

University of Windsor

## Scholarship at UWindor

---

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

---

3-10-2021

# Stochastic Modeling and Planning of Wind-Based Distributed Generators in Distribution System

Mohammad Borooshan  
*University of Windsor*

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

---

### Recommended Citation

Borooshan, Mohammad, "Stochastic Modeling and Planning of Wind-Based Distributed Generators in Distribution System" (2021). *Electronic Theses and Dissertations*. 8549.  
<https://scholar.uwindsor.ca/etd/8549>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

# Stochastic Modeling and Planning of Wind-Based Distributed Generators in Distribution System

By

Mohammad Borooshan

A Thesis

Submitted to the Faculty of Graduate Studies  
through the Industrial Engineering Graduate Program  
in Partial Fulfillment of the Requirements for  
the Degree of Master of Applied Science  
at the University of Windsor

Windsor, Ontario, Canada

2021

© 2021 Mohammad Borooshan

# Stochastic Modeling and Planning of Wind-Based Distributed Generators in Distribution System

by

Mohammad Borooshan

APPROVED BY:

---

R. Lashkari

Department of Mechanical, Automotive and Materials Engineering

---

S. Saad Ahmed

School of Computer Science

---

A. Azab, Co-Advisor

Department of Mechanical, Automotive and Materials Engineering

---

M. A. Azzouz, Co-Advisor

Department of Electrical and Computer Engineering

December 22, 2020

# DECLARATION OF CO-AUTHORSHIP AND PREVIOUS PUBLICATION

## i Co-Authorship

I hereby declare that this thesis is accomplished based on the result of joint research. Dr. Ahmed Azab and Dr. Maher Azzouz contributed to the research's overall organization, particularly in terms of optimization and power system, respectively. From Natural Resources Canada, Dr. Ahmed Awad contributed with the foundational guidance on which this thesis was built. Mahzan Dalawir contributed to the completion of the group paper including the power systems sections. Abdelrahman Amer contributed with a technical presentation [Abstract] in Windsor. In other cases, the main contributions, such as data analysis and programming, were accomplished by the author.

I am aware of the University of Windsor Senate Policy on Authorship, and I certify that I have correctly acknowledged the contribution of other researchers to my thesis and have obtained written permission from each of the co-authors to include the above materials in my thesis. I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my work.

## ii Previous Publication

A journal paper has been written using the research findings of this thesis and will be submitted soon. Moreover, an international conference paper was submitted in 2019.

Thesis chapters	Publication title/full citation	Publication status
3, 4	Borooshan Mohammad, Dalawir Mahzan, Azab Ahmed, Azzouz Maher, Awad Ahmed, “Optimal Allocation of Wind-Based Distributed Generators Using Stochastic program with Mixture of Probability Density Function,”	To be submitted to a journal
3, 4	Borooshan Mohammad, Abdelrahman Amer, Azab Ahmed, Awad Ahmed, “Optimal Location and Sizing of Wind-based Distributed Generators: Goodness of Fit for Electric Vehicle Charging and Wind Speed Profiles,”	Technical Presentation [Abstract] in Windsor

I certify that I have obtained written permission from the copyright owners to include the above-published materials in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

### iii General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the Canada Copyright Act's meaning, I certify that I have obtained written permission from the copyright owners to include such materials in my thesis.

I declare that this is a real copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office and that this thesis has not been submitted for a higher degree to any other University or Institution.

## ABSTRACT

The increasing strain on the Earth resulting from pollution, climate change, and finite resources has established the development of renewable energy sourcing methods, such as wind, solar and geothermal energy. By reorganizing the power system structures, and the growth in customer demand, the development of Distributed Generation (DG) play a vital role in the power system planning. Furthermore, because of the inexhaustibility and cleanliness of the renewable DG units, they are inevitably the key to a sustainable energy supply infrastructure. Nevertheless, the random nature associated with the renewable DG units produces specific challenges that have to be addressed to accelerate the expansion of the renewable DG units in the distribution system.

