Of course, it is because the feeder has to satisfy a higher load demand which means higher flows across the feeders and the laterals. The total load power met at the end of year 10 is 2341.1 kW, much higher than the base caseload with full compensation. The substation power factor also falls down to 0.955 because the source's reactive power requirement

has surged to 906 kVAr. The most troubling consequence of the load increase is the deteriorating voltage profile. The minimum voltage in the feeder drops down to 0.8 p.u, which is unacceptable. Hence with no new voltage compensation and an average load growth of 3% for the next ten years, the feeder does not meet the necessary requirements of a solved case.

Table 5.3: IEEE-34 Bus System Case Summary with 3% Load Growth for Ten years

Active Source Power	2.932 MW
Reactive Source Power	906 kVAr
Source Power Factor	0.955
Line Losses	574.3 kW
Transformer Losses	23.8 kW
Total Losses	598.1 kW
Total Load Power	2341.1 kW
Percentage Losses	25.5%

5.3.2 Quasi-static Time Series Simulation

To evaluate the impact on the voltage regulation equipment under the presence of a time-varying PV output and system loading, QSTS simulation tests are performed on the IEEE 34 node test feeder in OpenDSS. Since the feeder is located in Arizona, real residential load profile data is used with a time resolution of 1 hour. Residential load profiles of more than 1 hour are difficult to obtain and are not always made public by the utilities. Furthermore, it is assumed that all the loads on the feeder experience the same load profile. The time period selected for the load profile is one year, from January 1, 2012, to December 31, 2012. This takes into account the seasonal variations in the load, with the summer season experiencing heavy loads, most of which can be attributed to the air conditioning load. Figure 5.3 plots the feeder load for the first week of January 2012. Since the collected data is quite large, the load profile for one week is shown here to make visual sense.

The feeder load varies as the season's change, with peaks occurring mid-summer. The peak feeder load occurred on July 30, 2012, and is equal to 7800 kW. Throughout the year, the load varies from approximately 700 kW to more than 7500 kW.

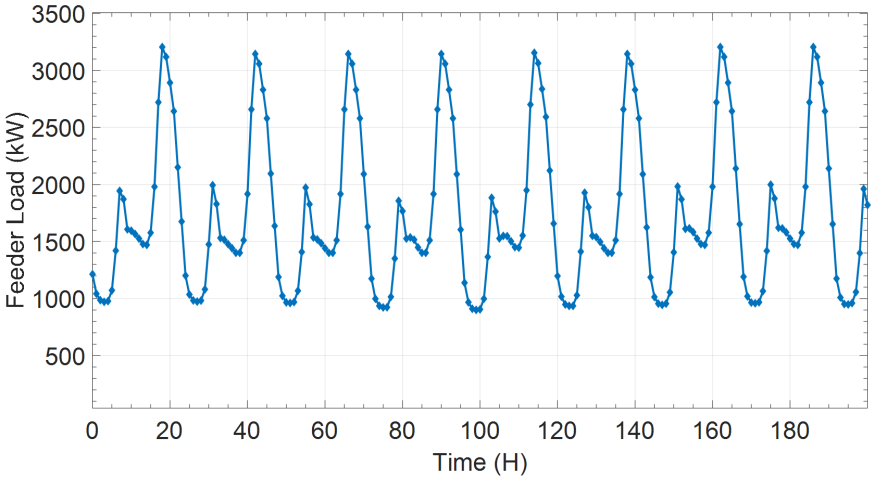

Figure 5.3: One Week Load Profile of IEEE-34 Bus Feeder

Figure 5.4 and Figure 5.5 show the active power flowing through Regulator 1 and the corresponding changes in the tap position. The plots indicate feeder conditions from January 1 to January 16, 2012. As the day begins on January 1, 2012, the PV output begins to ramp up which is reflected as a drop in active power flowing through the regulator is proportional to the degree of PV penetration on the feeder. Since the variability in power through the regulator are at a minimum, thus indicating a smooth ramp of the PV output, changes in tap positions are insignificant. However, as the week progresses, the stochastic nature of the PV generation introduces random fluctuations in the active power through the regulator. This leads to more tap operations to maintain the voltage in a defined band. It is also clear that at low PV penetration, the transformer tap does not vary much compared to high PV penetration. A small variation in the PV generation at high penetration can cause the transformer tap to rapidly change its position. It can also be observed that at high PV penetrations, reverse power flow can also take place across the regulator terminals.

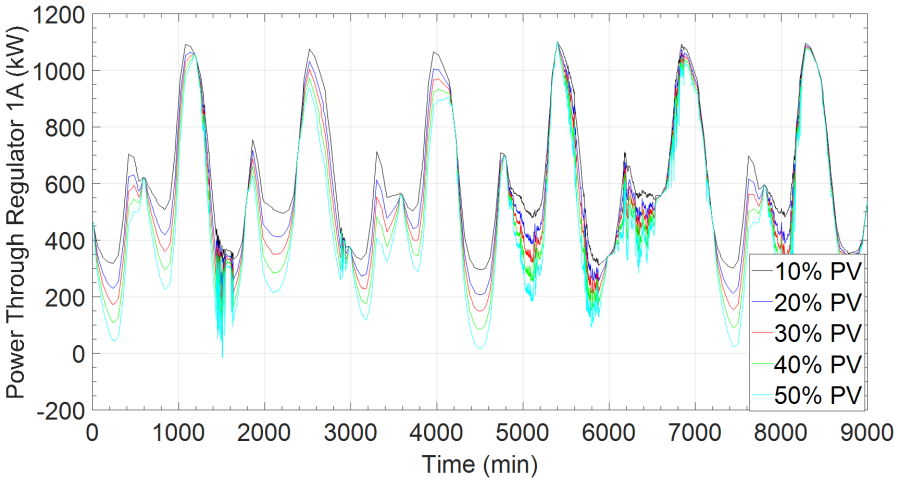

Figure 5.4: Active power through Voltage Regulator

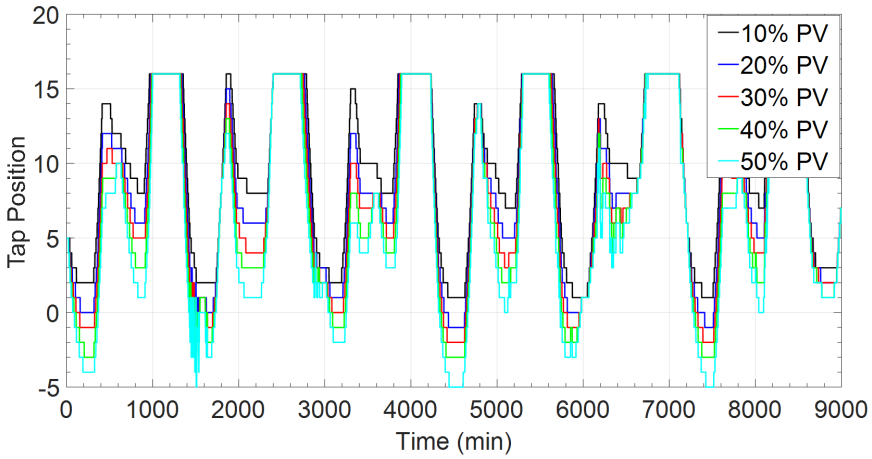

Figure 5.5: Changes in Tap Position

From Figure 7, a PV penetration of 50% reverse power flow happens roughly on January 3, 2012, for a brief period of time. On this day, the tilted irradiance hits the peak of approximately $1kW/m^2$, which explains the high PV output during this particular time of the year. The feeder at this point is relatively lightly loaded with a total system load of roughly 1400 kW, which is less than the base active power load of the system. Both the regulators exhibit similar behavior under high PV penetration, although the number of tap change operations differs. The location of regulator 2 makes it vulnerable to severely depressed voltages under peak load conditions. That explains the higher number of tap change operations in regulator two compared to regulator 1.

To quantify the annual tap changes with a load profile given in Figure 4 and the tilted irradiance profile shown in Figure 6, the IEEE 34 node feeder is subjected to time series simulations with varying degrees of PV penetration. The maximum increase in annual tap operations was found to occur in regulator 2A, with 33580 tap operations recorded under a PV penetration of 50%. In the absence of any PV generation, the number of annual tap operations in regulator 2A was found to be 9855. This represents an approximate increase of 3.5 times in the number of annual tap operations. Given a rough estimate of the average life of a load tap changer to be 500,000 operations, a 50% PV penetration on the feeder can shorten the life of the voltage regulator from 50 years to roughly about 15 years. Figure 5.6 shows the annual tap operations of the individual phases of the two regulators in the IEEE 34 node feeder as a function of PV penetration.

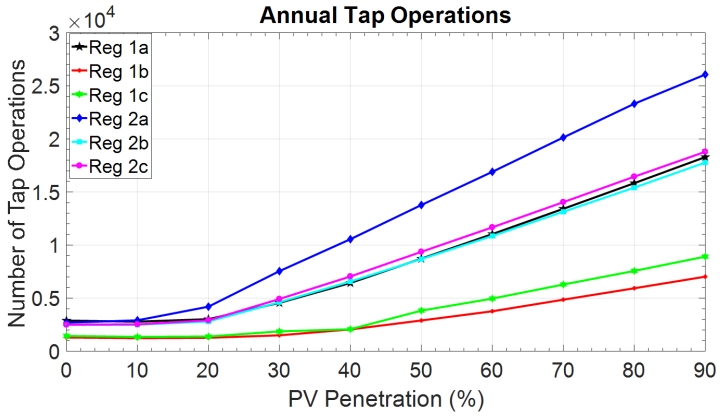

Figure 5.6: Cumulative Operations and Installed Solar Capacity

The voltage profile of Bus 890 is shown in Figure 5.7. This bus is located approximately 55 miles from the substation transformer and often exhibits depressed voltages, even with the compensation provided by the capacitors and the voltage regulators. The power to this bus is fed from the secondary of the in-line transformer, which steps down the line voltage from 24.9 kV to 4.16 kV. The rapid fluctuations in the voltage are a result of erratic power flow through the transformer. The percentage change in the voltage rise and drop increases with the increase in the PV penetration on the feeder. For a 50% PV penetration, the maximum rise and drop in the voltage exceed 1.5% during times when the intermittency of the tilt irradiance increases. On average, the changes in the voltage are in the order of 0.5%-1%. Such rapid change in the voltage can induce

flicker problems, especially in conditions of high PV penetration and increased solar irradiance variability. Given the upper threshold for the visibility of flicker to be 0.7% for the 1-minute interval, it can be concluded that flicker could be an issue under high penetration of photovoltaic power.

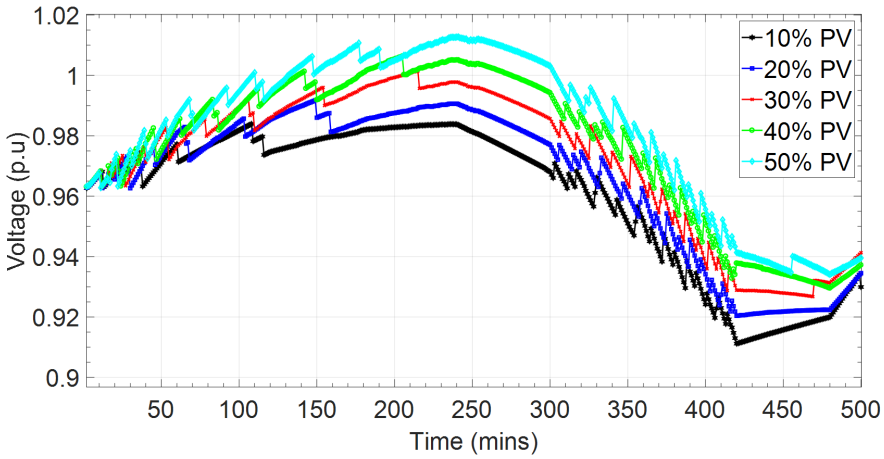

Figure 5.7: Bus Voltage Profile

5.4 Mathematical Modeling of Device Aging

This section presents a new statistical parametric model to predict the times-to-failure of broad classes of identical devices such as on-load tap changers, switched capacitors, breakers, etc. A two-parameter Weibull distribution with a scale parameter given by the inverse power law is employed to model the survivor functions and hazard rates of on-load tap changers. The resulting three-parameter distribution, called IPL-Weibull, is flexible enough to assume right, left, and symmetrical modal distribution. In this work, we propose an inferential method based on Bayes' rule to derive the point estimates of model parameters from the past right-censored failure data. It is possible to obtain such parameter estimates with high accuracy using the Monte Carlo integration technique.

For electric utilities, the problems of accurately identifying the service times of equipment and scheduling preventive maintenance are of critical importance as answers to these questions rep-

resent substantial savings to the utility [83],[84]. The service life of power equipment is affected by several factors that include but are not limited to insulation strength, thermal and non-thermal stress, moisture, etc. Many factors that impact the lifetime of a device are not practical for monitoring since failure databases with a complete list of all the failure-inducing factors rarely exist. This is further exacerbated by the impact of several failure-causing factors that are neither well documented nor the failure mechanism well understood. Given the limitations of the accurate logging of equipment failure, a probabilistic aging model that captures the most salient features of the aging process is the most practical for predicting the time to failure. Previous work in this area focused either on failure models derived from simple distributions or entirely ignored the impact of external stress on device aging. For example, in [45], an exponential failure model characterized by a single parameter is presented. The authors use Bayesian learning to estimate failure times based on historical data. The exponential distribution solves the problem of analytical tractability since it permits using a conjugate prior distribution for the parameter of interest. However, the exponential model lacks the flexibility that a Weibull distribution offers, making it less suitable to model device lifetimes. In [5], a distribution based on the Perks Hazard function is presented. The model, however, does not incorporate the impact of external stress on equipment aging. This chapter describes a failure model that combines the inverse power law and the Weibull distribution. This enables accurate modeling of non-thermal stress-related failures in devices like on-load tap changers, given a high penetration of intermittent non-scheduled generation. Since the failure database is often only partially available, we assume that the failure data contains only the following information: year of installation, the total number of assets, censoring number, and retirement history without replacement. An inferential technique based on Bayes' rule is developed to obtain the point estimates of model parameters. This enables us to predict the future performance of the assets that have survived based on the past failure history of similar devices, given similar failure mechanisms.

5.4.1 Cumulative Device Operations

We consider the problem of accelerated aging of devices like distribution transformers equipped with load tap-changers and switchable capacitor banks, given a high penetration of intermittent non-scheduled generation. The variable generation interferes with the regular operation of the tap-changers, causing them to operate much more frequently, usually outside the design limits. Due to the mechanical nature of the tap-changing devices, the increased frequency of operation leads to accelerated degradation of the device, which often results in premature failure of the equipment. Consider a substation transformer or a voltage regulator with $\lambda_(t_i)$ as the tap-ratio at the time instant,t_i. Let us consider a planning horizon, τ with K number of discrete time instants. If h is a fixed time step, the number of tap operations in a discrete interval, $[t_i, t_(i-h)]$ with $\triangle V$ as the step change in voltage per unit is

$$\delta_j = \frac{|\lambda(t_i) - \lambda(t_{i-h})|}{\triangle V}; i = 1, 2, ..., K \tag{5.3}$$

The cumulative tap operations over the planning period, τ is then given by

$$\zeta = \sum_{j=0}^{K-1} \delta_j \tag{5.4}$$

The interaction of the intermittent non-scheduled generation with the on-load tap changers results in heavy operational stresses imposed on the tap-changing devices. Since direct measurement of such time-dependent stress is difficult, if not entirely impossible, the change in the number of cumulative operations is a highly reliable indicator of such operational stress. We formulate a relationship between the mechanical stress imposed on tap-changing equipment and the resultant change in the number of cumulative operations. We use this relationship to develop an inverse power law-Weibull failure probability model of on-load tap-changers (OLTCs) and switchable capacitor banks. The parametric failure model can be used to forecast the remaining useful life and probability of failure of equipment given a high penetration of non-scheduled generation. It is im-

portant to note that in this work, we only consider the mechanical stresses imposed on the OLTCs and switched capacitors. If $p(t)$ is the probability density function of the time to failure, t of a device, then the probability of that device failing before time t is given by

$$F_T(t) = \Pr\{T \leq t\} = \int_0^t f(u)du \tag{5.5}$$

In the succeeding sections, we will examine the form of the function, f(u). In general, a parametric failure model takes the form [85]

$$\mathscr{F} = \{f(t;\overline{\theta}); \overline{\theta} \in \Theta; \Theta \subset \mathscr{R}^k\} \tag{5.6}$$

In 5.6, $\overline{\theta}$is the vector of model parameters that can take values in the parameter space, Θ The problem then reduces to one of estimating the parameters that characterize the failure model.

5.4.2 Characterization of Operational Stress

The development of the stress ratio factor is predicated on the understanding that an OLTC changes tap under the application of a force and hence stress on the contacts. Tap failure can either happen due to the asynchronous operation of the switches, usually caused by a broken axis, or due to the carbon formation and oxidation of contacts. While a broken axis may be a sudden event, the carbon formation on the contacts represents gradual aging, exacerbated by the intermittent non-scheduled generation [30]. The carbon formation and oxidation of contacts directly result from operational stresses imposed by the varying power flow conditions. It is possible to encode the information about the stresses and hence the gradual wear and tear of OLTCs in terms of the number of cumulative tap operations over the planning period. If m is the force (=stress) imposed on the contacts of the tap mechanism per tap operation, we can write for the total stress over the device lifetime, assuming ζ cumulative operations

$$M_0 = \zeta m \tag{5.7}$$

M_0 represents the total baseline mechanical stress over the device lifetime. The baseline stress indicates the device's wear and tear under normal conditions when accelerated aging of the device can be ignored. If ζ_{PV} is the cumulative operations of the tap-changer in the presence of solar generation, then

$$M_{PV} = \zeta_{PV} m \tag{5.8}$$

The stress ratio factor as a function of time t is

$$\gamma(t) = \frac{\zeta_{PV}(t)}{\zeta_{PV}(t)} \tag{5.9}$$

where, $\zeta_{PV}(t)$ is the cumulative tap operations till time instant, t given a high penetration of solar generation and $\zeta(t)$is the cumulative number of operations till the time, t in the absence of solar generation.

5.4.3 Proposed Failure Model and Failure Data

The reliability of a power distribution transformer essentially depends on the reliability of four components. These include 1) the reliability of paper winding insulation, 2) the reliability of the transformer tank, 3) the reliability of transformer bushes, and 4) the reliability of the load-tap changing mechanism. Among these, the reliability of load tap changers is a major concern and often makes up for a significant percentage of the transformer failure statistics. Several surveys [30],[39] focused on identifying the causes of distribution transformers' failures that OLTCs account for close to 50% of all the failed components. Given the growth in the installed capacity of renewable energy, in particular solar energy (solar farms and roof-top photovoltaic systems) over the past decade, and the fact that transformers equipped with OLTCs play a vital role in regulating feeder voltages and hence the flow of power in distribution systems, it becomes imperative to study the accelerated aging of such devices when exposed to the uncontrolled and fluctuating energy generation.

5.4.3.1 Failure Model

When accelerated aging is primarily driven by the excessive number of tap operations, whether in OLTCs or switched capacitor banks, the proposed failure model must be able to quantify and incorporate that information in the failure estimation process. Moreover, since the failure time of a device that is newly installed or of the one that has survived till the present moment is not completely known, any assessment of the remaining service life can only be a prediction. The error in such a prediction about the failure time of the device can be minimized, however, if historical failure data of a similar population is considered. Also, the information about the past failures can be used to fine-tune the parameters of the model or, in other words, to arrive at the best guess of the true parameter values. Needless to say, this will significantly improve the accuracy of failure predictions and result in a hazard rate consistent with the observed survival rate of such devices. In the light of these observations, it is natural to formulate the problem in a Bayesian setting, one that would allow for the incorporation of the failure data in ways that can lead to more accurate forecasting of the imminent and long-term failures. In this work, we restrict ourselves to nonthermal stress, also referred to as operational stress, resulting from frequent taps switching, usually outside the design parameters. We do not consider random failures due to external influences. Although the methods we present in this work have been applied to the lifetime estimation of mechanical assets, like distribution transformers equipped with a load tap-changer mechanism and switched capacitors, the theory's generality is preserved. As such, the parametric models developed here can be used to conduct failure estimation studies in any setting, where the gradual or accelerated aging due to mechanical stress is the main cause of equipment failure. In the case of distribution system transformers, aging-related failures contribute over 70% of the total failures, while random failures account for roughly 30% of the total failures [86]. Among the aging-related failures, the wearing out of On-load tap changers (OLTCs), has been singled out as a significant cause of transformer failure [87]. To address the problem, we consider the lifetime estimation of OLTCs and switched capacitor banks in a way that will help formulate replacement strategies.

In the literature, the lifetime of a device under nonthermal stress has been shown to follow the

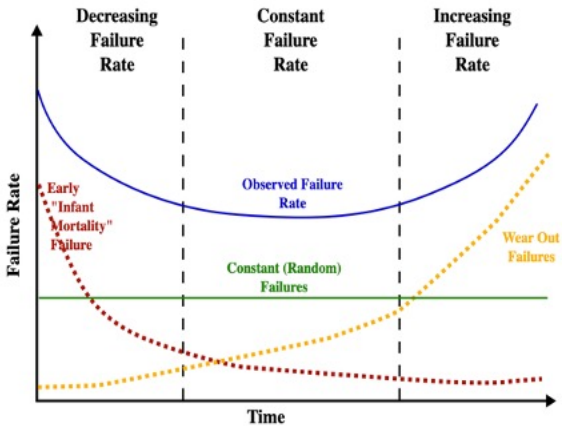

Figure 5.8: Typical Hazard Patterns

inverse-power law (IPL). If $L(M)$ is a lifetime of the device as a function of nonthermal stress M, the inverse-power law states that

$$L(M) = \left[\frac{M}{M_0}\right]^{-\eta} L_0 \tag{5.10}$$

L_0 refers to the device lifetime as a function of baseline stress M_0 below which accelerated device aging can be ignored. In (5.10), η is the stress endurance coefficient. To model equipment lifetimes, the Weibull hazard is consistent with the observed failures of broad classes of identical devices. The hazard function of the Weibull distribution is also very flexible and can model any of the hazard patterns shown in Fig. 5.8. The wear-out failures that result from the materials' decreasing mechanical strength are characterized by an increasing hazard, as shown on the far right of Fig. 5.8. The Weibull density function is also very flexible and can assume left, right, and symmetrical modal distribution.If we let T be the random variable that represents the unobserved failure time and supposing T follows a Weibull distribution, the density function of T includes three parameters as follow [88]

$$f_T(t|a,b,\beta) = \frac{\beta}{b}\left[\frac{t-a}{b}\right]^{\beta-1} \exp\left[-\left(\frac{t-a}{b}\right)^{\beta}\right] \tag{5.11}$$

In (5.11) a is the delay or minimum life of a device and hence the support of $f_T(t|a,b,c)$ is $t \geq a$. The domain of a is the interal $[0,\infty) \subset \Re$. The parameter b is the scale parameter in the domain $(0,\infty)$ and is measured in the same units as T. The third paramter β is the Weibull slope or

shape parameter. It is dimensionless with domain $(0,\infty)$. Our proposed failure model combines the inverse power law (5.10) with the density function (5.11) in a way that the three parameter Weibull, with $a = 0$ since devices begin to fail at age of $t = 0$, governs at each stress level and the scale parameter b varies inversely with the power of stress level. Combining (5.10) and (5.11) and setting $a = 0$ and $b = L(M)$ results in the density function of the form

$$f_T(t|\beta, L_0, \eta) = \begin{cases} \frac{\beta}{L_0^{\beta}} \left[\frac{M_0}{M}\right]^{-\beta\eta} t^{\beta-1} \exp\left[-\left(\frac{M_0}{M}\right)^{-\beta\eta} \frac{t^{\beta}}{L_0^{\beta}}\right] \\ 0; t < 0 \end{cases} \tag{5.12}$$

The information about the non thermal stress imposed on a tap-changing device is encoded in the cumulative operations the device registers over a specified planning horizon [89]. Letting $\gamma = \frac{M}{M_0}$, $\theta = L_0^{-\beta}$ and $\alpha = \beta\eta$ results in a re-parameterized model of the form

$$f_T(t|\beta, \theta, \alpha) = \begin{cases} \beta\theta t^{\beta-1}\gamma^{\alpha} \exp\left[-\gamma^{\alpha}\theta t^{\beta}\right]; t \geq 0. \\ 0; t < 0. \end{cases} \tag{5.13}$$

5.4.3.2 Description of Failure Data

The observed failure data with utilities usually includes two sets. For instance, the actual observed data of transformer lifetimes in PJM includes transformers installed and subsequently removed from service after failing and transformers that are in service until the present moment[5]. This type of data set is called censored or suspended data. In a censored data set, if a fixed number of failures, say r out of the total number of N assets, controls the test duration, the data is referred to as Type-II censored. This data type usually arises when the unit prices dominate, and failing is destructive. If we ignore replacement, the test duration, in this case, is the failure time of the r^{th} unit in a sample of N items

$$C = t_{r:N} \tag{5.14}$$

In (5.14), $t_{r:N}$ is the failure time of r^{th} unit. The sample size is N number of units. With Type-II censored data, the actual observed failure data set of N number of units is $\{(Y_1, \triangle_1), (Y_2, \triangle_2), ..., (Y_N, \triangle_N)\}$ where

$$Y_i = \min\{t_i, C\}; \triangle_i = \begin{cases} 1, t_i \leq C & \text{(Unit Failed)} \\ 0, t_i > C & \text{(Unit Censored)} \end{cases} \tag{5.15}$$

For a known number of failed/retired assets r, the Type-II failure data set with $N-r$ survived/censored units is

$$\mathscr{D} = \{(t_{1:N}, 1) < (t_{2:N}, 1) < ... < (t_{r:N}, 1) \leq (s_{r+1:N}, 0) \leq ... \leq (s_N, 0) | t_{r:N} = C\} \tag{5.16}$$

In (5.16) $s_{r+1:N}, ..., s_N$ are survival times of units that have not failed throughout the test duration. In the next section, we will develop an inferential scheme based on Bayes' rule to estimate the parameters of (5.13) with a failure data set of the form as given in (5.16).

5.4.4 Bayesian Inference Applied to Failure Estimation

The re-parameterized failure model (5.13) includes three unknown parameters, the shape β, scale θ and stress α. The parameter γ encodes information about the applied non-thermal stress. It is given by the ratio of the cumulative number of operations a device registers under stressed conditions, such as high penetration of distributed solar generation on the feeder and a baseline scenario (load injection only). It is possible to use the Bayesian method to evaluate the posterior probabilities of the unknown parameters in (5.13). Since the Bayesian method allows for incorporating subjective knowledge, it is a preferred estimation tool, especially when available data is insufficient. Besides, the Bayesian inference method allows for a more intuitive interpretation of parameter uncertainty in terms of probabilities that satisfy the likelihood principle. Once the posterior probabilities of unknown parameters are determined, point estimates such as conditional expectations or maximum a posteriori (MAP) estimates can also be obtained.

A fundamental difference between Bayesian inference and other statistical inference techniques such as maximum likelihood estimation (MLE) is that the unknown parameters are treated as

random variables in the Bayesian formulation. This is useful since we get a distribution of values of the unknown parameter instead of just one estimate. Hence various data interpretation ideas like credible intervals make more intuitive sense in a Bayesian paradigm. The principles of the Bayesian statistics allow us to infer the properties of the model parameters $\mathbf{X} = [\beta, \theta, \alpha]^T$ based on the potential failure data of the form given in (5.16). Assuming some prior distribution $p(\mathbf{X})$on $\mathbf{X}$ based on the subjective knowledge of an expert, the posterior of $\mathbf{X}$ is derived by using Bayes's theorem.

$$g(\mathbf{X}|\mathscr{D}) = \frac{f(\mathscr{D}|\mathbf{X})p(\mathbf{X})}{f(\mathscr{D})} = \frac{f(\mathscr{D}|\mathbf{X})p(\mathbf{X})}{\int_{\Omega} f(\mathscr{D}|\mathbf{X})p(\mathbf{X})d\mathbf{X}} \tag{5.17}$$

The marginal distribution $f(\mathscr{D})$ is independent of the parameter vector and acts as a normalization constant. Since the posterior probability depends on the assumed prior distribution, the Bayesian method offers the flexibility of interjecting expert knowledge into the estimation process. This is a desirable feature in reliability studies since the failure data of power equipment in general and OLTCs, in particular, are not readily available. In such situations, a strong prior based on experience can support the weak evidence that comes from insufficient data.

When the failure and the survival times are known, the failure model is represented using the likelihood function. For the proposed multi-parameter failure model given in (5.13), Appendix B describes the likelihood function considering Type-II censored data set. In (5.17) $g(\mathbf{X}|\mathscr{D})$ is the joint conditional posterior of the unknown parameters of (5.13). It is possible to obtain the marginal posterior of a parameter from the joint posterior by integrating out all parameters except the parameter of interest. The marginal posterior distribution of a parameter forms the basis for inferring a parameter's properties in the Bayesian paradigm.

5.4.4.1 Point Estimates of Model Parameters

One way to summarize the marginal posterior of an unknown parameter is to obtain the point estimates of the parameter. Point estimates represent the single best guess of the true parameter value, and based on the definition of the loss function considered, various point estimates can be derived. Some common ones include posterior mean, posterior mode, and posterior median [85].

In the Bayesian context, point estimates are also called Bayes estimators. Consider a failure model, $f(t|\mathbf{X})$ parameterized by $\mathbf{X}$ in the parameter space, Ω. In deriving a point estimate of the parameter, $\mathbf{X}$ referred to as $\hat{\mathbf{X}}$ the discrepancy between $\mathbf{X}$ and $\hat{\mathbf{X}}$is measured by the loss function, $L(\mathbf{X},\hat{\mathbf{X}})$. A loss function is a mapping from the parameter space to real space. To measure the risk associated with the point estimator, we consider a quadratic loss function

$$L(\mathbf{X},\hat{\mathbf{X}}) = (\mathbf{X}-\hat{\mathbf{X}})^2 \tag{5.18}$$

The posterior risk of the estimator, $\hat{\mathbf{X}}$is given by

$$R(\hat{\mathbf{X}}|\mathscr{D}) = \int L(\mathbf{X},\hat{\mathbf{X}}) f(\mathbf{X}|\mathscr{D}) d\mathbf{X} \tag{5.19}$$

Where $f(\mathbf{X}|\mathscr{D})$ is the marginal conditional posterior distribution of the parameter, $\mathbf{X}$ given the data in $\mathscr{D}$. With a squared loss function, the point estimate or the Bayes estimator, $\hat{\mathbf{X}}$ of the parameter, $\mathbf{X}$ is the expectation of $\mathbf{X}$.

$$\hat{\mathbf{X}}(\mathscr{D}) = \int \mathbf{X} f(\mathbf{X}|\mathscr{D}) d\mathbf{X} = \mathbb{E}[\mathbf{X}|\mathscr{D}] \tag{5.20}$$

5.4.4.2 Point Estimates with Uninformative Priors

If the prior information about the shape paramter β and the scale paramter θ is not available, Jefferey's vague prior is a good choice to model the prior lack of knowledge of the model parameters. Jeffrey's prior is an uninformative prior and is invariant to parameter transformation.

$$p(\beta) = 1/\beta; p(\theta) = 1/\theta \tag{5.21}$$

For the stress parameter, we assume a uniform prior.

$$p(\alpha|A,B) = \frac{1}{B-A}; A \leq \alpha \leq B; A,B > 0 \tag{5.22}$$

With this choice of priors, we can write for the joint posterior distribution of the model parameters

$$g(\beta,\theta,\alpha|t_1,...,t_r,s_{r+1},...,s_N,C,N,r) \propto \frac{1}{B-A}\beta^{r-1}\theta^{r-1}\prod_{i=1}^{r}t_{i:N}^{\beta-1}\gamma^{r\alpha}\exp[-\gamma^{\alpha}\theta P] \tag{5.23}$$

More information on how to obtain the likelihood function of the failure model and the joint distribution of the model parameters is given in Appendix A. To get the point estimates of the parameters, we integrate out all the parameters in (5.23) except the parameter of interest. With noninformative priors and the censored failure data, the Bayes Estimators of the shape parameter in is [**?**]

$$E(\beta|t_1,...,t_r,s_{r+1},...,s_N,N,C,r) = \hat{\beta} = K_{\beta,r}\int_0^{\infty}\beta^{r}\prod_{i}^{r}t_{i:N}^{\beta-1}P^{(-r)}d\beta \tag{5.24}$$

$$K_{\beta,r}^{-1} = \frac{1}{B-A}\int_0^{\infty}\int_A^B\beta^{r-1}\prod_{i}^{r}t_{i:N}^{\beta-1}P^{(-r)}d\beta \tag{5.25}$$

The conditional expectation of the scale parameter,θ with noninformative prior and Type-II censored data is

$$E(\theta|\mathscr{D},s_{r+1},...,s_N,N,C,r) = \hat{\theta} = \frac{1}{B-A}\frac{\gamma^A-\gamma^B}{\gamma^{A+B}\ln\gamma}\frac{I_{\theta,1}}{I_{\theta,2}} \tag{5.26}$$

Where

$$I_{\theta,1} = \int_0^{\infty}\beta^{r-1}\prod_{i=1}^{r}t_{i:N}^{\beta-1}P^{-(r+1)}d\beta \tag{5.27}$$

$$I_{\theta,2} = \int_0^{\infty}\beta^{r-1}\prod_{i=1}^{r}t_{i:N}^{\beta-1}P^{-(r)}d\beta \tag{5.28}$$

The Bayes Estimator for the stress coefficient α in (4.6) with noninformative priors and Type-II censored data is

$$E(\alpha|t_1,...,t_r,s_{r+1},...,s_N,N,C,r) = \hat{\alpha} = \frac{K_{\alpha,r}}{B-A}\int_0^{\infty}\int_0^{\infty}\beta^{r-1}\prod_{i}^{r}t_{i:N}^{\beta-1}P^{(-r)}d\alpha d\beta \tag{5.29}$$

$$K_{\alpha,r}^{-1} = \int_0^{\infty}\beta^{r-1}\prod_{i}^{r}t_{i:N}^{\beta-1}P^{(-r)}d\beta \tag{5.30}$$

5.4.4.3 Point Estimates with Informative Priors

For the proposed IPL-Weibull model, there does not exist any family of continuous conjugate prior distributions for the random shape, scale, and stress parameters that are closed under-sampling. In the case of the random scale parameter and a known shape parameter, the gamma prior for the scale parameter is closed under-sampling. We assume a gamma prior for the scale and shape parameter [90] and a uniform prior for the stress coefficient in the closed interval $[A,B]$.

$$p(\beta) \sim \text{Gamma}(\lambda_1,\delta_1); p(\theta) \sim \text{Gamma}(\lambda_2,\delta_2) \tag{5.31}$$

With hyperparameters $\lambda_1 > 0, \delta_1 > 0, \lambda_2 > 0$ and $\delta_2 > 0$. The joint conditional posterior of β, θ, and α with gamma prior on β and θ and a uniform prior on α is

$$g(\beta,\theta,\alpha|t_1,...,t_r,s_{r+1},...,s_N,C,N,r) \propto \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_s)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha} \prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^{\alpha}\theta P+\lambda_1\beta+\lambda_2\theta)] \tag{5.32}$$

The joint conditional posterior (5.32) can be used to obtain the marginal posterior and the conditional expectations of the parameters. The conditional expectation of β with priors defined in (5.31) and censored data defined in (2.6) is

$$E(\beta|\mathscr{D},s^j,N,C,r) = \hat{\beta} = \int_{\beta} \beta f(\beta|\mathscr{D},s^j,N,C,r)d\beta = \frac{\int_{\beta}\beta^{r+\delta_1}\exp(-\beta\lambda_1)\prod_{i=1}^{r} t_{i:N}^{\beta-1}P^{-(r+\delta_2+1)}d\beta}{\int_{\beta}\beta^{r+\delta_1-1}\exp(-\beta\lambda_1)\prod_{i=1}^{r} t_{i:N}^{\beta-1}P^{-(r+\delta_2+1)}d\beta} \tag{5.33}$$

The conditional expectation of θ with informative priors and censored data is

$$E(\theta|\mathscr{D},s^j,N,C,r) = \hat{\theta} = \int_{\theta} \theta f(\beta|\mathscr{D},s^j,N,C,r)d\theta = K_{\theta,r}\frac{PW_1-(r+\delta_2)\lambda_2 W_2}{PW_3-(r+\delta_2)\lambda_2 W_4} \tag{5.34}$$

Where

$$K_{\theta,r} = \frac{\int_{\beta}\beta^{r+\delta_1-1}\exp(-\beta\lambda_1)\prod_{i=1}^{r} t_{i:N}^{\beta-1}P^{-(r+\delta_2+2)}d\beta}{\int_{\beta}\beta^{r+\delta_1-1}\exp(-\beta\lambda_1)\prod_{i=1}^{r} t_{i:N}^{\beta-1}P^{-(r+\delta_2+1)}d\beta} \tag{5.35}$$

$$W_1 = \frac{\gamma^{-A(\delta_2+1)} - \gamma^{-B(\delta_2+1)}}{(\delta_2+1)\ln\gamma}; W_2 = W_4 \tag{5.36}$$

$$W_3 = \frac{\gamma^{B\delta_2} - \gamma^{A\delta_2}}{\delta_2\gamma^{\delta 2(A+B)}\ln\gamma}; W_4 = \frac{\gamma^{-A(\delta_2+2)} - \gamma^{-B(\delta_2+2)}}{(\delta_2+2)\ln\gamma} \tag{5.37}$$

Finally, the conditional expectation of the stress coefficient parameter with informative priors on β and θ and Type-II right censored data is

$$E(\alpha|\mathscr{D},s^j,N,C,r) = \hat{\alpha} = \int_\alpha \beta f(\alpha|\mathscr{D},s^j,N,C,r)d\alpha = \frac{\int_\alpha \alpha\gamma^{-\alpha(\delta_2+1)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha}{\int_\alpha \gamma^{-\alpha(\delta_2+1)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha} \tag{5.38}$$

In (5.33),(5.34) and (5.38), $s^j = [s_{r+1},...,s_N]$ is the vector of survival times of $N-r$ censored units. The mathematical derivation of the point estimates of the model parameters with informative priors is given in Appendix A. It is important to note that the point estimates given by 5.24, 5.26, 5.29,5.33, 5.34 and 5.38 do not have closed form solutions and must be solved using a numerical technique. The Bayesian methodology for failure estimation is presented in Fig. 5.9. It is important

Algorithm 2: Point Estimates of Model Parameters

Input: Joint Posterior of parameters $g(\mathbf{X}|\mathscr{D})$
for all i=1,..,d **do**
  Get marginal conditional posterior of parameter of interest from joint posterior
  $f(x_i|\mathscr{D}) = \int_{\mathbf{X}_{-x_i}} g(\mathbf{X}|\mathscr{D})d\mathbf{X}_{-x_i}$
  $\mathbf{X}_{-x_i} = [x_1,x_2,...,x_{i-1},x_{i+1},...,x_d]^T$
  Evaluate conditional expectation of parameter from marginal conditional posterior
  $E(x_i|\mathscr{D}) = \int_{x_i} f(x_i|\mathscr{D})dx_i$
end for
Output: Point Estimates of Parameters $E(x_i|\mathscr{D})$

to mention that in a Bayesian paradigm, information present in the failure data is combined with subjective knowledge to better forecast the in-service population. Subjective knowledge can either be informative, as is in the case of shape and scale parameters with gamma priors, or can even model ignorance, as in for stress coefficient where we do not have a clear prior knowledge. In

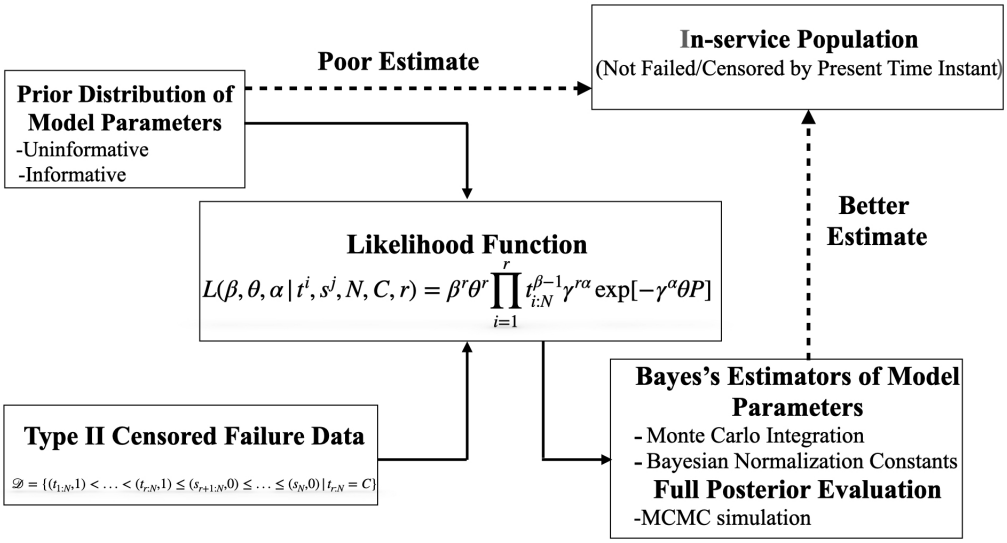

Figure 5.9: Failure Estimation via Bayesian Inference

many ways, Bayesian inference is a balancing act that works best when the choice of prior does not undermine the evidence in the data, nor does the evidence overwhelm the prior belief.

5.4.5 Simulation Technique Applied to Failure Estimation

The point estimates of the model parameters represent the single best guess of the actual parameter values. However, the point estimates of the parameters cannot account for the uncertainty in the estimation process. The uncertainty can, however, be quantified by obtaining the full posterior of the model parameters. Since the proposed failure model does not allow for any closed-form solutions of the posterior distribution of parameters, simulation procedures such as Metropolis-Hastings (M-H) algorithm draws samples from the target posterior. M-H algorithm is one of many Markov Chain Monte Carlo (MCMC) methods that enables sampling from a target posterior that cannot be obtained analytically. The MCMC methods circumvent the problem of analytical tractability by simulating a Markov chain on a continuous state space under conditions that guarantee convergence such that the stationary distribution of the simulated chain is the target posterior distribution.

5.4.5.1 *Evaluation of Full Posterior*

Our implementation of the M-H algorithm simulates a d-dimensional discrete-time Markov process $\{\mathbf{X}_n\}_{n\in\mathrm{N}}$ on a continuous state space such that at each time step, the proposed state vector $\mathbf{X}_{n+1}$ is generated conditional on the current state vector. The proposed state vector is accepted or rejected according to an appropriately defined acceptance probability. The proposed state vector is sampled from a suitably defined proposal distribution $q(\mathbf{X}_{n+1}|\mathbf{X}_n)$, and the algorithm is initiated by choosing a starting state vector $\mathbf{X}_0$ in the domain of the posterior of the parameter. An important assumption is that the target posterior is strictly positive in the entire state space. Our implementation of the M-H algorithm is described in Appendix C. To initiate the sampling procedure we let the point estimates of the model parameters be the starting state vector $\mathbf{X}_0 = [\hat{\beta}, \hat{\theta}, \hat{\alpha}]^T$ and the prior distributions of the parameters as the proposal distribution $q(\mathbf{X}_{n+1}|\mathbf{X}_n) = p(\mathbf{X}_{n+1})$. This leads to a new definition of the acceptance probability and (B.4) in Appendix C becomes

$$a(\mathbf{X}_n, \mathbf{X}^*) = \min\left\{1, \frac{(\beta^*)^r(\theta^*)^r\gamma^{\alpha^*}\prod_{i=1}^r t_{i:N}^{\beta^*-1}\exp\left[-\gamma^{\alpha^*}\theta^* P^*\right]}{\beta_n^r\theta_n^r\gamma^{\alpha_n}\prod_{i=1}^r t_{i:N}^{\beta_n-1}\exp\left[-\gamma^{\alpha_n}\theta_n P_n\right]}\right\} \tag{5.39}$$

The definition of the acceptance probability ensures that proposed state values of a parameter with a higher probability of being observed are accepted while the others are rejected. Although the optimal choice of a proposal distribution is still an open problem in MCMC methods, in a Bayesian setting, setting the proposal distribution equal to the prior distribution is a reasonable strategy since it makes the resulting numerical calculations tractable.