Firstly, a new method for the determination of the wind speed distribution based on hourly wind speed data is proposed. Thus, instead of using only the well-known unimodal distributions such as Weibull and Rayleigh, a combination of probability density functions (PDFs) is taken into account, considering four sets of parameters in which each set represents a distribution. Furthermore, this model enhances the likelihood of the estimated wind speed probabilities. The maximum likelihood estimation (MLE) method for finite mixture models through the expectation-maximization (EM) algorithm is used to estimate the optimal parameters of the mixture distribution. Then two types of error measurements assessed the performance of each unimodal and multimodal distribution. As a result, the mixture of Gamma (MoG) distribution returned the most accurate results.

Secondly, the results of wind speed modeling will be used in the siting and sizing wind-based DG units. The methodology addresses a probabilistic generation load model that combines all possible operating conditions of the wind-based DG units and load levels with their probabilities. The objective of siting and sizing formulation is to minimize the annual energy losses of the system as well as keeping the system constraints such as voltage limits at different buses (slack and load buses) of the system, feeder capacity, discrete size of the DG units, maximum investment on each bus, and maximum penetration limit of DG units in an acceptable limit.

## DEDICATION

This thesis is dedicated to my parents, my siblings, and my brother in law.

# TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP AND PREVIOUS PUBLICATION.....	iii
ABSTRACT.....	v
DEDICATION.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xii
ABBREVIATION.....	xiv
NOMENCLATURE.....	xvi
CHAPTER 1: INTRODUCTION.....	1
1.1 Preface.....	1
1.2 Thesis objectives.....	2
1.3 Outline.....	2
CHAPTER 2: LITERATURE SURVEY.....	3
2.1 Introduction.....	3
2.2 Siting and Sizing of Distributed Generators.....	3
2.3 Stochastic Modeling for Load and Generation.....	10
2.4 Application of Genetic Algorithm in Planning.....	13
2.5 Concluding Remarks.....	14
CHAPTER 3: MODELING OF WIND-BASED RENEWABLE RESOURCES..	15
3.1 Introduction.....	15
3.2 Factors Impacting Wind Speed.....	17
3.2.1 Impact of year on wind speed.....	18
3.2.2 Impact of season on wind speed.....	18
3.2.3 Impact of weather on wind speed.....	20
3.2.4 Impact of temperature on wind speed.....	20
3.2.5 Impact of time on wind speed.....	21



3.3 Wind Speed Stochastic Modeling .....	22
3.3.1 Unimodal probability density functions in wind modeling .....	23
3.3.2 Annual wind speed modeling for unimodal distributions .....	25
3.3.3 Goodness-of-fit Kolmogorov-Smirnov test.....	26
3.3.4 Average model performance .....	27
3.3.5 Multimodal distribution (mixture distribution) .....	28
3.3.6 Proper multimodal probability density function (PDF) in wind modeling .....	33
3.4 Wind Turbine Output Power & Modeling of Renewable Resources and Load Data .....	34
3.5 Load Modeling .....	37
3.5.1 Combined generation load model.....	37
3.6 Computational Results .....	39
3.6.1 Results of unimodality.....	39
3.6.2 Results of multimodality .....	43
3.6.3 Best candidate distributions.....	49
3.7 Combination of the Wind and Load State Probabilities.....	53
3.8 Discussion .....	59
CHAPTER 4: AN APPLICATION OF THE PROPOSED MIXTURE MODEL IN THE PLANNING OF DISTRIBUTION SYSTEMS .....	60
4.1 Introduction .....	60
4.2 Planning Problem Objective .....	61
4.3 Site Matching .....	62
4.4 Planning Problem Formulation .....	63
4.4.1 Objective function .....	64
4.4.2 Constraints .....	64
4.5 Case Study.....	66
4.5.1 System under study.....	66
4.6 Load Data .....	67
4.7 Wind Speed and Wind Turbine Data .....	68
4.8 Optimal Allocation of Wind-based DGs Using Genetic Algorithm (GA)....	70

4.9 Computational Results .....	73
4.10 Discussion .....	74
CHAPTER 5: CONCLUSION AND FUTURE WORK.....	75
5.1 Conclusion.....	75
5.2 Future Work .....	75
APPENDIX A: DISTRIBUTION SYSTEM DATA.....	77
BIBLIOGRAPHY.....	78
VITA AUCTORIS.....	84