5.4.5.2 *Component wise Sampling*

The proposed failure model is a multi-parameter model, and since parameters of interest are treated as random variables, we are essentially performing the multidimensional sampling. This is achieved by utilizing the component-wise sampling procedure, which involves looping the algorithm over the d dimensions of the parameter vector $\mathbf{X}$. Each dimension of the parameter vector $\mathbf{X}$ is sampled independently of the others. Component-wise sampling requires that the proposal distribution be univariate, and the sampling should occur in only one dimension: the dimension in

which the algorithm works.

Algorithm 3: Component wise Metropolis Hastings

Input:Time Index n, Starting State Vector $\mathbf{X}_0$, Proposal Density $q(\mathbf{I}^*|I^{n-1})$, Chain Length K
Set $n = 0$
Set initial state vector $\mathbf{X}_0 = [\hat{\beta}, \hat{\theta}, \hat{\alpha}]^T$
while $n \leq K$ **do**
  Set $n = n + 1$
  for all i=1,..,d **do**
    Generate a proposal state $x_i^{(*)} = q(x_i|x_i^{(n-1)})$
    Calculate acceptance probability, $a(\mathbf{X}_n, \mathbf{X}^*)$
    Draw $u \sim \text{Uniform}(0,1)$
    if $u \leq a$ **then**
      $x_i^{(n)} = x_i^{(*)}$
    else
      $x_i^{(n)} = x_i^{(n-1)}$
    end if
  end for
end while
Output: Posterior Distribution of Model Parameters

5.4.5.3 Burn-in and Autocorrelation

Since Markov chains are stochastic processes, the starting state's specification impacts the randomness of the process. Burn-in refers to the loss of memory of the initial conditions. In our implementation of the M-H algorithm, burn-in is achieved by discarding the first five hundred iterations of the algorithm. There is no definite answer to the problem of the optimal number of iterations that would signal burn-in occurrence. However, our results indicate that after the first five hundred iterations, the impact of specifying a starting state vector is mitigated considerably. Moreover, since the Markov process that we simulate is guaranteed to be irreducible, aperiodic, and positive recurrent, the basic limit theorem holds. This means that if the process is run long enough, the starting point is irrelevant.

Autocorrelation is cause for concern in MCMC methods and even more so for the M-H algo-

rithm that generates highly correlated draws. For a Markov process $\{\mathbf{X}_n : n \geq 0\}$ the autocorrelation at times n_1 and n_2 is defined as

$$\mathscr{V}(n_1, n_2) = \text{Cov}[\mathbf{X}_{n_1}, \mathbf{X}_{n_2}] = \text{E}[(\mathbf{X}_{n_1} - \bar{\mu}_{n_1}) - (\mathbf{X}_{n_2} - \bar{\mu}_{n_2})] \tag{5.40}$$

For a well-mixing chain, the autocorrelation must decrease and become negligible as $n_2 - n_1$ increases. The progressive decrease in autocorrelation with time indicates convergence of the M-H algorithm. A poorly-mixed chain with high autocorrelation or in which the autocorrelation does not reach zero after a substantial amount of lag may fail to explore the new areas of the posterior.

5.5 Failure Prediction of Mechanically-operated Devices

The proposed Bayesian method is applied to a model of the IEEE-34 Bus test system to estimate the lifetimes of voltage regulators with different penetration levels of solar generation. The test system has a nominal voltage of 24.9 kV. The feeder is characterized by long lines and light loads and requires two voltage regulators to keep the voltage within ANSI limits. Both the voltage regulators use line drop compensation (LDC) to measure the drop in voltage between the regulator and load center. To study the impact of solar generation on tap-changers, a full three-phase model of the circuit consisting of all circuit lines (single-phase and three-phase lines), regulators, customer loads, capacitor banks, substation, and in-line transformers, control elements of capacitors, and voltage regulators are developed in OpenDSS. For solar generation, a proportionally distributed configuration is chosen with the rated power of the PV systems proportional to the loads. Each PV system is interfaced with an inverter with a rating 1% higher than the PV panel. An increase in the solar capacity is expected to negatively impact the device lifetime due to a corresponding increase in the device wear and tear.

5.5.1 Model Validation

The analysis in the previous section assumes that the censored failure data is a sample from the Weibull distribution in which the scale parameter shares an inverse relationship with the applied stress. To check the validity of the assumption, we compute the Kolmogorov-Smirnov (KS) dis-

tance between the empirical distribution function and the fitted distribution function. KS test is a nonparametric test used to test the hypothesis that the sample (failure data) comes from a particular distribution. In our case, we hypothesize that the times to failure come from an IPL-Weibull distribution. The sample failure data is given in Table 5.4. The failure data assumes 40 assets installed ten years ago with a retirement history, as shown in Table 5.4.

Table 5.4: Censored Transformer Failure Data [5]

Year	Age	Retirements	Survivors	Survivor Rate
2009	0	0	40	100%
2010	1	1	39	98 %
2011	2	0	39	98 %
2012	3	0	39	98 %
2013	4	0	39	98 %
2014	5	1	38	95 %
2015	6	2	36	90 %
2016	7	1	35	88 %
2017	8	5	30	75 %
2018	9	4	26	65%
2019	10	6	20	50%

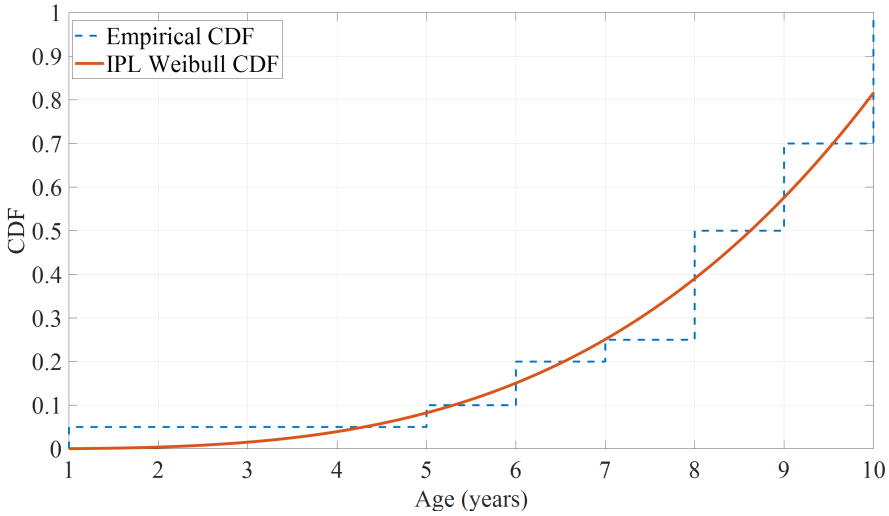

Figure 5.10: Empirical CDF and Model Predicted CDF

Figure 5.11 shows the result of the KS test. It is clear that the fitted IPL-Weibull provides an excellent fit to the failure data. Note that the fitted IPL-Weibull CDF uses point estimates of the parameters given by 5.24,5.26 and 5.29. Figure 5.11 confirms that the KS test accepts the hypothesis that the failure data of Table 1 is a sample from the Weibull distribution with a scale parameter given by the inverse power law.

To observe the impact of non-thermal stress or equivalently solar generation on the OLTC lifetimes, we designed annualized experiments over a 10-year planning horizon. The results of the tests are categorized into three scenarios which are 1) Load growth over the next ten years with no solar generation, 2) Load growth with low penetration of solar generation (30%), and 3) Load growth with high penetration of solar generation (90%). Note that the photovoltaic (PV) penetration level is defined as the ratio of the aggregate peak capacity of all PV systems and the total peak active load of the feeder. We consider a load growth of 3% for the first four years, followed by 5% and a 7% growth distributed equally for the remaining six years. This is within the conservative estimate of 3%-7% growth in feeder loads at the distribution level. We consider two different penetration levels for solar generation over the ten-year planning horizon. The accumulated stress on the OLTCs is proportional to the cumulative number of tap operations over the planning horizon.

5.5.2 Lifetime estimation results with Uninformative Priors

Since the failure data is censored on the right, we have two sets of observed lifetimes. One set contains the failed/retirement times $\{t_i \leq C; i = 1, 2, ..., r\}$ and the second set includes the survival times of the OLTCs not failed or retired by the present time instant. The Bayesian method enables us to make an inference on the second set of OLTCs based on the data in the first set since the failure/retirement time of the second set is unknown at the present time instant. The inference is valid because the OLTCs in both sets experience similar failure mechanisms. Figures 5.14 and 5.15 show the failure density function, hazard rate, and the survival function of the two voltage regulators of IEEE 34 bus system under different scenarios.

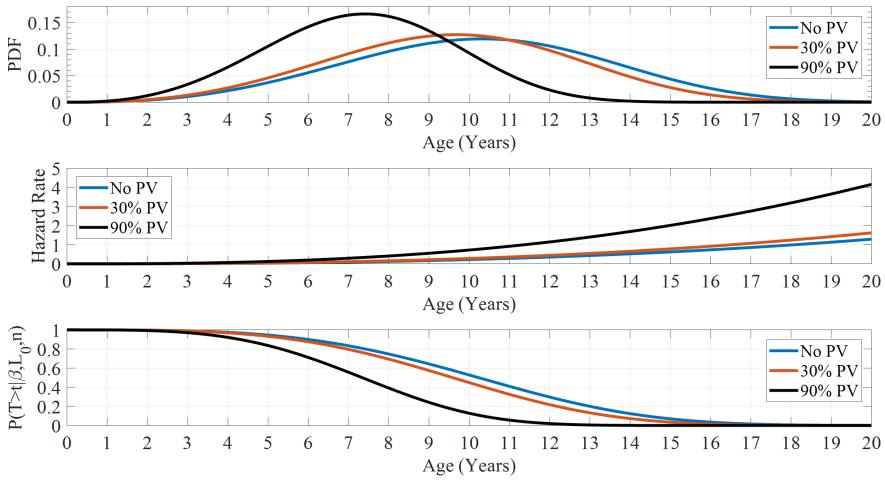

Figure 5.11: PDF, Hazard Rate and Survivor Function of Voltage Regulator-1

The regulators are assumed to have survived until the present instant and are not censored. From Figures 5.14 and 5.15 it is clear that while a low PV penetration may not significantly reduce the lifetimes of voltage regulators, sustained high PV penetration has a significant impact on the device's lifetime. This is further evinced by the hazard rates of both the regulators, which show a marked increase towards the end of the device lifetime with high PV penetration.

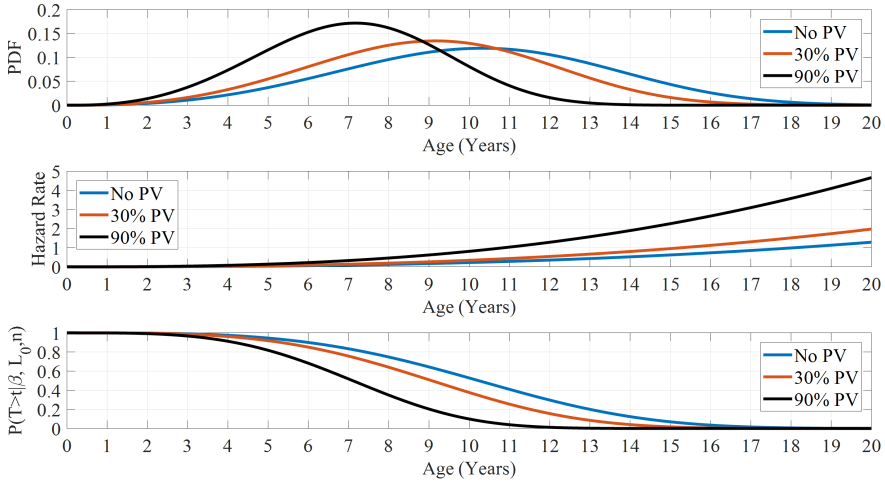

Figure 5.12: PDF, Hazard Rate and Survivor Function of Voltage Regulator-2

It is interesting to note how the failure density changes after acquiring new data. Figure 5 shows the impact of censoring number on the failure density function. The transition of failure

density from the poor initial estimate (blue curve) to the final estimate (green curve) is remarkably fast. The mean time to failure in the initial estimate with $r = 1$ is 114200 years, considering no PV penetration. With $r = 20$, the mean time to failure is 10.255 years. This suggests the method is very robust. Since the inference on the surviving assets is drawn from the set of failed/retired assets, the proposed method will perform better with a large amount of censored failure data. Table 5.5 lists the point estimates of the model parameters as a function of the censoring number. As more failure data are acquired, the failure prediction improves. The most likely $(\hat{\beta}, \hat{\theta}, \hat{\alpha}$ with $r = 20$ is $3.52, 9.16, 0.1108$ for voltage regulator-1 in the No PV scenario. Tables 5.6 and 5.7 list some statistical properties of the failure density function shown in Figure 5.11 and Figure 5.14. The impact of heavy PV penetration on device lifetime is quite apparent. This can be realized by observing the mean time to failure (MTTF) of the two voltage regulators under low PV and high PV penetration and comparing that with the No PV scenario. With a high PV penetration, the MTTF of voltage regulator 1 is 7.33 years; for voltage regulator 2, the MTTF with high PV penetration is 7.09 years.

Table 5.5: Impact of censoring information on Parameter Estimates

Parameter	$r = 1$	$r = 5$	$r = 10$	$r = 15$	$r = 20$
$\hat{\beta}$	0.43	1.11	1.83	2.65	3.5217
$\hat{\theta}$	7655	30.7	13.09	10.06	9.1617
$\hat{\eta}$	0.88	0.34	0.21	0.14	0.1108

Table 5.6: Statistical Properties of Figure 5.11

Statistical Property	No PV	30%PV	90% PV
Mean Time to Failure	10.25	9.58	7.33
Median	10.238	9.59	7.34
Mode	10.33	9.68	7.41

Table 5.7: Statistical Properties of Figure 5.14

Statistical Property	No PV	30%PV	90% PV
Mean Time to Failure	10.25	9.06	7.09
Median	10.238	9.07	7.1
Mode	10.33	9.15	7.17

5.5.3 Lifetime estimation results with Informative Priors

To predict the unobserved failure times, we compute the Bayes Estimators of the unknown parameters of the failure model using informative priors. A comparison between the performance of the estimators is made by computing the Mean Square Error (MSE) over 1000 Monte Carlo replications. The OLTCs of IEEE-34 feeder are assumed to be newly installed or have survived till the present moment, and thus their survival times are the same as the censoring time . We compute the stress ratio factor γ to complete the model identification by observing the cumulative tap operations registered by the OLTCs with and without PV generation over a planning horizon representing typical OLTC lifetimes. Since using informative priors involves assigning values to the hyperparameters, we consider two sets of values for the gamma hyperparameters. The high PV scenario-I corresponds to hyperparameters $\lambda_1 = \lambda_2 = \delta_1 = \delta_2 = 1$ and the high PV scenario-II refers to hyperparameter values $\lambda_1 = 3, \lambda_2 = \delta_1 = \delta_2 = 1$. The uniform interval considered for the stress parameter α is [0,1]. The Bayes estimators of model parameters and stress ratio factor are substituted in (5.13) to compute the statistical properties of the in-service OLTC units. An example is shown in Fig. 5.13, where the failure density, hazard rates, and survival probability of one of the in-service OLTCs in the IEEE-34 Bus test system are plotted. It is clear from Fig. 5.13 that a heavy PV penetration significantly impacts the OLTC lifetime. Compared to the baseline case, the reduction in a lifetime with a 90% PV penetration is anywhere from $6-10$ years while the probability of surviving beyond 20 years drops from the baseline value of 61.31% to just 3%

in High PV scenario-I and 27.25% in High PV scenario-II. This underscores the importance of expert knowledge in choosing the prior distributions. Since the gamma priors are a function of hyperparameters, informative priors offer more flexibility in incorporating expert knowledge with the inference process. This can lead to a better forecast of the in-service OLTC units' survivability given the past failure history of the units that failed/retired by the present time. Table 5.8 lists the statistical properties of Fig. 5.13.

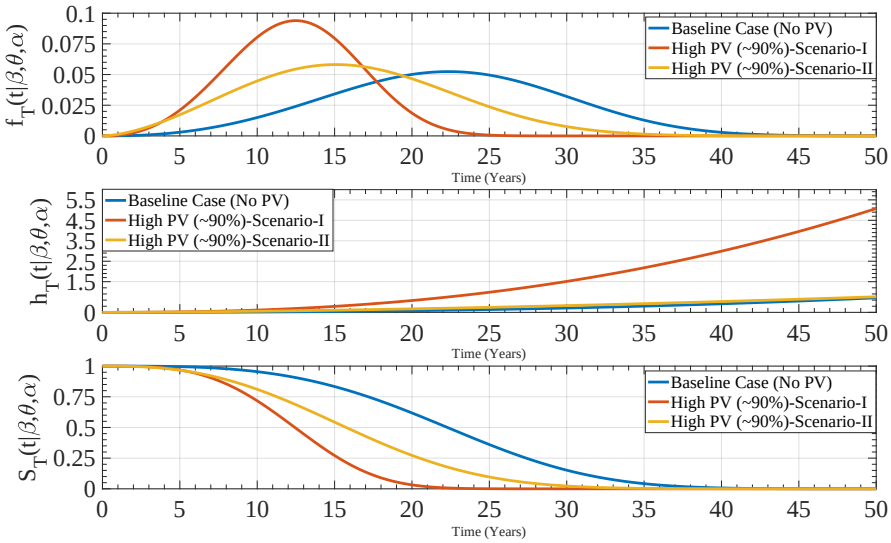

Figure 5.13: Failure Density, Hazard Rate and Survival Probability with N=30 and r=20

Table 5.8: Estimates of In-service OLTC Population

Statistical Property	Baseline Scenario	High PV (90%)	
		Scenario-I	Scenario-II
MTTF (Years)	22.31	12.42	16.08
Variance	24.85	13.86	18.105
Median (Years)	22.29	12.43	15.74
Mode (Years)	22.38	12.48	15.09
$P[T > 10\|\hat{\beta},\hat{\theta},\hat{\alpha}]$	95.45%	70.8%	81.06%
$P[T > 20\|\hat{\beta},\hat{\theta},\hat{\alpha}]$	61.31%	3%	27.25%

we also consider the impact of censoring on the inference process and the MSE of the estimators. From the experiments, we find that for fixed N and γ as the censoring number r increases, the failure probability curves result from using two different sets of hyperparameter values; thus reflecting two different prior beliefs, tend to converge. Note that an increase in the censoring number means an increase in the amount of information available to the failure model. This means that as more failure data are acquired, the posterior distributions of the parameters are more influenced by the data than the prior beliefs. This explains the convergence of the failure probability curves given different prior beliefs as we approach a complete dataset with $N = r$. This is a common theme in Bayesian analysis where the sample size of the data mitigates the impact of the prior belief. It is interesting to observe the changes in the failure density as r changes. Fig. 5.14 shows the impact of r with fixed N on the failure density.

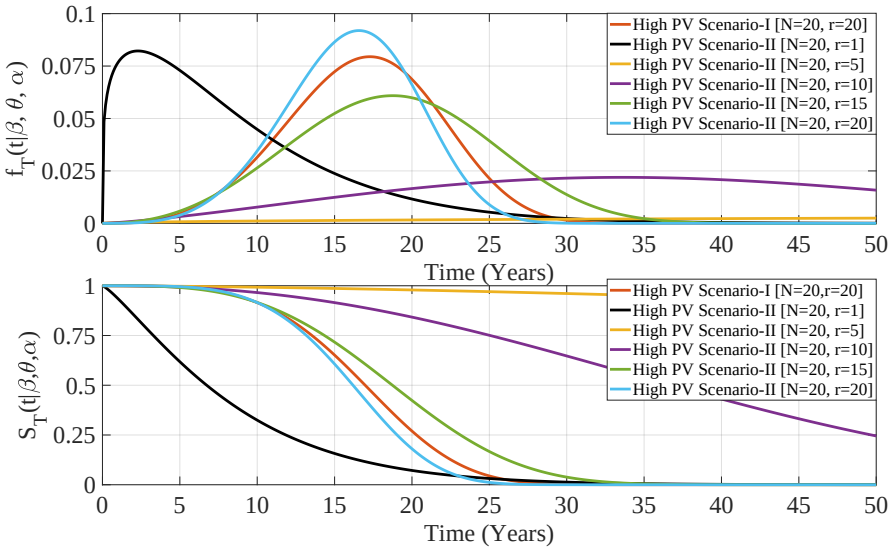

Figure 5.14: Impact of Censoring Information on Failure Prediction

The transition of failure density from a poor initial estimate (black curve) to the final estimate (green curve) is remarkably fast. The Bayes estimators of unknown parameters of the failure model and mean time to failures (MTTF) as a function of the censoring number r and fixed $N = 30$ are given in Table 5.9, 5.10.

Table 5.9: Impact of Censoring Number: High PV Scenario-I

Estimator	r=1	r=5	r=10	r=15	r=20
$\hat{\beta}$	0.61	1.11	1.78	2.54	3.36
$\hat{\theta}$	0.0212	0.0225	0.0241	0.025	0.0236
$\hat{b}$	1521.58	141.735	47.19	25.07	13.86
$\hat{\alpha}$	0.533	0.59	0.691	0.927	1.90
$\hat{\eta}$	0.9	0.532	0.38	0.364	0.560
MTTF (Years)	2221.4	135.92	41.99	22.18	12.44

Table 5.10: Impact of Censoring Number: High PV Scenario-II

Estimator	r=1	r=5	r=10	r=15	r=20
$\hat{\beta}$	0.38	0.825	1.37	1.97	2.6
$\hat{\theta}$	0.0211	0.022	0.0235	0.0242	0.0231
$\hat{b}$	32187.56	372.09	77.215	35.1994	18.1015
$\hat{\alpha}$	0.533	0.59	0.684	0.88	1.59
$\hat{\eta}$	1.449	0.72	0.38	0.449	0.60
MTTF (Years)	121285.59	412.508	70.6	31.20	16.08

The process of acquiring more failure data improves the accuracy of the estimation process and reduces the dependency of the inference scheme on the prior information. In [89], the authors present a case study on the impact of censoring on the MSE using noninformative priors. In this work, we draw a comparison between noninformative and informative priors. The results of the experiments, as shown in Fig. 5.15 seem to suggest that the Bayes estimators and hence the marginal posteriors are affected by observed data and the expert knowledge when informative priors are used, which is a desirable outcome. In Fig. 5.15, it is clear that larger sample sizes of observed failures have the effect of reducing the MSE of the estimators. Further, the MSE is less

when informative priors are used than noninformative priors. This underscores the importance and impact of prior knowledge available to the distribution system engineer.

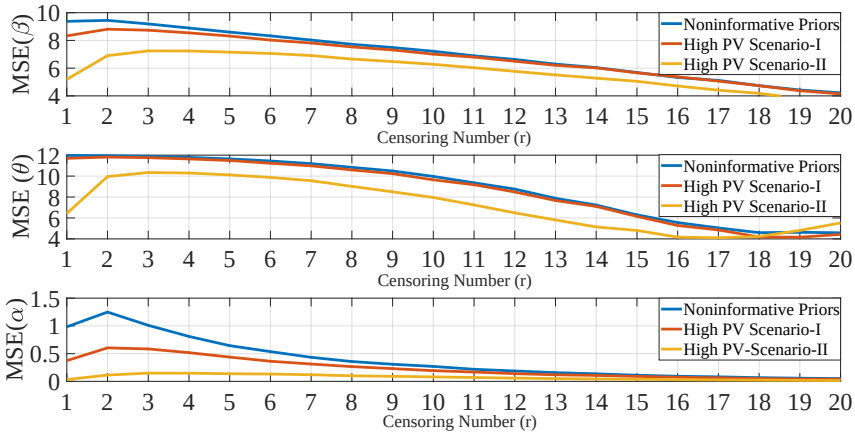

Figure 5.15: Mean Square with Censoring Number

5.5.4 Uncertainty Characterization of Model Parameters

To generate samples from the marginal posterior of parameters that characterize the failure model, we consider informative priors (gamma prior)on the shape and scale parameter of (5.13) and uninformative prior (uniform prior) on the stress parameter of (5.13). The proposal density for generating new states is equal to the parameters' prior density. With this choice of the proposal density, the probability that the proposed state vector will be accepted or rejected is given by (5.39). Since the point estimates of the parameters are in the domain of the parameter posteriors, the starting state vector is set equal to the vector of the point estimates. To calculate the likelihood ratio in (5.39), the failure data is normalized to make the numerical approximations tractable. The results of the MCMC simulations that include the trace plots, marginal posteriors and the autocorrelation plots of the model parameters are shown in Figure 5.16, Figure 5.17, Figure 5.18 and Figure 5.19.

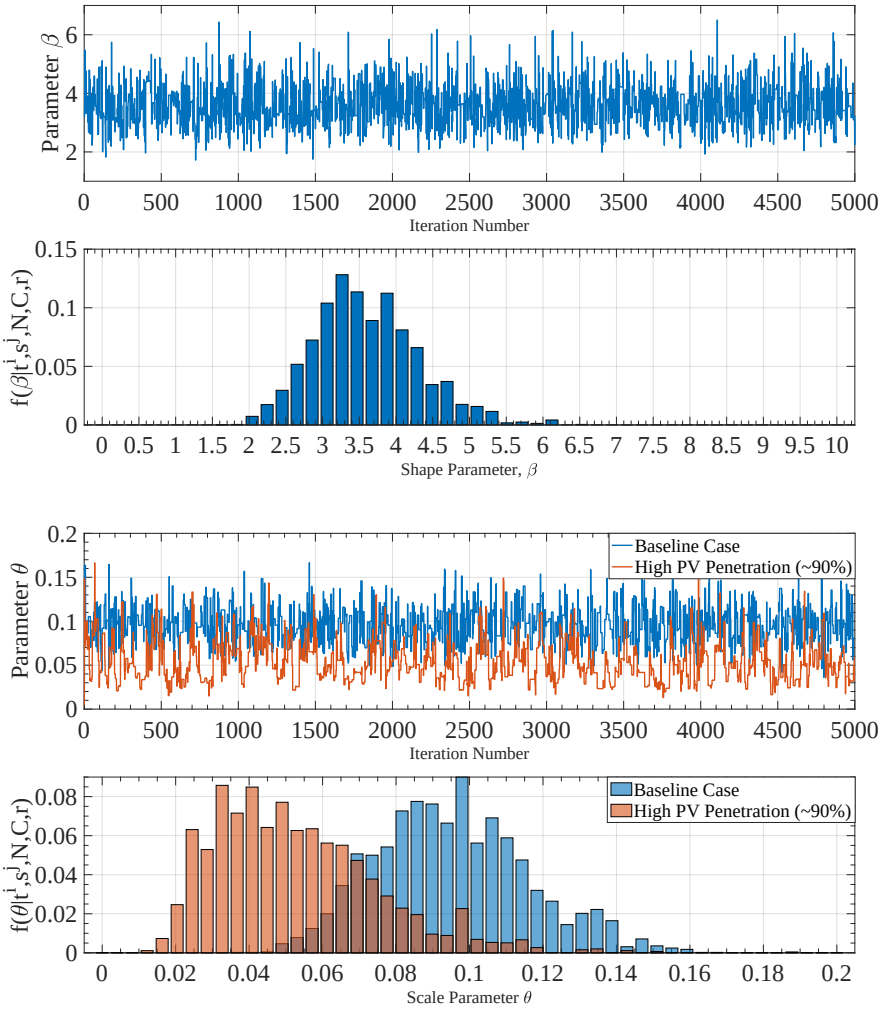

Figure 5.17: Trace Plot and Marginal Posterior of θ with $N = 30, r = 20$

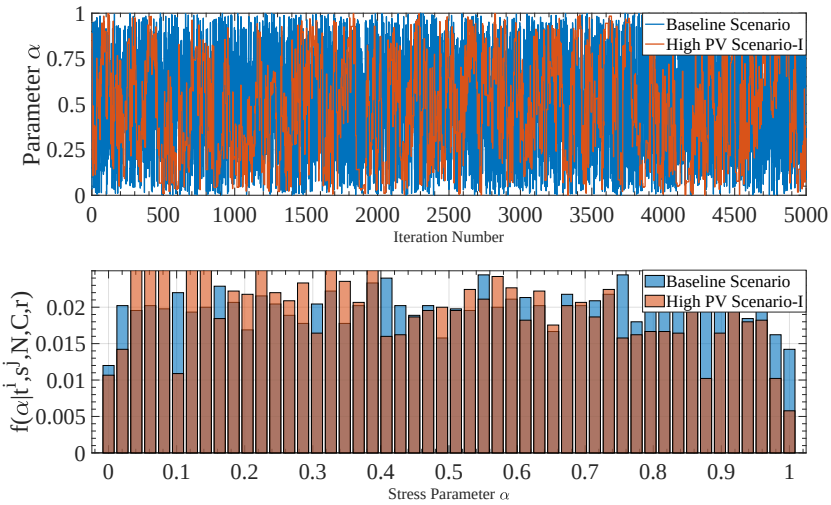

Figure 5.18: Trace Plot and Marginal Posterior of α with $N = 30, r = 20$

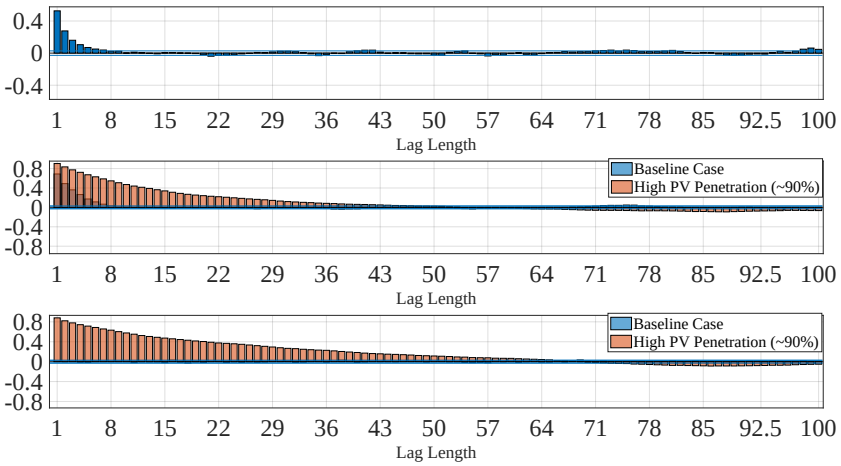

Figure 5.19: MCMC Autocorrelation with $N = 30, r = 20$

5.5.4.1 Convergence Diagnostics

Since the marginal posterior of a parameter forms the basis for inferring the statistical properties of the unknown parameter, it is important to ensure that the inference based on the MCMC sample is valid. The inference process's validity depends on the Markov chain's convergence. Trace and autocorrelation plots are the two most accessible convergence diagnostics techniques that can be visually inspected. The trace plots of the model parameters as shown in Fig. 5.16, Fig. 5.17, Fig.5.18 have well mixing chains. The shape parameter β is invariant to changes in the operational stress, whereas the scale parameter θ and the stress parameter α are affected as the operational stress changes. At a higher stress level corresponding to high penetration of distributed solar, the Markov chain that traces the sample path of the parameter θ experiences some difficulty converging to the steady-state distribution. Also, the marginal posterior of θ is shifted to the left under conditions of high PV penetration, thus suggesting a decrease in the equipment life at a higher stress level. The marginal posterior of the parameter α resembles a quasi-uniform distribution, and the Markov chain in both scenarios shows good mixing. The acceptance rate of the Markov chains for all the model parameters is listed in Table 5.11. The convergence of the Markov chains can also be assessed from the autocorrelation plot shown in Fig. 5.19. The shape

parameter β autocorrelation becomes negligible very quickly after about eight lags. The same is

Table 5.11: Convergence Diagnostics

Model Parameter	Acceptance Rate		
	Baseline Scenario	Scenario-I	Scenario-II
Shape β	33.72%	33.72%	33.88%
Scale θ	22.76%	18.92%	18.54%
Stress α	90.02%	36.86%	37.50%

true for the scale parameter θ where the autocorrelation gets fairly close to zero after about ten lags in the baseline scenario. However, the autocorrelation becomes negligible at a higher stress level after about fifty lags. This results from relatively poor mixing of the chain compared to the baseline scenario. It is important to mention that even though the autocorrelation takes longer to become negligible in a high PV scenario, it does so after approximately fifty lags, suggesting that the Markov chain has converged to a steady-state distribution. For the stress parameter α, the autocorrelation is essentially negligible for all positive lags for the baseline case, but at a higher stress level, the Markov chain experiences similar convergence issues as that of scale parameter θ. However, as was true for the scale parameter, the autocorrelation becomes negligible after approximately sixty-four lags.

5.5.4.2 Credible Intervals

The idea of credible intervals or credible region for a vector-valued parameter space has a very intuitive understanding in a Bayesian paradigm. For a one-dimensional parameter $\Theta \in \Omega$ or a vector-valued $\hat{\Theta} \in \Omega \subset \mathrm{R}^d$, and a confidence level $z \in (0,1)$, the interval $I_z \subset \Omega$ or the region $I_z \subset \Omega$ that contains the proportion of $1-z$ of the probability mass of the posterior is called a credible interval or a credible region respectively.

$$\Pr(\hat{\Theta} \in I_z | T = t) = 1 - z \tag{5.41}$$

Equation (5.41) is a general expression and does not determine a unique credible interval. It is possible to impose some additional conditions for choosing credible intervals or regions. For example, an equal-tailed interval of confidence level z is an interval

$$I_z = [q_{z/2}, q_{(1-z)/2}] \tag{5.42}$$

In (5.42) q_z is $z-$ quantile of the posterior distribution. In this work we utilize 95% equal-tailed interval defined as

$$I_{0.05} = [q_{0.025}, q_{0.975}] \tag{5.43}$$

where $q_{0.025}$ and $q_{0.075}$ are quantiles of marginal posterior of the parameters. The quantile functions

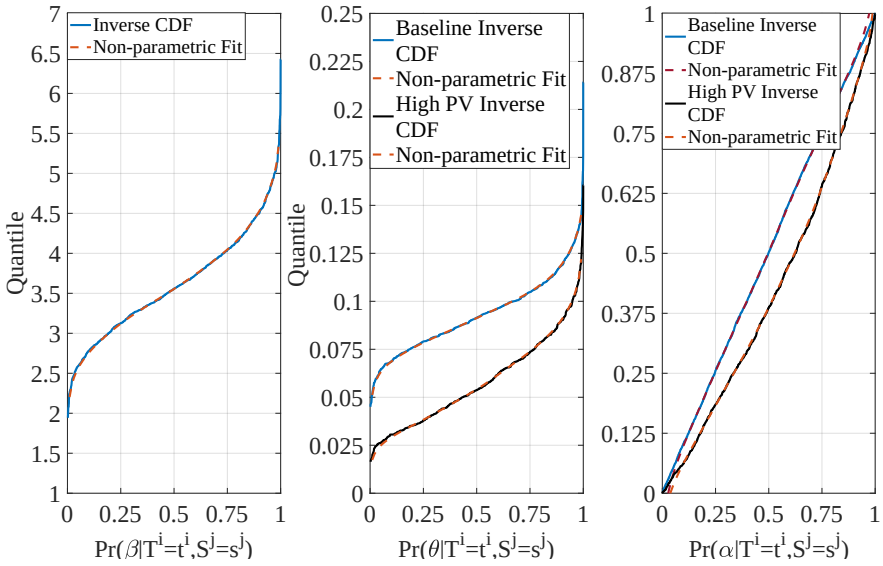

Figure 5.20: Marginal Inverse CDFs of Model Parameters, $N = 30, r = 20$

of the model parameters are shown in Fig. 5.20 and Table 5.12 lists the 95% equal-tailed credible intervals of the parameters.

Table 5.12: Credible Intervals

Model Parameter	Credible Intervals		
	Baseline Scenario	Scenario-I	Scenario-II
Shape β	[2.05,4.83]	[1.996,4.94]	[2.29, 5.16]
Scale θ	[0.0563,0.138]	[0.0184,0.106]	[0.0182,0.106]
Stress α	[0.0031,0.97]	[0.0045,0.985]	[0.004,0.99]

5.6 Conclusions

This chapter proposes a statistical parametric model to accurately assess the impact of non-thermal stress on mechanically operated voltage control assets such as OLTCs and switched capacitor banks. High penetration of distributed proportional roof-top PV generation is used to model the non-thermal stress. The non-thermal or operational stress is quantified in terms of the number of cumulative operations over a defined planning horizon. The failure model is a combination of the inverse power law, typically used to model the device lifetimes under non-thermal stress, and a two-parameter Weibull distribution, the hazard function of which is consistent with the degradation induced by the gradual aging. We consider Type-II censored survival data and propose an inferential scheme based on Bayes's rule to estimate the parameters of the proposed failure model. In this work, we develop point estimates of the model parameters based on a quadratic loss function using informative prior beliefs. The resulting point estimates are shown not to possess closed-form solutions and thus evaluated using numerical integration techniques. A simulation procedure based on the MCMC sampling technique is also proposed to evaluate the full posterior of the parameters using informative priors and censored failure data set.

The proposed methods are applied to a model of the IEEE-34 test feeder assumed to have either newly installed OLTCs or have survived till the present moment. Based on the simulation results obtained, some of the conclusion that can be drawn are as follows

- A high penetration of distributed solar PV negatively impacts the service life of devices such

as OLTCs. The simulation results indicate a shortening of the device lifetime by as much as $27\% - 44\%$, depending on the choice of the prior belief.

- In the proposed inferential scheme, the properties of the model parameters are inferred by combining the prior beliefs with the survival data. It has been shown that different prior beliefs can lead to different predictions. However, as the sample size grows, the model predictions from different prior beliefs begin to converge. This is consistent with the Bayesian paradigm, in which a large sample size dominates prior beliefs.

- A comparison between the performance of noninformative priors and informative priors is made by numerically computing the MSE of the estimators. It is shown that for larger sample sizes, the prior choice does not significantly impact the MSE. For small sample sizes, however, informative priors are associated with lower MSE than noninformative priors.

- The full posterior of the parameters is evaluated using the MCMC sampling procedure. The Markov chains constructed show good convergence characteristics for baseline and high PV scenarios. While the shape and the stress parameters are largely unaffected by changes in the operational stress, the scale parameter experiences a shift in the posterior distribution at higher stress levels. This shift is indicative of a shorter lifespan at increased stress levels.

6. POWER QUALITY IMPROVEMENT WITH DISTRIBUTION STATIC COMPENSATOR

The increasing adoption of distributed energy resources (DERs), particularly solar generation and the use of unconventional loads such as plug-in electric vehicles (PHEVs), has a profound impact on the planning and operation of electric distribution systems. In particular, PHEV charging introduces stochastic peaks in energy consumption, while solar generation is fraught with variability during intermittent clouds. The stochastic nature of such DERs renders the operation of mechanical assets such as on-load tap changers and switched capacitor banks ineffective. A possible solution to mitigate the undesirable effects of DERs is using solid-state-based devices such as a distribution static synchronous compensator (D-STATCOM). This chapter focuses on the operation and the impact of a solid-state-based device called capacitor-less Distribution Static Compensator (D-STATCOM) in electrical distribution systems. Various steady-state models of the D-STATCOM in different operating modes are presented. An important problem is the capacity usage of a capacitor-less D-STATCOM when integrated with the distribution system. This chapter formulates the capacity usage problem in probabilistic terms taking into account the uncertainties of the distributed energy resources. A Monte Carlo simulation framework is proposed to study the capacity usage problem with DER inputs sampled from the proposed underlying distributions.

In electric distribution systems, reactive power compensation in voltage regulation and power factor correction is typically realized using electromechanical assets such as on-load tap changers (OLTCs) and switched capacitor banks (SCBs). However, given the mechanical nature of the operation, such legacy devices are not suitable for providing reactive power compensation at shorter time scales of minutes or seconds. The output of a PV system is a function of weather primarily at a given location and can experience rapid changes that necessitate a continuously adjustable reactive power compensation for precise voltage control. One class of devices capable of providing dynamic voltage control is the smart PV inverters. However, the PV inverters are limited in providing continuous reactive support constrained by the inverter's apparent power sizing [**?**]. The PV inverters are also not owned by the utility and, in most situations, do not respond to the dispatch

commands issued by the network operator [91].

Other examples of power electronics-based solid-state devices that mitigate the adverse impact of DERs include static var compensators (SVCs) at transmission or sub-transmission voltage levels and D-STATCOMs based on voltage source converter (VSC) technology. A major disadvantage of the SVCs is the lack of harmonic control functionality and exorbitant capital costs in low voltage regimes [2]. Due to these reasons, SVCs are inherently unsuitable for use in low-voltage networks plagued by power quality issues due to the increased use of nonlinear loads. On the other hand, the VSC-based D-STATCOMs rely on electrolytic capacitors (E-caps) for energy storage. The reliance on E-caps adversely affects the reliability of the VSC-based D-STATCOMs, especially in locations with tropical climate conditions [2]. The literature on the reliability of the power electronics devices has established that nearly 30% of all the failures in power electronics-based devices are caused by the E-caps [91].

Table 6.1: Functionality comparison of Capacitor-less D-STATCOM with incumbent technologies [2]

Functionality	OLTC	SCB	SVC	Hybrid VARs	**Smart PV Inverters**	D-STATCOM	**Capacitor-less D-STATCOM**
Dynamic Power Factor Correction		✓	✓	✓	✓	✓	✓
Dynamic Voltage Regulation			✓		✓	✓	✓
Network Dispatchability	✓	✓	✓	✓		✓	✓
Longer Service Life	✓						✓
Harmonic Filtering					✓	✓	✓

In light of these observations, a capacitor-less D-STATCOM based on a matrix converter (MC) has been proposed recently to address the dual challenges of fast reactive support and increased reliability [92]. The newly proposed capacitor-less D-STATCOM uses inductive storage and is controlled using a finite control set model predictive control (MPC). The capacitor-less D-STATCOM is a multi-functional device that can be used simultaneously for power factor correction, voltage regulation, and harmonic compensation with local autonomous control or directly controlled by

the distribution network operator (DSO). A comparison between the capacitor-less MC-based D-STATCOM and the incumbent technologies is given in [2]

A major significance of this chapter is that it studies the interaction of a power electronics converter in terms of its capacity usage when integrated with an electric distribution system without neglecting the details. A complete model of the IEEE-34 bus distribution test system [50] is developed in OpenDSS [65]. The steady-state power system models of the capacitor-less D-STATCOM in different modes of operation are developed in a MATLAB environment and interfaced with the OpenDSS solver. It is important to emphasize that while the focus of this paper is to develop a probabilistic capacity usage model of the capacitor-less D-STATCOM, a rigorous treatment of the computational methods used for uncertainty quantification of the inputs of the Monte Carlo simulation, which in this case are bus load, PHEV charging, and rooftop solar generation, is crucial to answering the central question.

6.1 Capacitor-less D-STATCOM

The capacitor-less D-STATCOM performs the same functions as the existing VSC-based D-STATCOM. A distinction, however, can be drawn in terms of the circuit elements used for energy storage. The capacitor-less D-STATCOM uses inductors for energy storage as opposed to electrolytic capacitors. This enables the capacitor-less D-STATCOM to achieve a much longer service life, especially in regions with harsh climatic conditions [91]. The capacitor-less D-STATCOM is designed to provide reactive power compensation to simultaneously address the multiple challenges of power factor correction, voltage regulation, and harmonic mitigation. The proposed converter can work either in a local autonomous mode and determine the amount of reactive power to be injected, absorbed, or centrally dispatched by the system operator. This multiplicity of operation is a benefit that a utility can derive from the converter, and the resulting monetary savings can offset the high capital cost of the device for use in low-voltage distribution systems.

6.2 Converter Topology

The three-phase matrix converter (MC) with nine bidirectional switches, a three-phase input filter, and output chokes is the fundamental building block of the proposed converter. The bidirectional switches are realized using two anti-parallel IGBT-diode pairs for bidirectional flow. The shunt-connected MC-based capacitor-less D-STATCOM is shown in Figure 6.1 .

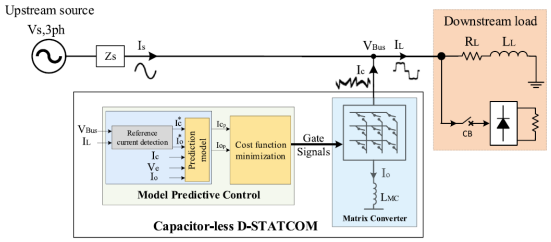

Figure 6.1: Matrix Converter-based capacitor-less D-STATCOM [2]

The converter design uses Model Predictive Control (MPC) strategy to achieve phase current inversion between the input and output currents. The input and output current relations of the MC are given in (6.1)

$$\begin{bmatrix} I_{cA} \\ I_{cB} \\ I_{cC} \end{bmatrix} = \begin{bmatrix} S_{Aa} & S_{Ab} & S_{Ac} \\ S_{Ba} & S_{Bb} & S_{Bc} \\ S_{Ca} & S_{Cb} & S_{Cc} \end{bmatrix} \begin{bmatrix} I_{0A} \\ I_{0B} \\ I_{0C} \end{bmatrix} \tag{6.1}$$

The input and output voltages share the same relationship. The switching function S_{ij} can be either 0 or 1, where $i \in [A,B,C]$ and $j \in [a,b,c]$. The phase reversals are achieved by making good choices about S_{ij}, which makes inductive energy storage appear capacitive at the input of the MC. More details on the MC are given in [92].