## LIST OF TABLES

Table 3.1: Parameters of the unimodal distributions .....	26
Table 3.2: Selected wind speed states .....	36
Table 3.3: Load model .....	37
Table 3.4: P-value of the Unimodal pdf.....	39
Table 3.5: Root mean square error of Unimodal pdfs .....	39
Table 3.6: Mean absolute error of Unimodal pdfs .....	41
Table 3.7: BIC & Log-likelihood of mixture distributions with different # of parameters .....	43
Table 3.8: BIC & Log-likelihood of mixture distributions with different # of parameters .....	44
Table 3.9: Parameters and weights of MoW distributions.....	44
Table 3.10: Parameters and weights of MoN distributions.....	44
Table 3.11: Parameters and weights of MoG distributions.....	45
Table 3.12: P-value of the Multimodal pdfs .....	45
Table 3.13: Root mean square error of Multimodal pdfs.....	45
Table 3.14: Mean absolute error of Multimodal pdfs .....	47
Table 3.15: RMSE of two candidate distributions.....	49
Table 3.16: MAE of two candidate distributions.....	51
Table 3.17: The wind state probabilities of Johnson SB distribution .....	54
Table 3.18: The wind state probabilities of Weibull distribution .....	54
Table 3.19: The wind state probabilities of MoW distribution.....	55
Table 3.20: The wind state probabilities of MoN distribution.....	55
Table 3.21: The wind state probabilities of MoG distribution.....	56
Table 3.22: The combination of wind and load state probabilities of Johnson SB distribution .....	57
Table 3.23: The combination of wind and load state probabilities of Weibull distribution .....	57
Table 3.24: The combination of wind and load state probabilities of MoW distribution .....	58

Table 3.25: The combination of wind and load state probabilities of MoN distribution .....	58
Table 3.26: The combination of wind and load state probabilities of MoG distribution .....	59
Table 4.1: Probabilistic load model .....	68
Table 4.2: Wind speed probabilities .....	69
Table 4.3: Wind power probabilities .....	70
Table 4.4: Losses and installed wind DGs for each scenario .....	74

## LIST OF FIGURES

Figure 3.1: Canada’s installed wind power capacity .....	16
Figure 3.2: Ontario installed capacity .....	16
Figure 3.3: Average wind speed of years 2016 and 2017.....	18
Figure 3.4: Average wind speed of different seasons .....	19
Figure 3.5: Average wind speed of different months .....	19
Figure 3.6: Average wind speed of different weather's type .....	20
Figure 3.7: Wind speed at different temperatures.....	21
Figure 3.8: Wind speed in different time segments .....	22
Figure 3.9: The flowchart of the EM algorithm.....	32
Figure 3.10: Wind turbine power curve.....	35
Figure 3.11: RMSE of Unimodal pdfs in Windsor.....	40
Figure 3.12: RMSE of Unimodal pdfs in Hamilton.....	40
Figure 3.13: MAE of Unimodal pdfs in Windsor.....	41
Figure 3.14: MAE of Unimodal pdfs in Hamilton.....	42
Figure 3.15: Pdfs of unimodal distributions in Windsor.....	42
Figure 3.16: Pdfs of unimodal distributions in Hamilton .....	43
Figure 3.17: RMSE of Multimodal pdfs in Windsor .....	46
Figure 3.18: RMSE of Multimodal pdfs in Hamilton.....	46
Figure 3.19: MAE of Multimodal pdfs in Windsor .....	47
Figure 3.20: MAE of Multimodal pdfs in Hamilton.....	48
Figure 3.21: Pdfs of Multimodal distributions in Windsor.....	48
Figure 3.22: Pdfs of Multimodal distributions in Hamilton .....	49
Figure 3.23: RMSE of candidate pdfs in Windsor.....	50
Figure 3.24: RMSE of candidate pdfs in Hamilton .....	50
Figure 3.25: MAE of candidate pdfs in Windsor.....	51
Figure 3.26: MAE of candidate pdfs in Hamilton .....	52
Figure 3.27: Pdfs of candidate distributions in Windsor .....	52
Figure 3.28: Pdfs of candidate distributions in Hamilton.....	53