6.3 Power Factor Correction Operation

In OpenDSS, the capacitor-less D-STATCOM in power factor correction mode is modeled as an ideal current source. The compensating currents and the reactive powers for each phase of the load are calculated based on phase admittance and the phase voltage. In general, with a three-phase

load with phase A admittance $Y_A = G_A + jB_A$, and phase voltage $\overline{V}_A$the phase current is

$$\bar{I}_A = \overline{V}_A Y_A = V_A \angle \theta_A (G_A + jB_A) = I_{RA} + jI_{XA} \tag{6.2}$$

The capacitor-less D-STATCOM as an ideal current source is shunt connected to compensate for the phase quadrature component I_{XA} of the phase A load current. If the phase A compensating current is I_{CA} and $I_{CA} = -I_{XA}$, the phase A converter rating is

$$S_{CA}^{pf} = \overline{V}_A \bar{I}_{CA}^{*} = \overline{V}_A (-j\overline{V}_A B_A) = j\overline{V}_A^2 B_A \tag{6.3}$$

Equation (6.3) suggests that only the reactive power demand of phase A of the load is compensated while the real power demand remains unaffected. Since the reactive power demand of phase A of the load is $Q_A = -\overline{V}_A^2 B_A$, the converter rating can be expressed solely in terms of Q_A.

$$S_{CA}^{pf} = -jQ_A = jQ_{CA}^{pf} \tag{6.4}$$

From (6.4) we can conclude that if the compensator, the ideal current source modeling the steady state behavior of the the capacitor-less D-STATCOM in this case, is desired to provide power factor correction, the condition for unity power factor operation of phase A of the load is $Q_{CA}^{pf} = -Q_A$. Partial compensation can be achieved if $|Q_{CA}^{pf}| = \alpha |Q_A|, 0 < \alpha < 1$. The phasor compensating current required to achieve full compensation (unity power factor) of Phase A is

$$\bar{I}_{CA} = \frac{S_{CA}^{pf}}{\overline{V}_A^{*}} = \frac{-jQ_{CA}^{pf}}{\overline{V}_A^{*}} = \frac{-j\left(S_A \sqrt{1 - \cos^2 \phi_A}\right)}{\overline{V}_A^{*}} \tag{6.5}$$

In (6.5) Q_{CA}^{pf} is the reactive power rating of the converter, S_A is the apparent power of phase A of the load, and $\cos \phi_A$ is the phase A power factor. The compensatory ideal current source model of (6.2-6.5) is implemented in MATLAB environment and interfaced with OpenDSS solver engine via the Component Object Model (COM) interface. The compensating currents for each phase of the load are updated after successfully converging the power flow and injected into the load. The total three-phase reactive power rating of the converter in power factor correction mode is the sum

of reactive power ratings for each phase. That is $Q_{3\phi}^{pf} = Q_{CA}^{pf} + Q_{CB}^{pf} + Q_{CC}^{pf}$.

6.4 Voltage Regulation Operation

The power system steady-state model of the capacitor-less D-STATCOM in voltage regulation mode is based on the reactive power mismatch equations. A proportional-integral controller is used to minimize the mismatch between the reference voltage (voltage set-point) and the measured bus voltage. The first two equations model the reactive power exchange between the converter and the load bus, and the third equation models the control scheme.

$$0 = \begin{bmatrix} Q_{3\phi}^{vr} - V_L I_{line} \sin \theta_{V_L} - \theta_{I_{line}} \\ Q_{3\phi}^{vr} + |V_{PCC}|^2 B_l - |V'| G_l \sin \delta + |V'| B_l \cos \delta \\ V_L - V_{sp} \end{bmatrix} \tag{6.6}$$

In (6.6) $Q_{3\phi}^{vr}$ is the total three phase reactive power exchanged between the converter and the load bus, B_l and G_l are the line susceptance and conductance respectively connecting the converter and the bus, V_L, θ_{V_L} is the bus voltage magnitude and angle respectively,I_{line}, $\theta_{I_{line}}$ are the line current magnitude and angle respectively, $|V'| = |V_L||V_{conv}|$, V_{conv} is the converter voltage magnitude, $\delta = \theta_{V_L} - \theta_{V_{conv}}$, $\theta_{V_{conv}}$ is the converter voltage angle and V_{sp} is the voltage set-point of the converter.

6.5 Harmonic Filtering

Under normal operation conditions, the voltage and current waveforms at any node of a power system are considered to be perfectly sinusoidal at a frequency of 60 or 50 Hz. However, when a source of electrical power is connected to a nonlinear device, the current drawn by the nonlinear device is not a perfect sinusoid. The non-sinusoidal current interacts with the system impedance to create voltage harmonics and in some cases, harmonic resonance. More commonly, voltage or current harmonics are components at frequencies integer multiples of the fundamental frequency. It is possible to reconstruct a distorted and periodic waveform by an infinite summation of such spectral components via Fourier analysis. In power systems, the problem of harmonics dates back to the 20th century, when the subject was first developed to understand the waveform distortion

caused by static converters.

Over the last five decades, the nature of the loads in medium and low-voltage industrial and residential distribution feeders has radically changed with rapid advances in semiconductor technology. As a result, the distribution systems now experience higher levels of harmonic distortion due to the higher penetration of power electronic controlled equipment. This problem is exacerbated by the use of power factor correction capacitor banks in distribution feeders, making phenomena like harmonic resonance much more likely. The resonance curve of a typical distribution feeder is often broad-shaped, which renders the feeder sensitive to a range of harmonic orders. Moreover, it is not uncommon to have multiple capacitor banks installed over the length of the feeder for power factor and voltage control applications. Hence, computer simulations are required to accurately assess the feeder behavior at frequencies at or near the resonant frequency. The operation of a distribution feeder close to resonant frequency can amplify the voltage and current distortion levels. Furthermore, the change in feeder capacitance affects the resonance curve of a distribution feeder. The optimization of the capacitor banks to regulate the power factor at the substation or the local level can interfere with the resonant frequency of the feeder. In such a scenario, there is a need to track the resonant frequency as the feeder loads change to accurately predict the distortion levels in voltage and current waveforms. The modeling of the harmonic filtering functionality of the D-STATCOM is based on the computation of harmonic flows and the use of the harmonic load model in OpenDSS.

6.5.1 Computation of Harmonic Flows

A nonlinear load connected to the power grid, which supplies sinusoidal voltages with no harmonics, will draw a non-sinusoidal current from a sinusoidal source. Mathematically, the non-sinusoidal current can be expressed as the sum of infinite sinusoids with different frequencies, where higher frequencies are integer multiples of some fundamental frequency. At the point of interconnection, the distorted current and voltage waveform can be expressed in the time domain

as

$$i(t) = \sum_{k=1}^{\infty} I_k \cos(k\omega_1 t + \phi_{ki}) \tag{6.7}$$

$$v(t) = \sum_{k=1}^{\infty} V_k \cos(k\omega_1 t + \phi_{kv}) \tag{6.8}$$

In power systems, the transformers block the flow of DC, and loads do not produce a DC component. In (6.7) and (6.8), I_k and V_k are the peak magnitudes of k^{th} harmonic and ϕ_{kI} and $\phi_k V$ are the associated phase angles, respectively. The distorted current drawn by the load has to be supplied from the source in the absence of any harmonic filtering equipment. Hence, the harmonic distortion propagates upstream into the network. The computation of harmonic flows begins with the conventional power flow solution of the test system under consideration. The snap-shot power flow solution establishes the fundamental voltage magnitudes and phase angles. The power flow solution can also be used to look for any irregularities in the test system or in the input data. The slack bus in the system provides the phase angle reference and is assumed to have no distortion. Once the fundamental nodal voltages are known, the harmonic components are considered, one at a time. Calculating the harmonic nodal voltages is the same for all harmonic orders. To summarize, for each harmonic, k

- The nodal admittance, Y_k with appropriate positive, negative, and zero sequence network is constructed. The admittance matrix is of the form

$$Y_k = \begin{bmatrix} Y_{11} & \cdots & Y_{1n} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \cdots & Y_{nn} \end{bmatrix}$$

  where $Y_i j$ is the mutual admittance between buses i and j and $Y_i i$ is the self-admittance of bus i at $k^t h$ harmonic order. Furthermore, in a three-phase unbalanced system, the elements

of the admittance are 3×3 matrices, which consist of self and transfer admittances, i.e.

$$Y_k = \begin{bmatrix} Y_{aa} & Y_{ab} & Y_{ac} \\ Y_{ba} & Y_{bb} & Y_{bc} \\ Y_{ca} & Y_{cb} & Y_{cc} \end{bmatrix}$$

- The $k^t h$ injection current of each nonlinear load is determined from the harmonic spectrum of the load. The harmonic spectrum of the load is a user-defined object which contains the percentage contribution of each harmonic with respect to the fundamental. The magnitude and the angle of $k^t h$ harmonic current is given by

$$I_k = I_1 * k \tag{6.9}$$

$$\phi_k = \phi_k + k(\phi_1 - \theta) \pm 180^\circ \tag{6.10}$$

  where, I_k is the magnitude of $k^t h$ harmonic current, I_1 is the magnitude of the fundamental, K is the corresponding multiplier as defined in the harmonic spectrum of the nonlinear load, ϕ_k is the phase angle of the $k^t h$ harmonic injection, ϕ_1 is the phase angle of the fundamental and θ is the phase angle of the slack bus.

- Once I_k and ϕ_k are known for each nonlinear load, the system harmonic voltages are determined by solving

$$[\bar{I}_k] = [\overline{Y}_k][\overline{V}_k]; k \neq 1 \tag{6.11}$$

The steady-state harmonic filtering model of the capacitor-less D-STATCOM is developed in OpenDSS using a component object model (COM) server for writing custom code. To compute the harmonic flows, the feeder loads are converted into Norton Equivalents, consisting of a current source and an admittance. Figure 6.2 shows the harmonic load model used in the OpenDSS. The shunt admittance comprises a series R-L part and a parallel R-L part. The parameters G, B, R, and L are obtained from the load's user-specified real and reactive powers at 100% rated voltage.

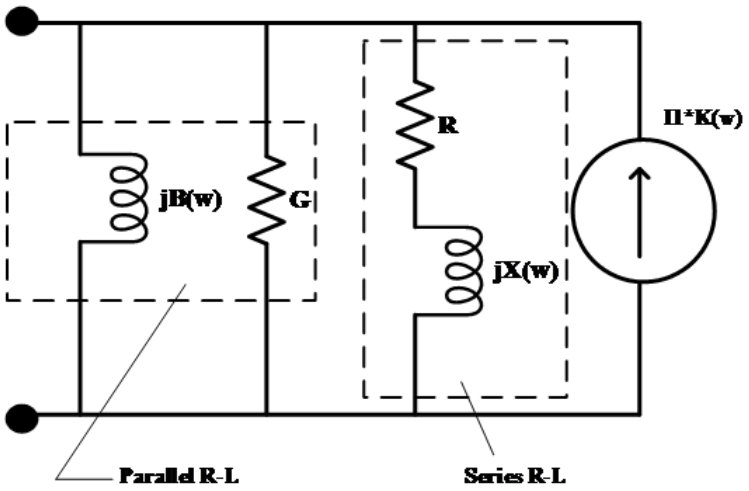

Figure 6.2: Harmonic Load Model

The shunt admittance of the load plays an important role in damping out the harmonic components of the distorted load current at frequencies near or equal to the resonant frequency. This is because in the case of parallel resonance, the system impedance seen by the harmonic source increases dramatically and a significant portion of the harmonic load current is siphoned off in the shunt admittance of the load model. At frequencies away from the resonant frequency, the load impedance is much greater than the system equivalent impedance, and very little current is channeled into the shunt admittance.

Usually, the percentage mix of the series and the parallel R-L part of the load is not known but in most scenarios, a 50/50 distribution tends to produce more realistic-looking results. A 100% parallel R-L load provides the highest damping and hence the lowest distortion while a 100% series R-L load yields high values of harmonic distortion. The series R-L part of the load admittance can also be used to model a rotating machine load. In this case, the resistance and inductance parameters are determined by blocked rotor impedance or sub-transient impedances. While it is true that the load admittance does not impact the resonant frequency in any significant manner, it does, however, provide damping to the harmonics, the amount of which depends on the type of admittance model used.

6.5.2 Harmonic Filtering with D-STATCOM

To achieve harmonic filtering the D-STATCOM is modeled as a variable current source and is programmed to cancel the harmonic currents drawn by the nonlinear load. These are steady-state models and the phasor quantities of voltage and current at the PCC are used to define the system parameters. The task of harmonic mitigation involves measuring the harmonic content of the load current and injecting the canceling current. An example of harmonic cancellation is given in Table 6.2. The input to the compensator is the measured harmonic currents at the terminals of the non-linear load. Given the harmonic spectrum of the load and the daily variation, the sequential-time

Table 6.2: Harmonic Filtering with D-STATCOM

Harmonic Order	Load Current	Cancelling Current
3	10∠30	-10∠-30+180
5	20∠ 60	-20∠ 60+180
7	14∠ -7	-14∠ -7+180
9	2∠ 51	-2∠ -50+180

harmonic simulations are carried out by computing the harmonic flows as described by 6.9, 6.10 and 6.11. The distribution system solver evaluates the power system response at each frequency taken separately and at each time step. In the next step, the loads on the feeder are updated, and the process of computing the flows is performed sequentially.

6.6 Capacity Usage Determination of D-STATCOM*

[1] The probabilistic capacity usage model of the capacitor-less D-STATCOM is developed by Monte Carlo sampling of the input variables. For a given penetration level of the solar generation, the input variables of the Monte Carlo simulation are the bus load profile, PHEV charging profile, and PV generation profile. The use of Monte Carlo sampling entails that the capacity usage Q be a random variable that can assume a range of values. When the random variable Q is indexed by time, it represents a stochastic process $\{Q_t; 0 \leq t < \infty\}$ defined on the probability space Ω_Q. The stochastic process Q_t is continuous and its distribution is described by the probability density

[1] Reprinted with permission from A. Peerzada, M. Begovic and D. Ostojic, "Carbon Tax and Utility-scale Solar Deployment," 2022 IEEE Power & Energy Society General Meeting (PESGM), 2022, pp. 01-05, doi: 10.1109/PESGM48719.2022.9917021

$f_{Q_t}(q,t)$. The density function $f_{Q_t}(q,t)$ describes the joint distribution of the random variables $\{Q_t; 0 \leq t < \infty\}$. The Monte Carlo simulation estimates the expectation of any Borel-measurable function $h : \mathbb{R} \longrightarrow \mathbb{R}$ of the stochastic process Q_t.

$$\mathbf{E}[h(Q_t)] = \int_{q \in \Omega_Q} h(Q_t) f_{Q_t}(q,t) dq \tag{6.12}$$

The Monte Carlo estimator of (6.12) can be obtained by running Monte Carlo simulations with inputs sampled from the underlying distributions. The inputs considered for estimating (6.12) are the PHEV charging, load profile, and solar generation profile.

6.6.1 Generating Load Scenarios

The generalized Gaussian mixture model described in chapter 2 is used to generate load scenarios with similar statistical properties as the measured (empirical) load data. Bus 890 of the test feeder is assumed to be a commercial load. This assumption is based on the peak active and reactive load ratings provided in the original feeder data developed by the IEEE distribution subcommittee. The measured historical annual commercial load profile (without PHEV charging) is obtained from the OpenEI website [60]. The web page is sponsored by the U.S Department of Energy (DoE) in support of the Open Government Initiative to make energy data transparent and collaborative. The measured historical annual data set is used to estimate the parameters of the generalized Gaussian mixture model using the E-M algorithm. The E-M algorithm is coded in MATLAB and initialized using the K- means algorithm. A random number generator function generates statistically similar load profile scenarios. The random number generator function can be evaluated by taking the inverse of the CDF of the mixture model. The CDF of the mixture model in (3.6) is

$$F_Y(y|\Psi) = \int_{-\infty}^{y} \sum_{j=1}^{M} \pi_j \frac{\beta_j}{2 s_j \Gamma\left(\frac{1}{\beta_j}\right)} \exp\left[-\frac{|t-\mu_j|^{\beta_j}}{s_j}\right] dt \tag{6.13}$$

A closed form solution of $\hat{y} = F_Y^{-1}(y|\Psi)$ does not exist and hence numerical techniques such as Newton-Raphson must be used to generate random samples from the fitted model. Figure 6.3

shows a hundred examples of the load at bus 890 for one day. It is important to note that the load scenarios shown in Figure 6.3 also account for the PHEV charging.

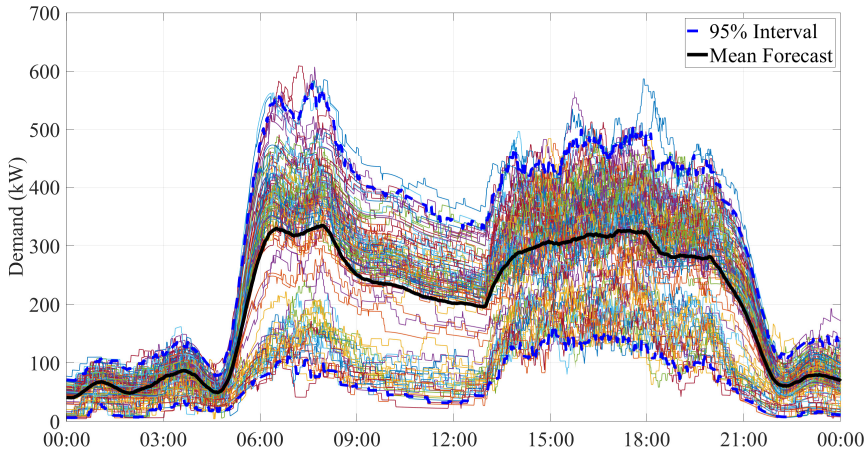

Figure 6.3: Load Scenarios at Bus 890 including PHEV charging Profiles

6.6.2 Distribution Generation Scenarios

The scenarios for solar generation at bus 890 are generated using the Monte Carlo sampling of the fitted time series ARMA model. We assume a three-phase PV system at bus 890 with a peak power rating of 450 kW proportional to the peak active load. The voltage rating of the PV system is 4.16 kV. The PV system is oriented at an azimuth of 180 °(south-facing) and a tilt angle of 30 °. Since the IEEE-34 bus test feeder is based on an existing distribution system located in the state of Arizona (AZ), the hourly average TMY3[62] data of AZ is used to synthesize a high resolution (1-minute) irradiance profile. The solar output in 1-minute intervals is estimated using the PV system model in OpenDSS[65]. The high-resolution solar output data is used to fit the ARMA model. The fully realized ARMA model is used to generate solar forecasts. Figure 6.4 shows one hundred 1-minute ahead solar output scenarios for a typical spring day in AZ along with the mean forecast and 95% confidence interval. The parameters of the realized ARMA model are given in Table 6.3

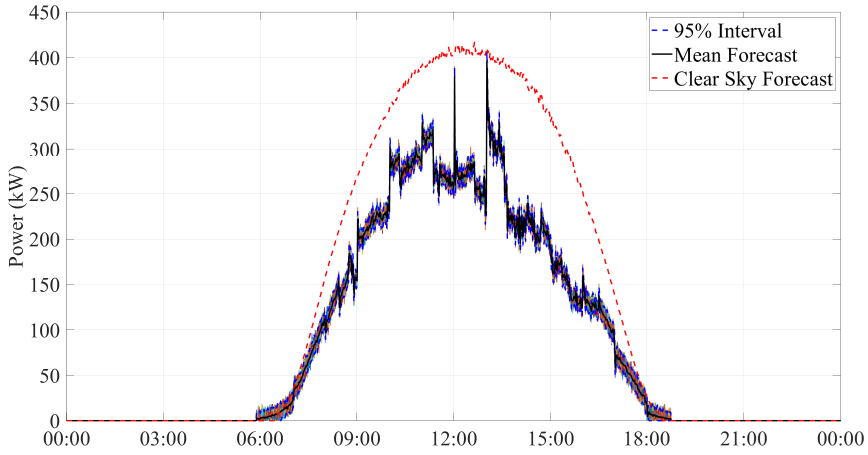

Figure 6.4: 1-minute ahead solar forecasts from ARMA model from Table 6.3

Table 6.3: The Realized ARMA (p,q) Model

p	q	ϕ_i	θ_j
3	3	$\phi_1 = 0.376$ $\phi_2 = -0.3722$ $\phi_3 = 0.9848$	$\theta_1 = 0.6151$ $\theta_2 = 0.9845$ $\theta_3 = -0.005$

The generated scenarios of the input variables are used to calculate the Monte Carlo estimator of (**??**). This is done by taking the i.i.d sample $\{q_\gamma(t); \gamma = [1,2,...H,], t = [1,2,...,T]\}$. H is the total number of Monte Carlo runs and T is the time horizon. The time horizon considered in this work is one year. The mean of $(g(q_\gamma(t)))$ over the chosen sample is the estimate of $\mathbf{E}[h(Q_t)]$.

$$\hat{h}_\gamma\left(q_\gamma(t); t \in [1,T]\right) = \frac{1}{\gamma}\sum_{v=1}^{\gamma} h\left(q_v(t); t = \in [1,T]\right) \tag{6.14}$$

Assuming the expectation $\mathbf{E}[h(Q_t)]$ is finite, the weak law of large numbers implies for an arbitrarily small ε

$$\lim_{\gamma\to\infty}\left(\Pr\{\,\hat{h}_\gamma\left(q_\gamma(t); t \in [1,T]\right) - \mathbf{E}[h(Q_t)] \geq \varepsilon\}\right) = 0 \tag{6.15}$$

Equation (6.15) implies that as γ gets large, the Monte Carlo estimator converges in probability to

the true expectation. The i.i.d sample $\{q_\gamma(t); t = [1,2,...,T]\}$ for fixed γ is obtained by executing the power flow on the model of IEEE 34 bus test feeder in OpenDSS interfaced with MATLAB via COM. The capacitor-less D-STATCOM is placed at bus 890 of the test feeder. The power factor correction and voltage regulation programs are developed in MATLAB. The circuit is solved in OpenDSS with control actions suspended at first. This is done to sample the quantities of interest, i.e., bus voltage phasors and load reactive power for D-STATCOM initialization. Based on the retrieved values of the bus voltages and load, the programs populate the current injection sources that model D-STATCOM operation in power factor correction mode to compensate for the reactive power of the load. At the same time, the voltage regulation subroutine uses the reactive power mismatch equations to calculate the reactive power for maintaining bus voltage at a predefined set point. The modified circuit with D-STATCOM is solved again in OpenDSS with control actions enabled. Once the control queue clears, the program steps through the next solution.

The probability distributions of the capacity usage for power factor correction application $Q_{3\phi}^{pf}$ and voltage regulation $Q_{3\phi}^{vr}$ application considering the uncertainty in PHEV charging, commercial reactive demand and the solar generation is shown in Figure 6.5 . The capacity usage, as shown

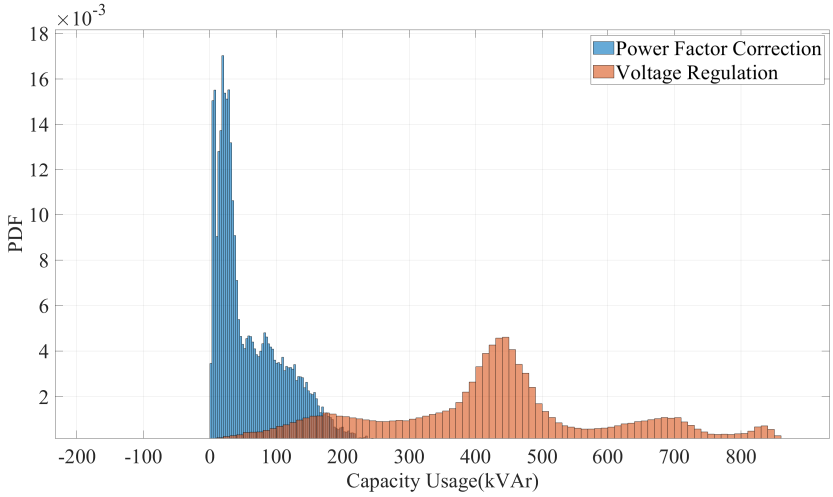

Figure 6.5: Annual Capacity Usage with unity load power factor and voltage 1 p.u

in Figure 6.5 represents the unity power factor operation of the load at bus 890 of the feeder with a voltage set point of 1.0 p.u. The D-STATCOM adjusts its reactive output in accordance with (6.2-6.5) and (6.6). This ensures that the power factor of the load is unity at all times, and the voltage is held nearly constant at 1.0 p.u. Figure 6.5 seems to suggest that voltage regulation is a more capacity intense operation than power factor correction. The expected value of the capacity usage for power factor correction is 121.95 kVAr, and for voltage regulation, the expected value of capacity usage is 474.62 kVAr. This is approximately four times higher than the power factor application. Furthermore, the standard deviation of $Q_{3\phi}^{pf}$ is 88.6 kVAr while the standard deviation of $Q_{3\phi}^{vr}$ is 162.95 kVAr. A higher standard deviation of $Q_{3\phi}^{vr}$ suggests greater uncertainty in its estimation. This could be because of the dual impact of intermittent solar generation and PHEV charging on the bus voltage. Solar power and PHEV charging are both characterized by a higher variability and hence increased uncertainty which propagates further and results in an output with a higher variance. Figure 6.6 shows the capacity usage when the power factor of the load is 0.8 lagging and the bus voltage is held constant at 0.95 p.u.

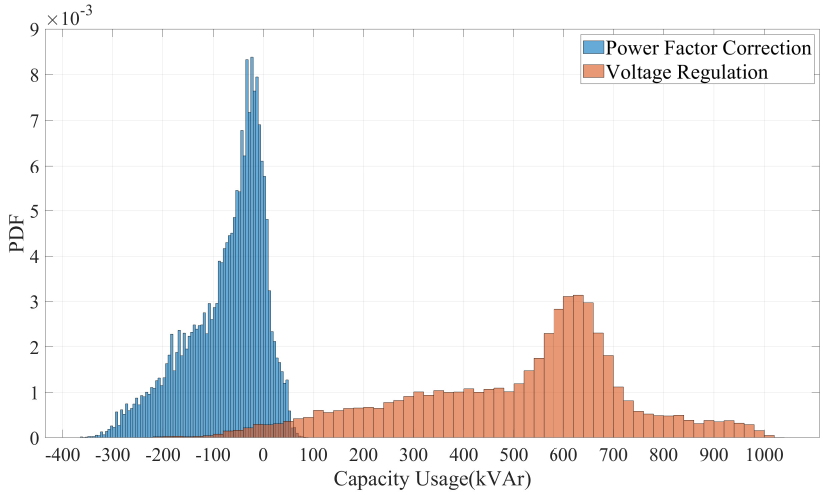

Figure 6.6: Annual Capacity Usage with non-unity load power factor and voltage 0.95 p.u

An interesting application of the capacity usage results in Figure **??** and **??** is the determination of the energy loss due to the curtailment of the PV output when capacity sizing constraints on

the converter are considered. The results in Figure **??** are valid for an unrestrained D-STATCOM. However, a restrained D-STATCOM will be limited in its ability to provide reactive power compensation. The limits placed on the converter capacity could either be due to the high manufacturing cost or the technology limitations that could preclude scaling the converter capacity. Since voltage regulation application utilizes more capacity, any limits on the converter capacity would primarily impede the converter's ability to regulate voltage. In order to maintain the bus voltage within acceptable limits, especially the upper limit of $1.05p.u$, additional measures such as curtailing the PV system's output must be put into effect. To observe the impact of the capacity reduction of the D-STATCOM on the energy loss due to the curtailment of PV output, we use the volt-watt control functionality in OpenDSS. It provides a flexible mechanism to regulate the active power output of a PV system based on a user-configured volt-watt control curve. For our application, we use a volt-watt curve that reduces the PV output whenever the bus voltage exceeds $1.05p.u$. If $\triangle g^i_{PV}(\omega)$ is the curtailed PV power at bus i for scenario ω, the energy loss over the planning horizon is

$$E_{loss}(\omega) = \int_0^T \triangle g^i_{PV}(\omega) dt \tag{6.16}$$

$E_{loss}(\omega)$ is a random variable and we are interested in the expectation $\mathbf{E}[E_{loss}(\omega)]$ as a function of the capacity reduction of the D-STATCOM. Figure 6.7 shows the cdf plots of energy loss of a 450 kW PV system at bus 890. The converter is programmed to maintain the bus voltage at $1p.u$ and the capacity is reduced in steps of 10% from the baseline capacity of $+/-900$ kVAr. It can be inferred from Figure 6.7 that the energy loss probability increases with the converter capacity reduction. For capacity reductions up to 20% there is roughly 70% probability of incurring an annual energy loss of 5% or less. However, the probability of energy loss dramatically increases as the converter capacity is further reduced.

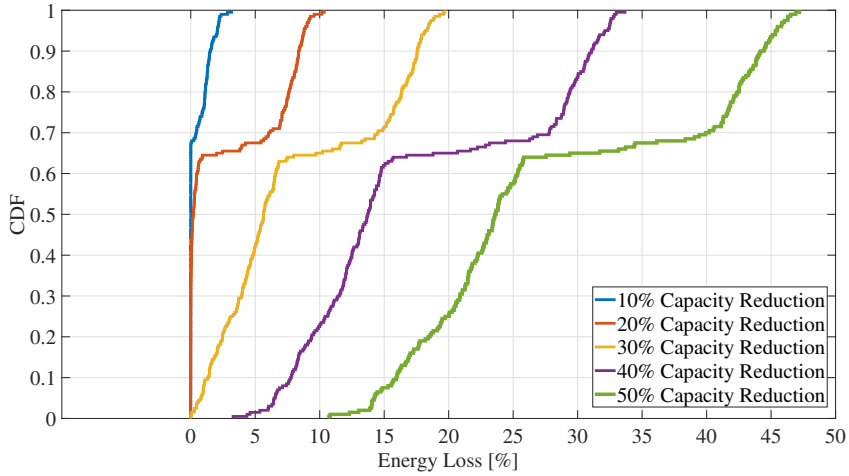

Figure 6.7: CDF plots of $E_{loss}(\omega)$ for different converter capacities

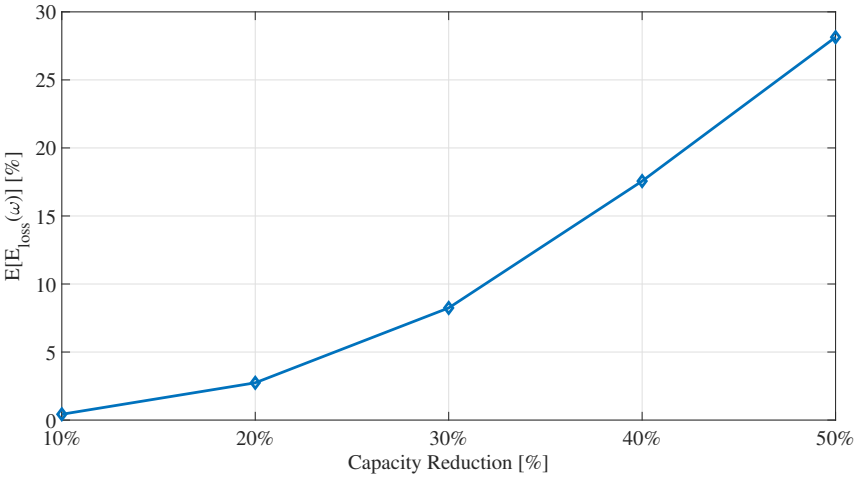

Figure 6.8: Expected value of Annual Energy loss of a 450 kW PV System

Figure 6.8 plots the expectation $\mathbf{E}[E_{loss}(\omega)]$ as a function of the capacity reduction of the proposed converter. A 50% reduction in the converter capacity would result in an annual energy loss of nearly 30% of the value if the converter was operating at full capacity with a voltage setpoint of $1.0p.u.$ On the other hand, a capacity reduction up to 20% could be acceptable as the energy loss is less than 5%.

6.7 Conclusions

This chapter studied the capacity usage of the proposed capacitor-less D-STATCOM in a distribution system, considering the uncertainties in PHEV charging, system demand, and solar generation. Two different modes of operation of the converter are presented. The power factor correction operation is modeled as an ideal current source shunt connected to the load. The voltage regulation mode of operation is modeled based on the reactive power mismatch equations between the converter and the load bus. To assess the uncertainty in capacity usage, Monte Carlo simulations are designed with inputs sampled from their underlying distributions. Various scenario generation algorithms consistent with the physics and based on real-world measured data are presented to sample the input parameters. Based on the results of the Monte Carlo runs, it is concluded that voltage regulation operation utilizes more capacity than power factor correction and is characterized by a higher variance. Furthermore, restraining the converter's capacity increases the energy loss due to the curtailment of the PV output. More precisely, the expected value of the energy loss shares a nonlinear relationship with the capacity reduction of the converter. The energy loss increases with the progressive decrease in the converter's capacity.

The methodology and the results presented in this paper are important and can be used as a tool by the manufacturers to improve the design and functionalities of the power electronic converters for distribution systems. In the future, we plan on extending this work to include more complex distribution systems with more components, such as the IEEE-123 bus feeder and the 8500 node test feeder.

7. SUMMARY AND CONCLUSIONS

Today's electric utility industry is in the midst of an aggressive expansion and an unprecedented evolution due to the proliferation of distributed energy resources (DERs) and the implementation of the smart grid concept. In addition, the growing realization of a reliable and secure electric power system as a fundamental component of the next industrial revolution has ushered in transformative changes in the operation and planning of transmission and distribution grids. The changing landscape of the energy flow has prompted utilities to invest in critical transmission and distribution infrastructure to improve the grid "resiliency." The shift from centralized energy infrastructures to a distributed one and coupled with the growing capacity of renewable sources of energy, has not only led to a redefinition of the term "resiliency", but has also helped shape an obscure theoretical term from just being a buzzword to a cornerstone of the infrastructure planning. In 2017, the electric utilities in the United States spent $50 billion dollars in capital expenditures to maintain and upgrade the distribution grid, followed by annual average spending of $50 billion dollars through 2018 [**?**].

Every major economy on the planet has witnessed a rapid growth in the penetration of renewable sources and, more recently, in energy storage. More importantly, these resources are increasingly being deployed in a distributed manner, thus disrupting the more traditional "top-down" structure of the electrical grid. Even though the contribution of the distributed resources might not be significant at present, these are projected to grow at an accelerated rate. The evidence for the accelerated expansion of such distributed resources is plenty and is indicative of a new beginning in the power sector-one that will result in major upheavals and restructuring. In light of these facts, the primary goal of this research is to investigate the operational and planning issues associated with the adoption of high-capacity distributed energy resources. Integrating DERs, especially high capacity distributed generation, introduces multiple layers of complexity to the normal operation of transmission or distribution networks. To ensure the security and reliability of the system, these complexities need to be rigorously investigated, and solutions must be sought that would facilitate

increased adoption of DERs while at the same time improving network resiliency.

7.1 Summary

This dissertation focuses primarily on the applications of computational statistics in power systems. More specifically, a decision-making framework to facilitate the widespread adoption of distributed energy resources (DERs) is presented. The successful integration of DERs with the electric grid is a complex problem that nearly impacts every aspect of power system planning, operation, and reliability. The work presented in this dissertation studies a small subset of the challenges imposed by DER integration, particularly in load modeling and asset management. Also, the impact of a power electronics-based solid-state power quality converter on the operation and planning of distribution systems is studied. In addition, a case study on the impact of a carbon tax on the profitability of investing in a utility-scale solar generation is presented.

The work presented in chapter 2 has interesting applications. Since the proposed mixture model is parametric and hence "generative"; the model can generate synthetic load data with similar statistical properties as the measured data. Since the proposed model considers the EV charging load, the mixture models proposed have far-reaching applications such as probabilistic load flow with EV charging and distribution system state estimation (DSSE), where many pseudo measurements are used to run the state estimation algorithms. The mixture models studied can model the non-Gaussian distributed measurements, especially scenarios involving heavy EV penetration and distributed generation. Another important application of the proposed mixture models is designing Monte Carlo simulations where the inputs are sampled from some underlying distributions. The mixture models can be used to create different realizations of the electric demand curve and as input to run stochastic optimization algorithms. A multivariate formulation of the mixture model is presented to address the problem of modeling load correlation. The update equations are obtained using the E-M algorithm. The multivariate model can generate correlated random numbers and thus can be used to model correlation among loads in close proximity to each other. This is significant since the information about one load can be used to predict the load at a nearby location in a power system.

Chapter 3 of the dissertation focuses on the uncertainty characterization of solar generation. Two computational models are presented to quantify the uncertainty associated with solar generation. The Markov Chain Monte Carlo simulation model synthesizes high-resolution solar data from the low-resolution TMY3 data. This is achieved by extracting the transition probabilities from the measured high-resolution non-TMY3 data and feeding the probability matrices to a first-order Markov chain. The second model is based on time-series modeling and uses an auto-regressive moving average model to develop new scenarios of the solar data. It is essentially a scenario generation algorithm that exploits the time-series structure of the measured solar data.

In chapter 4, a case study on the profitability of investing in a solar generation is presented, leveraging the modeling capabilities of chapters 2 and 3. The case study is an investigation of the impact of carbon tax policy on the total operational cost incurred to the utility when investing in a large-scale solar generation to offset the carbon emissions from the coal-fired and gas plants. A Monte Carlo simulation framework is developed to assess the impact of a carbon tax on the marginal benefit of carbon abatement. The results demonstrate a complex interplay between several factors and thus call for a more nuanced approach toward policy-making and regulation. A major implication of this work is that the targets set for a specific technology to decarbonize the power grid cannot be applied as a "one size fits all" across different adoption levels of the same technology and different power utilities.

Chapter 5 of the dissertation focuses on an interesting asset management problem in power systems. Specifically, the impact of high-capacity solar generation on the equipment lifetimes such as on-load voltage tap changers and switched capacitor banks is studied in detail. A predictive failure model that learns from the distribution of past failures and can model the life-shortening effect of high-capacity solar generation is developed. This is a significant improvement over the simple failure models in the literature, which largely ignore the problem of accelerated aging. A Bayesian framework to estimate the parameters of the failure distribution is presented. A computational algorithm based on Markov chain Monte Carlo sampling of the known function is utilized for the evaluation of full posteriors. Determining the full posterior distributions of the model parameters is

important to assess the uncertainty in the parameter estimation process. The failure model is tested on a real distribution test system to get the failure probabilities of the devices still in operation based on the past failure data and the future growth of the solar generation. It is shown that the failure model is quite robust and the convergence is fast as more failure data are acquired.

In chapter 6, the computational models developed in the earlier chapters are utilized to derive a probabilistic capacity-usage model of a power electronics-based solid state power quality converter when interfaced with a distribution system model. The power quality converter, referred to as a capacitor-less D-STATCOM is a multi-functional device used for power factor correction, voltage regulation, and harmonic filtering. Steady-state models of the D-STATCOM are developed to observe the impact on the distribution grid. A Monte Carlo simulation is designed to assess the capacity usage of the D-STATCOM operation in different modes with inputs sampled from the underlying distributions. The energy loss due to the solar curtailment with different capacity reductions of the D-STATCOM is studied. Estimating the energy loss is important to appropriately size the power quality converter, given the constraints on the capital cost of the device.

7.2 Conclusions

The research presented in this dissertation explores some of the operational and reliability challenges associated with integrating DERs with the electric grid. To that end, computational models are developed to properly assess those challenges and design mitigation measures enabling large-scale adoption of DERS while satisfying various operational and economic constraints. Some of the broad conclusions of the research work can be summarized as follows

- The first operational problem addressed in this work is the impact of DERs on the overall power system load. It is observed that systems with high-capacity DERs have distinct load profiles with a pronounced "peaky" behavior. The density plots of the load exhibit multi-modal characteristics with multiple high-density regions. A probability mixture model with component densities derived from generalized versions of probability distributions such as Gaussian and beta is formulated to accurately model the multimodal behavior of the load. Generalized distributions such as a generalized Gaussian or a generalized beta have addi-

tional parameters that control the shape of the distribution and are thus better suited to model the peak load behavior. However, the parameter estimation of the generalized mixture model is computationally intense due to the nonlinear parameter update equations. This is a major disadvantage, and the decision to use generalized mixture models to model the pdf of the DER-impacted load ultimately depends on the desired accuracy, which further depends on the application.

- The second major problem addressed in this dissertation is the economic viability of large-scale solar farms considering policies such as carbon tax. A carbon tax rate is an economic policy that penalizes electric utilities for each unit of carbon dioxide (metric ton) released into the atmosphere by burning fossil fuels such as coal and gas to produce electric power. The challenge is addressed by formulating a metric called marginal benefit of carbon abatement that quantifies the percentage change in the total operational cost incurred to the utility. It is concluded that the overall economic benefit of investing in solar generation depends on the amount of the solar generation capacity and the carbon tax rate. In the scenario where coal-fired plants are progressively retired, solar generation capacity greater than 20% of the peak system load is necessary to return an overall profit on the investment.

- The third problem addressed in this dissertation is related to the accelerated wear and tear of the mechanically-operated equipment in distribution systems. It is concluded that equipment such as on-load tap changers and switched capacitor banks placed on the distribution systems for voltage control and loss reduction can experience a severe derating of the lifetime if exposed to high-capacity solar generation. The simulation results indicate a shortening of the device lifetime by as much as $27\% - 44\%$, depending on the choice of the prior belief.

- The fourth problem addressed in this dissertation is the investigation of the impact of a solid-state power quality converter on the distribution system operation. A major conclusion drawn from the Monte Carlo simulations is that the voltage regulation function of the converter is the most capacity-intense function, followed by power factor correction. The

expected value of the capacity usage in voltage regulation mode is roughly twice that of the capacity usage in power factor correction mode. The analysis of expected power loss due to solar curtailment reveals that in the absence of the D-STATCOM the expected power loss is roughly 30% higher than with the D0STATCOM running at full capacity. This underscores the importance of using solid-state-based devices in conjunction with mechanically-switched devices in distribution systems.

REFERENCES

[1] C. Martin, F. Starace, and J. P. Tricoire, "The Future of Electricity: New Technologies Transforming the Grid Edge," *World Economic Forum*, 2017.

[2] W. Rohouma, M. Metry, R. S. Balog, A. A. Peerzada, M. M. Begovic, and D. Zhou, "Analysis of the Capacitor-Less D-STATCOM for Voltage Profile Improvement in Distribution Network With High PV Penetration," *IEEE Open Journal of Power Electronics*, vol. 3, pp. 255–270, 2022.

[3] O. Hafez and K. Bhattacharya, "Queuing analysis based PEV load modeling considering battery charging behavior and their impact on distribution system operation," *IEEE Transactions on Smart Grid*, 2018.

[4] "EIA Hourly Grid Monitor."

[5] Q. Chen and D. M. Egan, "A bayesian method for transformer life estimation using Perks' hazard function," *IEEE Transactions on Power Systems*, 2006.

[6] US EPA, "Inventroy of U.S Greenhouse Gas Emissions and Sinks," *Epa 430-R-21-005*, 2021.

[7] M. Begovic, A. Peerzada, R. Nuqui, and B. Picone, "Locational accuracy of VIP indices for voltage collapse margin estimation," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2018.

[8] A. Soroudi and T. Amraee, "Decision making under uncertainty in energy systems: State of the art," 2013.

[9] Z. Fikri, *Statistical load analysis for distribution network planning*. PhD thesis, Royal Institue of Technology Stockholm Sweden, 1975.

[10] EPRI, "Selected Statistical Methods for Analysis of Load Research Data- EA-3467," tech. rep., Electric Power Research Institute, 1984.

[11] G. Irwin, W. Monteith, and W. Beattie, "Statistical electricity demand modelling from consumer billing data," *IEE Proceedings C Generation, Transmission and Distribution*, 1986.

[12] A. Seppala, "Statistical distribution of customer load profiles," in *Proceedings of the International Conference on Energy Management and Power Delivery, EMPD*, 1995.

[13] R. Herman and J. J. Kritzinger, "The statistical description of grouped domestic electrical load currents," *Electric Power Systems Research*, 1993.

[14] A. K. Ghosh, D. L. Lubkeman, M. J. Downey, and R. H. Jones, "Distribution circuit state estimation using a probabilistic approach," *IEEE Transactions on Power Systems*, 1997.

[15] S. W. Heunis and R. Herman, "A Probabilistic Model for Residential Consumer Loads," 2002.

[16] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Transactions on Power Systems*, 2010.

[17] G. Valverde, A. T. Saric, and V. Terzija, "Probabilistic load flow with non-Gaussian correlated random variables using Gaussian mixture models," *IET Generation, Transmission and Distribution*, 2012.

[18] B. Stephen, A. J. Mutanen, S. Galloway, G. Burt, and P. Jarventausta, "Enhanced load profiling for residential network customers," *IEEE Transactions on Power Delivery*, 2014.

[19] J. Quiros-Tortos, A. Navarro-Espinosa, L. F. Ochoa, and T. Butler, "Statistical representation of EV charging: Real data analysis and applications," in *20th Power Systems Computation Conference, PSCC 2018*, 2018.

[20] M. Cui, C. Feng, Z. Wang, and J. Zhang, "Statistical representation of wind power ramps using a generalized Gaussian mixture model," *IEEE Transactions on Sustainable Energy*, 2018.

[21] B. Srbinovski, A. Temko, P. Leahy, V. Pakrashi, and E. Popovici, "Gaussian mixture models for site-specific wind turbine power curves," *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 2021.

[22] F. Ge, Y. Ju, Z. Qi, and Y. Lin, "Parameter estimation of a Gaussian mixture model for wind power forecast error by riemann L-BFGS optimization," *IEEE Access*, 2018.

[23] J. Zhang, S. Zhang, F. Gao, X. Li, H. Li, Z. Wang, and C. Shen, "An analytical method for probabilistic load flow using Gaussian mixture model," in *2016 IEEE International Conference on Power System Technology, POWERCON 2016*, 2016.

[24] Z. Wang, C. Shen, F. Liu, J. Wang, and X. Wu, "An Adjustable Chance-Constrained Approach for Flexible Ramping Capacity Allocation," *IEEE Transactions on Sustainable Energy*, 2018.

[25] Z. Guo, P. Pinson, S. Chen, Q. Yang, and Z. Yang, "Chance-Constrained Peer-to-Peer Joint Energy and Reserve Market Considering Renewable Generation Uncertainty," *IEEE Transactions on Smart Grid*, 2021.

[26] K. G. Hollands and R. G. Huget, "A probability density function for the clearness index, with applications," *Solar Energy*, 1983.

[27] D. McCracken Alexander, *Synthetic High Resolution Solar Data*. PhD thesis, University of Strathclyde, 2011.

[28] M. Jurado, J. M. Caridad, and V. Ruiz, "Statistical distribution of the clearness index with radiation data integrated over five minute intervals," *Solar Energy*, 1995.

[29] G. Y. Saunier, T. A. Reddy, and S. Kumar, "A monthly probability distribution function of daily global irradiation values appropriate for both tropical and temperate locations," *Solar Energy*, 1987.

[30] A. v. Schijndel, *Power Transformer Reliability Monitoring*. PhD thesis, Eindhoven University of Technology, 2010.

[31] G. K. Ari and Y. Baghzouz, "Impact of high PV penetration on voltage regulation in electrical distribution systems," in *3rd International Conference on Clean Electrical Power: Renewable Energy Resources Impact, ICCEP 2011*, 2011.

[32] R. Yan, B. Marais, and T. K. Saha, "Impacts of residential photovoltaic power fluctuation on on-load tap changer operation and a solution using DSTATCOM," *Electric Power Systems Research*, 2014.

[33] M. Lave, M. J. Reno, and R. J. Broderick, "Characterizing local high-frequency solar variability and its impact to distribution studies," *Solar Energy*, 2015.

[34] J. Bank and B. Mather, "Analysis of the impacts of distribution connected PV using high-speed datasets," in *IEEE Green Technologies Conference*, 2013.

[35] C. Hanley, G. Peek, J. Boyes, G. Klise, J. Stein, D. Ton, and T. Duong, "Technology development needs for integrated grid-connected PV systems and electric energy storage," in *Conference Record of the IEEE Photovoltaic Specialists Conference*, 2009.

[36] Y. P. Agalgaonkar, B. C. Pal, and R. A. Jabr, "Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators," *IEEE Transactions on Power Systems*, 2014.

[37] J. Yan, "Bayesian Survival Analysis," *Journal of the American Statistical Association*, 2004.

[38] D. Martin, J. Marks, T. K. Saha, O. Krause, and N. Mahmoudi, "Investigation into Modeling Australian Power Transformer Failure and Retirement Statistics," *IEEE Transactions on Power Delivery*, 2018.

[39] M. S. Minhas, J. P. Reynders, and P. J. de Klerk, "Failures in power system transformers and appropriate monitoring techniques," in *IEE Conference Publication*, 1999.

[40] V. Sokolov, S. Tsurpal, and A. Drobyshevski, "Reliability problems with lange power transformers and shunt reactors. Typical failure modes and failure causes," in *Cigré A2 Colloqium*, 2005.

[41] D. Zhou, Z. Wang, P. Jarman, and C. Li, "Data requisites for transformer statistical lifetime modelling - Part II: Combination of random and aging-related failures," *IEEE Transactions on Power Delivery*, 2014.

[42] W. Li, "Evaluating Mean Life of Power System Equipment with Limited End-of-Life Failure Data," *IEEE Transactions on Power Systems*, 2004.

[43] Y. Hong, W. Q. Meeker, and J. D. McCalley, "Prediction of remaining life of power transformers based on left truncated and right censored lifetime data," *Annals of Applied Statistics*, 2009.

[44] M. Begovic, P. Djuric, J. Perkel, B. Vidakovic, and D. Novosel, "New probabilistic method for estimation of equipment failures and development of replacement strategies," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2006.

[45] M. Begovic and P. Djuric, "On predicting the times to failure of power equipment," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2010.

[46] C. M. Affonso and M. Kezunovic, "Probabilistic assessment of electric vehicle charging demand impact on residential distribution transformer aging," in *2018 International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2018 - Proceedings*, 2018.

[47] C. D. M. Affonso and M. Kezunovic, "Technical and economic impact of pv-bess charging station on transformer life: A case study," *IEEE Transactions on Smart Grid*, 2019.

[48] F. Pascal, L. Bombrun, J. Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized gaussian distributions," *IEEE Transactions on Signal Processing*, 2013.

[49] F. Pascal, Y. Chitour, J. P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, 2008.

[50] K. P. Schneider, B. A. Mather, B. C. Pal, C. W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. Pereira, L. F. Ochoa, L. R. De Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott,

and W. Kersting, “Analytic Considerations and Design Basis for the IEEE Distribution Test Feeders,” *IEEE Transactions on Power Systems*, 2018.

[51] M. Muratori, “Impact of uncoordinated plug-in electric vehicle charging on residential power demand,” *Nature Energy*, 2018.

[52] A. Papoulis and J. G. Hoffman, “ Probability, Random Variables, and Stochastic Processes ,” *Physics Today*, 1967.

[53] R. Pasupathy, “Generating Nonhomogeneous Poisson Processes,” in *Wiley Encyclopedia of Operations Research and Management Science*, 2011.

[54] P. A. Lewis and G. S. Shedler, “SIMULATION OF NONHOMOGENEOUS POISSON PROCESSES BY THINNING.,” *Naval research logistics quarterly*, 1979.

[55] Z. Chen, X. Y. Xiao, C. S. Li, Y. Zhang, and Z. X. Zheng, “Study on Unit Commitment Problem Considering Large-Scale Superconducting Magnetic Energy Storage Systems,” *IEEE Transactions on Applied Superconductivity*, 2016.

[56] C. a. Driver, “Best Selling EVs,” 2021.

[57] M. G.-V. . J. M. E. Gómez, “A multivariate generalization of the power exponential family of distributions, Communications in Statistics - Theory and Methods,” *Communications in Statistics - Theory and Methods*, vol. 27:3, 589-, 2007.

[58] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions: Second Edition*. 2007.

[59] D. Steinley and M. J. Brusco, “Evaluating Mixture Modeling for Clustering: Recommendations and Cautions,” *Psychological Methods*, 2011.

[60] Office of Energy Efficiency & Renewable Energy, “Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States - Datasets - OpenEI Datasets,” 2015.

[61] T. MathWorks, “MATLAB (R2020b),” *The MathWorks Inc.*, 2020.

[62] S. Wilcox, “National Solar Radiation Database 1991–2010 Update: User’s Manual ,” *Nrel/Tp-5500-54824*, 2012.

[63] D. A. Dickey and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *Journal of the American Statistical Association*, vol. 74, p. 427, 6 1979.

[64] "System Identification Toolbox Documentation."

[65] R. C. Dugan, "The Open Distribution System Simulator (OpenDSS)," tech. rep., 2012.

[66] "Energy Information Agency: Today in Energy," 2021.

[67] M. Liu, P. Reed, and C. L. Anderson, "Stochastic Synthetic Data Generation for Electric Net Load and Its Application," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021.

[68] G. B. Shrestha, K. Song, and L. Goel, "Strategic self-dispatch considering ramping costs in deregulated power markets," *IEEE Transactions on Power Systems*, 2004.

[69] R. Bhat, M. Begovic, I. Kim, and J. Crittenden, "Effects of PV on conventional generation," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2014.

[70] I. Kim, M. Begovic, H. Jeong, and J. Crittenden, "Impact of photovoltaic distributed generation on generation resource allocation," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2013.

[71] W. Katzenstein and J. Apt, "Air emissions due to wind and solar power," *Environmental Science and Technology*, 2009.

[72] S. Blumsack and K. Richardson, "Cost and Emissions Implications of Coupling Wind and Solar Power," *Smart Grid and Renewable Energy*, 2012.

[73] I. Kim and M. Begovic, "On impact of randomly distributed PV systems on distribution networks," in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2016.

[74] Sandia National Laboratories, "PVPMC, PV_LIB Toolbox."

[75] D. Lew, G. Brinkman, N. Kumar, P. Besuner, D. Agan, and S. Lefton, "Impacts of Wind and Solar on Fossil-Fueled Generators," *Proceedings of the IEEE PES General Meeting 2012*, 2012.

[76] International Energy Agency, "Projected Costs of Generating Electricity 2020 Edition," tech. rep., International Energy Agency, 2020.

[77] A. Peerzada, M. Begovic, and D. Ostojic, "On the Environmental and Economic Impact of Utility Scale Renewable Energy Deployment [Accepted for Publication]," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022.

[78] X. Liu, A. Aichhorn, L. Liu, and H. Li, "Coordinated control of distributed energy storage system with tap changer transformers for voltage rise mitigation under high photovoltaic penetration," *IEEE Transactions on Smart Grid*, 2012.

[79] Y. Liu, J. Bebic, B. Kroposki, J. De Bedout, and W. Ren, "Distribution system voltage performance analysis for high-penetration PV," in *2008 IEEE Energy 2030 Conference, ENERGY 2008*, 2008.

[80] CIGRE, "An International Survey of Failures in Large Power Transformers in Service," 1983.

[81] A. Drobyshevski, V. Sokolov, and S. Tsurpal., "Reliability problems with large power transformers and shunt reactors. typical failure modes and failure causes.," in *Transformer Reliability and Transients Moscow Russia*, (Moscow), pp. 20–23, 0.

[82] P. d. Klerk, M. Minhas, and J. Reynders, "Failures in power system transformers and appropriate monitoring techniques.," in *High Voltage Engineering Symposium*, p. No 467, 0.

[83] R. E. Brown, *Electric power distribution reliability*. 2017.

[84] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*. 1996.

[85] Wasserman, "All of Statistics : A Concise Course in Statistical Inference Brief Contents," *Simulation*, 2004.

[86] X. Zhang, E. Gockenbach, V. Wasserberg, and H. Borsi, "Estimation of the lifetime of the electrical components in distribution networks," *IEEE Transactions on Power Delivery*, 2007.

[87] M. Begovic and A. Peerzada, "Impact of Large Distributed Solar PV Generation on Distribution Voltage Control," in *Proceedings of the 52nd Hawaii International Conference on System Sciences.*, pp. 3473–3482, HICSS, 2019.

[88] H. Rinne, *The Weibull Distribution*. 2008.

[89] A. Peerzada, M. M. Begovic, W. Rohouma, and R. Balog, "On Estimation of Equipment Failures in Electric Distribution Systems Using Bayesian Inference," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021.

[90] A. Banerjee and D. Kundu, "Inference based on type-II hybrid censored data from a Weibull distribution," *IEEE Transactions on Reliability*, 2008.

[91] W. Rohouma, R. S. Balog, A. A. Peerzada, and M. M. Begovic, "D-STATCOM for harmonic mitigation in low voltage distribution network with high penetration of nonlinear loads," *Renewable Energy*, 2020.

[92] W. Rohouma, M. Metry, R. S. Balog, A. A. Peerzada, and M. M. Begovic, "Adaptive Model Predictive Controller to Reduce Switching Losses for a Capacitor-Less D-STATCOM," *IEEE Open Journal of Power Electronics*, vol. 1, pp. 300–311, 2020.

APPENDIX A

INVERSE POWER LAW-WEIBULL FAILURE MODEL

A.1 Survival Function, Hazard Rate and CDF of IPL-Weibull Failure Model

This appendix describes the survival function and other related formulas of the IPL-Weibull distribution.

A.1.1 Survival Function of IPL-Weibull

The survival function is the probability of surviving beyond a certain time. A random variable T that characterizes the failure time of a device, the survival function is given by

$$S_T(t) = \Pr(T > t|\mathbf{X}) = \int_t^\infty f_T(u)du \tag{A.1}$$

If T has a Weibull distribution with the scale parameter replaced by the inverse power law, the failure density function $f_T(u)$ in (A.1) is the density of the IPL-Weibull distribution and $\mathbf{X}$ is the vector of parameters. This gives for the survival function of the IPL-Weibull distribution

$$S_T(t) = \int_t^\infty \beta\theta t^{\beta-1}\gamma^\alpha \exp(-\gamma^\alpha \theta t^\beta)dt \tag{A.2}$$

Consider the integral

$$I = \lim_{l\to\infty}\int_t^l t^{\beta-1}\exp(-\gamma^\alpha\theta t^\beta)dt = \frac{1}{\beta\gamma^\alpha\theta}\exp(-\gamma^\alpha\theta t^\beta) \tag{A.3}$$

Thus the survival function of an IPL-Weibull distributed random variable T is

$$S_T(t) = \exp(-\gamma^\alpha\theta t^\beta) = \exp\left[-\gamma^{\beta\eta}\frac{t^\beta}{L_0^\beta}\right] \tag{A.4}$$

A.1.2 Hazard Rate and CDF of IPL-Weibull

The hazard function of a random variable T with IPL-Weibull distribution can be obtained using the definition of hazard rate

$$h_T(t) = \lim_{dt \to 0} \frac{\Pr(t < T \leq t + dt | T > t)}{dt} \tag{A.5}$$

With this definition, the hazard function of IPL-Weibull distributed random variable is

$$h_{T(t)} = \frac{f_{T(t|\mathbf{X})}}{S_T(t)} = \beta \theta t^{\beta-1} \gamma^{\alpha} = \frac{\beta}{L_0^{\beta}} t^{\beta-1} \gamma^{\beta\eta} \tag{A.6}$$

The CDF or the failure probability can be easily obtained from the survival function. It is in the form of the following equation

$$F_T(t) = 1 - S_T(t) = 1 - \exp\left[-\gamma^{\beta\eta} \frac{t^{\beta}}{L_0^{\beta}}\right] \tag{A.7}$$

A.2 Likelihood Function of the IPL-Weibull Failure Model

The observed failure data of OLTCs or switched capacitors includes two sets: the first set is the set of observed failures and the second set includes devices that have survived until now and are in-service. Let $\{T_i | i = 1, 2, ..., r\}$ be a random sample of observed failures and $\{S_j | j = r+1, r+2, ..., N\}$ be the random set of transformers in-service. The random samples $\{T_i\}$ and $\{S_j\}$ are from T ; the unobserved failure time and the Bayes estimators of the model parameters are functions of the random samples. If $\mathbf{X} = [\beta, \theta, \alpha]^T$ is the vector of model parameters taking values in the parameter space $\Omega \subset \mathbb{R}^d$ then

$$\hat{\mathbf{X}} = h(T_i, S_j); i = 1, 2, ..., r, j = r+1, r+2, ..., N \tag{A.8}$$

$\hat{\mathbf{X}}$ is a random vector and after the random samples are taken i.e., the observations of failure and survival times take numeric values, $t_1, t_2, ..., t_r$ and $s_{r+1}, s_{r+2}, ..., s_N$ respectively, $\hat{\mathbf{X}}$ assumes a numeric value. Assuming censored data, the joint density of the random samples of failure times , survival times and the vector of model parameters is

$$p_{T_i,S_j}(t_1,t_2,..,t_r,s_{r+1},s_{r+2},...,s_N,\mathbf{X}) = p_{T_i,S_j}(t_1,t_2,...,t_r,s_{r+1},s_{r+2},...,s_N|\mathbf{X})p(\mathbf{X}) \tag{A.9}$$

In (A.9), $p(\mathbf{X})$ is the joint prior distribution of the model parameters. Since $T_i, i = 1,2,...,r$ and $S_j, j = r+1, r+2, ..., N$ are i.i.d sequences of random variables, the likelihood function that characterizes the failure model has the following form

$$\begin{aligned} p_{T_i,S_j}(t_1,t_2,...,t_r,s_{r+1},s_{r+2},...,s_N|\mathbf{X}) &= f_{T_1}(t_1|\mathbf{X})f_{T_2}(t_2|\mathbf{X})...f_{T_r}(t_r|\mathbf{X})f_{S_{r+1}}(s_{r+1}|\mathbf{X})...f_{S_N}(s_N|\mathbf{X}) \\ &= \prod_{i=1}^{r} f_{T_i}(t_i|\mathbf{X}) \prod_{j=r+1}^{N} [1 - F_{S_j}(s_j|\mathbf{X})] \end{aligned} \tag{A.10}$$

$1 - F_{S_j}(s_j|\mathbf{X})$ is the probability that the j^{th} unit will last at least s_j units of time. For the re-parameterized IPL-likelihood model the likelihood function can be analytically obtained by substituting the re-parameterized failure density, survival function and the observed failure data in (A.10).

$$\begin{aligned} \mathscr{L}(\beta,\theta,\alpha|t_1,t_2,...,t_r,s_{r+1},s_N,N,C,r) &= \prod_{i=1}^{r} \beta\theta t_{i:N}^{\beta-1}\gamma^{\alpha} \exp\left[-\gamma^{\alpha}\theta t_{i:N}^{\beta}\right] \prod_{j=r+1}^{N} \exp\left[-\gamma^{\alpha}\theta s_{j:N}^{\beta}\right] \\ &= \beta^r\theta^r \prod_{i=1}^{r} t_{i:N}^{\beta-1}\gamma^{\alpha} \exp\left[-\gamma^{\alpha}\theta \sum_{i=1}^{r} t_{i:N}^{\beta}\right] \exp\left[\gamma^{\alpha}\theta \sum_{j=r+1}^{N} s_{j:N}^{\beta}\right] \\ &= \beta^r\theta^r\gamma^{\alpha} \prod_{i=1}^{r} t_{i:N}^{\beta-1} \exp\left[-\gamma^{\alpha}\theta\left(\sum_{i=1}^{r} t_{i:N}^{\beta} + (N-r)s_{r:N}^{\beta}\right)\right] \end{aligned} \tag{A.11}$$

Thus we can write the likelihood function of the IPL-Weibull model as

$$\mathscr{L}(\beta,\theta,\alpha|t_1,t_2,...,t_r,s_{r+1},s_N,N,C,r) = \beta^r\theta^r\gamma^{\alpha} \prod_{i=1}^{r} t_{i:N}^{\beta-1} \exp[-\gamma^{\alpha}\theta P] \tag{A.12}$$

Where P is the re-scaled total test time in which the observed failure times $t_{i:N}$ and the survival times $s_{r:N}$ are re-scaled by raising to power β. The re-scaled total test time is given by

$$P := \sum_{i=1}^{r} t_{i:N}^{\beta} + (N-r)s_{r:N}^{\beta} \tag{A.13}$$

A.3 Joint Conditional Posterior and Marginal Posterior Distribution of Model Parameters

The joint conditional posterior density of the model parameters can be obtained from the joint density of the failure data and the model parameters. Applying the law of conditional probability to (A.9) we get for the joint conditional posterior of model parameters

$$\begin{aligned} g(\mathbf{X}|t_1,t_2,...,t_r,s_1,s_2,...,s_N,N,C,r &= \frac{p_{T_i,S_j}(t_1,t_2,..,t_r,s_{r+1},s_{r+2},...,s_N,\mathbf{X})}{p_{T_i,S_j}(t_1,t_2,..,t_r,s_{r+1},s_{r+2},...,s_N)} \\ &= \frac{\beta^r\theta^r\gamma^{\alpha}\prod_{i=1}^{r}t_{i:N}^{\beta-1}\exp\left[-\gamma^{\alpha}\theta P\right]p(\mathbf{X})}{p_{T_i,S_j}(t_1,t_2,..,t_r,s_{r+1},s_{r+2},...,s_N)} \end{aligned} \tag{A.14}$$

We assume gamma priors on the shape, the re-parameterized scale parameter, and a uniform prior in the closed interval $[A,B]$ on the stress coefficient parameter. Since the parameters are independent, we assume independent priors. We can write for the joint prior distribution $p(\mathbf{X})$

$$p(\mathbf{X}) = \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{\delta_1-1}\theta^{\delta_2-1}\exp[-(\lambda_1\beta+\lambda_2\theta)] \tag{A.15}$$

Where $\lambda_1 > 0, \delta_1 > 0$ are hyperparameters on β and $\lambda_2 > 0, \delta_2 > 0$ are hyperparameters on θ. The joint prior definition of (A.15) when combined with the (A.14) gives the following joint conditional posterior of $\mathbf{X} = [\beta,\theta,\alpha]^T$

$$\begin{aligned} g(\beta,\theta,\alpha|t_1,...,t_r,s_{r+1},...,s_N,C,N,r) \propto \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha} \\ \prod_{i=1}^{r}t_{i:N}^{\beta-1}\exp[-(\gamma^{\alpha}\theta P+\lambda_1\beta+\lambda_2\theta)] \end{aligned} \tag{A.16}$$

The marginal conditional posterior of model parameters can be obtained from the joint conditional posterior by integrating out all the parameters except the parameter of interest. For the shape parameter β the marginal conditional posterior is

$$f(\beta|t_1,t_2,..,t_r,s_{r+1},...,s_N,C,N,r) = \int_\theta \int_\alpha g(\beta,\theta,\alpha|t_1,...,t_r,s_{r+1},...,s_N,C,N,r)d\theta d\alpha$$
$$\propto \int_\theta \int_\alpha \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$
$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\theta d\alpha$$
$$= K_{\beta,r}\frac{1}{B-A}\frac{\lambda_1^{\delta}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\int_\theta \int_\alpha \beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$

$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\theta d\alpha \quad \text{(A.17)}$$

Where $K_{\beta,r}$ is the normalization constant and is equal to

$$K_{\beta,r}^{-1} = \frac{1}{B-A}\frac{\lambda_1^{\delta}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\int_\beta \int_\theta \int_\alpha \beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$
$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\beta d\theta d\alpha \quad \text{(A.18)}$$

Since the domain of θ is $(0,\infty)$ consider the definite integral

$$I_1 = \lim_{l\to\infty}\int_0^l \theta^{r+\delta_2-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\theta = -\exp(-\beta\lambda_1)[\Gamma(r+\delta_2, l\gamma^\alpha P+l\lambda_2)-$$
$$\Gamma(r+\delta_2,0)](\gamma^\alpha P+\lambda_2)^{-(r+\delta_2)} \quad \text{(A.19)}$$

Since the total re-scaled test time is much larger than the total number of failures, we have for $\gamma^{\alpha}P >> r$,

$$I = \exp(-\beta\lambda_1)(\gamma^{\alpha}P+\lambda_2)^{-(r+\delta_2)} \tag{A.20}$$

$$= \exp(-\beta\lambda_1)(\gamma^{\alpha}P)^{-(r+\delta_2)}\left[1+\frac{\lambda_2}{\gamma^{\alpha}P}\right]^{-(r+\delta_2)} \tag{A.21}$$

Since $\gamma^{\alpha}P >> \lambda_2$, and ignoring higher order terms, we have

$$I = \exp(-\beta\lambda_1)(\gamma^{\alpha}P)^{-(r+\delta_2)}\left[1-(r+\delta_2)\frac{\lambda_2}{\gamma^{\alpha}P}\right] \tag{A.22}$$

$$= \exp(-\beta\lambda_1)(\gamma^{\alpha}P)^{-(r+\delta_2+1)}[\gamma^{\alpha}P-(r+\delta_2)\lambda_2] \tag{A.23}$$

Substituting (A.23) in (A.18) gives the following for the normalization constant, $K_{\beta,r}$

$$K_{\beta,r}^{-1} = \frac{1}{B-A}\frac{\lambda_1^{\delta}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\int_{\beta}\beta^{r+\delta_1-1}\exp(-\beta\lambda_1) \prod_{i=1}^{r} t_{i:N}^{\beta-1} P^{-(r+\delta_2+1)}d\beta \int_{\alpha}\gamma^{-\alpha(\delta_2+1)}[\gamma^{\alpha}P-(r+\delta_2)\lambda_2]d\alpha \tag{A.24}$$

Equations (A.17) and (A.24) together give the marginal conditional posterior of the shape parameter β. The marginal conditional posterior of re-parameterized scale parameter θ can be obtained in a similar manner by integrating out β and α from the joint conditional posterior (A.16).The

marginal conditional posterior of θ is

$$f(\theta|t_1,t_2,..,t_r,s_{r+1},...,s_N,C,N,r) = \int_\beta \int_\alpha g(\beta,\theta,\alpha|t_1,...,t_r,s_{r+1},...,s_N,C,N,r)d\beta d\alpha$$
$$\propto \int_\beta \int_\alpha \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$
$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\beta d\alpha$$
$$= K_{\theta,r}\int_\beta \int_\alpha \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$
$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\beta d\alpha \quad \text{(A.25)}$$

Where the normalization constant $K_{\theta,r}^{-1} = K_{\beta,r}^{-1}$. In a similar manner we can obtain the marginal conditional of re-parameterized stress parameter α by integrating out β and θ from (A.16). This gives for the marginal conditional of α

$$f(\alpha|t_1,t_2,..,t_r,s_{r+1},...,s_N,C,N,r) = K_{\alpha,r}\int_\beta \int_\theta \frac{1}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$
$$\prod_{i=1}^{r} t_{i:N}^{\beta-1}\exp[-(\gamma^\alpha\theta P+\lambda_1\beta+\lambda_2\theta)]d\beta d\theta \quad \text{(A.26)}$$

A.4 Bayes Estimators of Model Paramters

In deriving the Bayes estimators of the model parameters, the discrepancy between the true value of the parameter vector $\mathbf{X}$ and the estimated value $\hat{\mathbf{X}}$ is measured by the loss function $\mathrm{L}(\mathbf{X},\hat{\mathbf{X}})$. We assume a quadratic loss function $\mathrm{L}(\mathbf{X},\hat{\mathbf{X}}) = (\mathbf{X}-\hat{\mathbf{X}})^2$ under which the point estimates that minimize the posterior risk of the estimator vector $\hat{\mathbf{X}}$ are conditional expectations of the individual estimators [85].

The conditional expectation or Bayes estimator of β can be obtained as

$$\hat{\beta} = \int_{\beta} \beta f(\beta|t_1,t_2,...,t_r,s_{r+1},...,s_N,N,C,r)d\beta = \int_{\beta} K_{\beta,r} \frac{1}{B-A} \frac{\lambda_1^{\delta} \lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)} \int_{\theta}\int_{\alpha} \beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$

$$\prod_{i=1}^{r} t_{i:N}^{\beta-1} \exp[-(\gamma^{\alpha}\theta P + \lambda_1\beta + \lambda_2\theta)]d\theta d\alpha d\beta$$

$$= K_{\beta,r} \frac{1}{B-A} \frac{\lambda_1^{\delta_1} \lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)} \int_{\beta}\int_{\theta}\int_{\alpha} \beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\gamma^{r\alpha}$$

$$\prod_{i=1}^{r} t_{i:N}^{\beta-1} \exp[-(\gamma^{\alpha}\theta P + \lambda_1\beta + \lambda_2\theta)]d\theta d\alpha d\beta \quad \text{(A.27)}$$

To get the Bayes estimator of β we substitute (A.21) and (A.24) in (A.27). This gives the following form of the Bayes estimator of β

$$\hat{\beta} = K_{\beta} \int_{\beta} \beta^{r+\delta_1} \exp(-\beta\lambda_1) \prod_{i=1}^{r} t_{i:N}^{\beta-1} P^{-(r+\delta_2+1)}$$

$$K_{\beta}^{-1} = \int_{\beta} \beta^{r+\delta_1-1} \exp(-\beta\lambda_1) \prod_{i=1}^{r} t_{i:N}^{\beta-1} P^{-(r+\delta_2+1)} \quad \text{(A.28)}$$

The Bayes estimator of the θ is

$$\hat{\theta} = \int_{\theta} \theta f(\theta|t_1,t_2,...,t_r,s_{r+1},...,s_N,N,C,r)d\theta = K_{\theta,r} \frac{1}{B-A} \frac{\lambda_1^{\delta_1} \lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)} \int_{\beta}\int_{\alpha} \beta^{r+\delta_1-1}\gamma^{r\alpha} \prod_{i=1}^{r} t_{i:N}^{\beta-1} d\beta d\alpha$$

$$\int_{\theta} \theta^{r+\delta_2} \exp[-(\gamma^{\alpha}\theta P + \lambda_1\beta + \lambda_2\theta)]d\theta \quad \text{(A.29)}$$

Consider the definite integral

$$I_2 = \lim_{l\to\infty} \int_0^l \theta^{r+\delta_2} \exp[-(\gamma^{\alpha}\theta P + \lambda_1\beta + \lambda_2\theta)]d\theta = \exp(-\beta\lambda_1)\gamma^{-\alpha(r+\delta_2+2)} P^{-(r+\delta_2+2)}$$

$$(\gamma^{\alpha} P - (r+\delta_2+1)\lambda_2) \quad \text{(A.30)}$$

Substituting the value of $K_{\theta,r}$ and (A.30) in (A.29) we have for the conditional expectation of θ

$$\hat{\theta} = K_\theta \frac{\int_\alpha \gamma^{-\alpha(\delta_2+2)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha}{\int_\alpha \gamma^{-\alpha(\delta_2+1)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha} \tag{A.31}$$

Where

$$K_\theta = \frac{\int_\beta \beta^{r+\delta_1-1}\exp(-\beta\lambda_1)\prod_{i=1}^r t_{i:N}^{\beta-1}P^{-(r+\delta_2+2)}d\beta}{\int_\beta \beta^{r+\delta_1-1}\exp(-\beta\lambda_1)\prod_{i=1}^r t_{i:N}^{\beta-1}P^{-(r+\delta_2+1)}d\beta} \tag{A.32}$$

The integrals in the numerator and denominator of (A.31) can be further simplified and thus we have for the Bayes estimator of θ

$$\hat{\theta} = K_\theta \frac{PI_{\hat{\theta},1} - (r+\delta_2)\lambda_2 I_{\hat{\theta},2}}{PI_{\hat{\theta},3} - (r+\delta_2)\lambda_2 I_{\hat{\theta},4}} \tag{A.33}$$

$$I_{\hat{\theta},4} = I_{\hat{\theta},2}; I_{\hat{\theta},1} = \frac{\gamma^{-A(\delta_2+1)} - \gamma^{-B(\delta_2+1)}}{(\delta_2+1)\ln\gamma}$$

$$I_{\hat{\theta},2} = \frac{\gamma^{-A(\delta_2+2)} - \gamma^{-B(\delta_2+2)}}{(\delta_2+2)\ln\gamma}, I_{\hat{\theta},3} = \frac{\gamma^{B\delta_2} - \gamma^{A\delta_2}}{\delta_2\gamma^{\delta_2(A+B)}\ln\gamma} \tag{A.34}$$

Similarly we can find the conditional expectation or the Bayes estimator of the re-parameterized stress coefficient by evaluating the integral

$$\hat{\alpha} = \int_\alpha f(\alpha|t_1,t_2,...,t_r,s_{r+1},...,s_N)d\alpha$$

$$= \frac{K_{\alpha,r}}{B-A}\frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)}\int_\beta\int_\theta\int_\alpha \beta^{r+\delta_1-1}\theta^{r+\delta_2-1}\prod_{i=1}^r t_{i:N}^{\beta-1}$$

$$\alpha\gamma^{r\alpha}\exp[-(\gamma^\alpha P\theta + \lambda_1\beta + \lambda_2\theta)]d\beta d\theta d\alpha \tag{A.35}$$

Substituting the value of $K_{\alpha,r}$ and the integral I_1 in (A.35), we have for the Bayes estimator of α

$$\hat{\alpha} = \frac{\int_\alpha \alpha\gamma^{-\alpha(\delta_2+1)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha}{\int_\alpha \gamma^{-\alpha(\delta_2+1)}(\gamma^\alpha P - (r+\delta_2)\lambda_2)d\alpha} \tag{A.36}$$

The integrals in the numerator and denominator of (A.36) have closed-form solutions and this gives the final form of the Bayes estimator of α

$$\hat{\alpha} = \frac{PI_{\hat{\alpha}_1,r} - (r+\delta_2)\lambda_2 I_{\hat{\alpha}_2,r}}{PI_{\hat{\alpha}_3,r} - (r+\delta_2)\lambda_2 I_{\hat{\alpha}_4,r}} \tag{A.37}$$

Where

$$\begin{aligned} I_{\hat{\alpha}_1,r} &= \frac{G_1 - G_2}{\delta_2^2 \ln^2 \gamma}; I_{\hat{\alpha}_2,r} = \frac{G_3 - G_4}{(\delta_2+1)^2 \ln^2 \gamma} \\ I_{\hat{\alpha}_3,r} &= \frac{\gamma^{B\delta_2} - \gamma^{A\delta_2}}{\delta_2 \gamma^{(A+B)\delta_2} \ln \gamma}; I_{\hat{\alpha}_4,r} = \frac{\gamma^{-A(\delta_2+1)} - \gamma^{-B(\delta_2+1)}}{(\delta_2+1)\ln \gamma} \end{aligned} \tag{A.38}$$

G_1, G_2, G_3 and G_4 in (A.38) are given by

$$\begin{aligned} G_1 &= \exp(-A\delta_2 \ln \gamma)(A\delta_2 \ln \gamma + 1) \\ G_2 &= \exp(-B\delta_2 \ln \gamma)(B\delta_2 \ln \gamma + 1) \\ G_3 &= \exp(-A\delta_2 \ln \gamma - A \ln \gamma)[1 + \ln \gamma(A(\delta_2+1))] \\ G_4 &= \exp(-B\delta_2 \ln \gamma - B \ln \gamma)[1 + \ln \gamma(B(\delta_2+1))] \end{aligned} \tag{A.39}$$

APPENDIX B

METROPOLIS HASTING ALGORITHM-MCMC SIMULATION

This appendix describes the Markov Chain Monte Carlo-based Metropolis-Hastings algorithm used to obtain the full posterior of the model parameters. The Metropolis-Hastings algorithm enables sampling from a target posterior distribution by performing a random walk on a Markov Chain in a manner that the target distribution is the stationary distribution of the simulated Markov Chain. Given a target distribution π^*, the Metropolis-Hastings algorithm constructs a transition kernel that converges to π^* in the limit.

B.1 Irreducible and Aperiodic Markov Chain

Consider a stochastic process $\{\mathbf{X}(n); n \geq 0\}$ as a family of Ω-valued random vectors defined on the same probability space $(\mathbb{R}^d, \mathbb{B}, \Phi)$. $\mathbb{B}$ is the Borel σ-field on $\mathbb{R}^d$ and Φ is the probability measure defined on the subsets of $\mathbb{B}$.

The stochastic process $\{\mathbf{X}(n); n \geq 0\}$ is a Markov chain in discrete time taking values in finite continuous state space, $\Omega \subset \mathbb{R}^d$ if it satisfies

$$\Pr(\mathbf{X}_{n+1} \in A | \mathbf{X}_n = \mathbf{I}_n, \mathbf{X}_{n-1} = \mathbf{I}_{n-1}, ..., \mathbf{X}_0 = \mathbf{I}_0) = \Pr(\mathbf{X}_n \in A | \mathbf{X}_n = \mathbf{I}_n) \quad \text{(B.1)}$$

$$\forall n \geq 0, \mathbf{I}_n, \mathbf{I}_{n-1}, ..., \mathbf{I}_0 \in \mathbb{R}^d, \forall A \in \mathbb{B}$$

For the d-dimensional Markov process defined in (B.1), we consider a transition kernel $P(\mathbf{I}, A)$ where $\mathbf{I} \in \mathbb{R}^d$ and $A \subset \mathbb{R}^d$. $P(\mathbf{I}, A)$ is the probability that the d-dimensional Markov process will arrive in in the set $A \subset \mathbb{R}$ given that the process is in state vector $\mathbf{I} \in \mathbb{R}^d$.

$$P^n(\mathbf{I}, A) = \Pr(\mathbf{X}_n \in A | \mathbf{X}_0 = \mathbf{I}) \forall n \in \mathbb{N} \quad \text{(B.2)}$$

To ensure the irreducibility and aperiodiciy of the simulated Markov Chain the transition kernel

must converge to a stationary distribution. This implies that the transition kernel must be irreducible and aperiodic. To accomplish this consider two extremely small and disjoint subsets $\Delta_{\mathbf{I}_n,\Phi}$ and $\Delta_{\mathbf{I}^*,\Phi}$ of Ω that contain the state current state vector $\mathbf{I}_n \in \Delta_{\mathbf{I}_n,\Phi}$ and the proposed state vector $\mathbf{I}^* \in \Delta_{\mathbf{I}^*,\Phi}$ with $\Phi(\Delta_{\mathbf{I}_n,\Phi}) > 0$ and $\Phi(\Delta_{\mathbf{I}^*,\Phi}) > 0$. The transition kernel $P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}^*,\Phi} | \mathbf{X}_n \in \Delta_{\mathbf{I}_n,\Phi})$ will be Φ-irreducible and aperiodic if it has the form

$$P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}^*,\Phi} | \mathbf{X}_n \in \Delta_{\mathbf{I}_n,\Phi}) = a(\mathbf{I}_n, \mathbf{I}^*) q(\mathbf{I}^* | \mathbf{I}_n) \tag{B.3}$$

Where $a(\mathbf{I}_n, \mathbf{I}^*)$ is the acceptance probability of the candidate state vector $\mathbf{I}^* \in \Delta_{\mathbf{I}^*,\Phi}$ and $q(\mathbf{I}^* | \mathbf{I}_n)$ is the multivariate proposal distribution of the candidate state vector $\mathbf{I}^* \in \Delta_{\mathbf{I}^*,\Phi}$ conditioned on the current state vector $\mathbf{I}_n \in \Delta_{\mathbf{I}_n,\Phi}$. To preserve the ergodicity of the simulated Markov process the Metropolis Hastings algorithm defines the acceptance probability of the candidate state vector, $\mathbf{I}^* \in \Delta_{\mathbf{I}^*,\Phi}$ as

$$a(\mathbf{I}_n, \mathbf{I}^*) = \min \left\{ 1, \frac{\pi^*(\mathbf{I}^*) q(\mathbf{I}_n | \mathbf{I}^*)}{\pi^*(\mathbf{I}_n) q(\mathbf{I}^* | \mathbf{I}_n)} \right\} \tag{B.4}$$

Assuming that the target posterior is strictly positive in the entire state space, it follows that $a(\mathbf{I}_n, \mathbf{I}^*) > 0$ and $q(\mathbf{I}^* | \mathbf{I}_n) = \Pr(\mathbf{I}^* \in \Delta_{\mathbf{I}^*,\Phi} | \mathbf{I}_n \in \Delta_{\mathbf{I}_n,\Phi}) > 0$ and hence

$$P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}^*,\Phi} | \mathbf{X}_n \in \Delta_{\mathbf{I}_n,\Phi}) > 0 \tag{B.5}$$

Thus a Markov Process with the transition kernel given by (B.3) is irreducible and aperiodic and hence converges to a unique invariant distribution.

B.2 Target Posterior as Invariant Distribution

A probability distribution π^* on the finite continuous state space Ω satisfies detailed balance or reversibility conditions with respect to the transition kernel in (B.3) if for any disjoint $\Delta_{\mathbf{I}_n,\Phi} \subset \Omega$ and $\Delta_{I^*,\Phi} \subset \Omega$

$$\pi^*(\Delta_{\mathbf{I}_n,\Phi}) P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}^*,\Phi} | \mathbf{X}_n \in \Delta_{\mathbf{I}_n,\Phi}) = \pi^*(\Delta_{I^*,\Phi}) P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}_n,\Phi} | \mathbf{X}_n \in \Delta_{\mathbf{I}^*,\Phi}) \tag{B.6}$$

If a target posterior π^* satisfies (B.6), then π^* is the invariant or stationary distribution defined on Ω with respect to the transition kernel of (B.3). With the transition kernel defined as in (B.3) and the acceptance probability in (B.4) it can be shown that the target distribution π^* satisfies the detailed balance conditions and thus is a sample from the invariant distribution. To see this suppose $a(\mathbf{I}_n, \mathbf{I}^*) = 1$. This implies that

$$a(\mathbf{I}^*, \mathbf{I}_n) = \frac{\pi^*(\Delta_{I_n,\Phi})q(\mathbf{I}^*|\mathbf{I}_n)}{\pi^*(\Delta_{I^*,\Phi})q(\mathbf{I}_n|\mathbf{I}^*)} \tag{B.7}$$

With $a(\mathbf{I}_n, \mathbf{I}^*) = 1$ the left hand side of (B.6) becomes

$$\pi^*(\Delta_{\mathbf{I}_n,\Phi})P(\mathbf{X}_{n+1} \in \Delta_{\mathbf{I}^*,\Phi}|\mathbf{X}_n \in \Delta_{\mathbf{I}_n,\Phi}) = \pi^*(\Delta_{\mathbf{I}_n,\Phi})q(\mathbf{I}^*|\mathbf{I}_n) \tag{B.8}$$

The right hand side of (B.6) thus equates to

$$\pi^*(\Delta_{I^*,\Phi})a(\mathbf{I}^*, \mathbf{I}_n)q(\mathbf{I}_n|\mathbf{I}^*) = \pi^*(\Delta_{I^*,\Phi})\frac{\pi^*(\Delta_{I_n,\Phi})q(\mathbf{I}^*|\mathbf{I}_n)}{\pi^*(\Delta_{I^*,\Phi})q(\mathbf{I}_n|\mathbf{I}^*)}q(\mathbf{I}_n|I^*) = \pi^*(\Delta_{I_n,\Phi})q(\mathbf{I}^*|\mathbf{I}_n) \tag{B.9}$$

Hence the detailed balance or the reversibility condition is satisfied. Similarly it can be easily shown that the detailed balance is satisfied in the case $a(\mathbf{I}_n, \mathbf{I}^*) = \frac{\pi^*(\mathbf{I}^*)q(\mathbf{I}_n|\mathbf{I}^*)}{\pi^*(\mathbf{I}_n)q(\mathbf{I}^*|\mathbf{I}_n)}$

B.3 Ergodic Theorem

he ergodic theorem dictates the long term probabilistic behavior of an irreducible Markov process. An ergodic Markov process converges to a unique invariant distribution regardless of the initial distribution. If the total number of visits made by an irreducible Markov process to a state vector $\mathbf{I}_n \in \Delta_{\mathbf{I}_n,\Phi}$ be $N_{\Delta_{\mathbf{I}_n,\Phi}}(n)$ then by the Ergodic theorem

$$\Pr\left(\lim_{n\to\infty} \frac{N_{\Delta_{\mathbf{I}_n,\Phi}}(n)}{n} = \pi^*(\Delta_{\mathbf{I}_n,\Phi})\right) = 1 \tag{B.10}$$

Equation (B.10) says that if an ergodic Markov process is run for a long time, the steady-state probability of being in state vector $\mathbf{I}_n \in \Delta_{I_n,\Phi}$ is the same as the fraction of the time the process

spends in state vector $\mathbf{I}_n \in \Delta_{I_n,\Phi}$. Since the Markov process is never explicitly constructed, the ergodic theorem allows us to sample from the target posterior by walking the Markov process and recording the states.