Figure 4.1: System understudy .....	67
Figure 4.2: The chromosome of the eight decision variables .....	72
Figure 4.3: Adopted hybridize solution approach.....	73

## ABBREVIATION

DGs	Distributed Generators
MINLP	Mixed Integer nonlinear programming
PSO	Particle swarm optimization
PGS	Particle generating set
WTGU	Wind turbine generation unit
PV	Photovoltaic
SA	Simulated annealing
RSM	Response surface methodology
MDEP	Multistage distribution expansion planning
MILP	Mixed integer linear programming
CCP	Chance constrained programming
OPF	Optimal power flow
GA	Genetic algorithm
BESS	Battery energy storage system
FGP	Fuzzy goal programming
LSF	Loss sensitivity factor
ALOA	Ant lion optimization algorithm
GMM	Gaussian mixture modeling
MAE	Mean absolute error
RMSE	Root mean square error
MTTND	Mixture of two truncated normal distribution

PDF	Probability density function
EM	Expectation Maximization
KDE	Kernel density estimation
BMA	Bayesian model averaging
MCMC	Markov chain monte carlo
DKDM	Diffusion based kernel density model
MLE	Maximum likelihood estimation
K-S	Kolmogorov Smirnov
MAPE	Mean absolute percentage error
MBE	Mean bias error
ACO	Ant colony algorithm
FCSs	Fast charging stations
MoN	Mixture of normal distribution
MoW	Mixture of weibull distribution
MoG	Mixture of Gamma distribution
NR	Newton-Raphson
BIC	bayesian information criterion
WTG	Wind turbine generator
LDCs	local distribution companies
SOP	Standard Offer Program
CF	Capacity factor of wind turbine



## NOMENCLATURE

$k$	Shape parameter of the weibull distribution
$\lambda$	Scale parameter of the weibull distribution
$k$	Shape index
$h$	Scale index of rayleigh distribution
$\alpha$	Shape parameter of the gamma distribution
$\theta$	Scale parameter of the gamma distribution
$\beta$	Inverse scale parameter of the gamma distribution
$\xi$	Location parameter of the Johnson sb distribution
$\gamma$	Shape parameter of the johnson sb distribution
$\lambda_o$	Range parameter of the johnson sb distribution
$\delta$	Johnson sb parameter
$F_0(x)$	cumulative distribution function (CDF)
$F_{data}(x)$	empirical distribution function (EDF)
$n$	number of bins of the specific data set
$x_i$	number of observations
$\hat{x}_i$	probability of the wind speed
$\pi_j$	weight of the $j$ th component of mixture distributions
$\theta_j$	parameter of the $j$ th component of mixture distributions
$\Phi$	a vector of all the parameters of the mixture model
$n_i$	number of observations that fall in the $i$ th bin

$r$	the total number of bins
$z$	a vector of zeros
$z^{(p)}$	expected value of $z$ given the estimated value of $\Phi$
$\hat{\Phi}_0$	Null hypothesis
$\hat{\Phi}_1$	Alternative hypothesis
$B$	Number of iterations in bootstrap likelihood ratio test
$w_0$	likelihood ratio test statistic
$M$	the number of components
$v_{w1}$ and $v_{w2}$	the speed limits of state $w$
$v_{ci}$	cut-in speed of wind turbine
$v_r$	rated speed of wind turbine
$v_{co}$	cut-off speed of wind turbine
$P_{ave}$	Average wind power
$P_{rated}$	Rated wind power
$P(L_y)$	Load level probabilities
$P(C_g)$	Combination of wind and load state probabilities
$P_w(G)$	Wind state probabilities
$R$	the entire annual generation load model of $m$ turbines
$C$	combinations of the wind output power states
$N$	The total number of states in model $R$
Comp	Component of the mixture model

### Sets & indices

$n_i$	Set of buses
$c_i$	Set of candidate buses selected for DG placement
$N_g$	Set of all states $g$
$L_{i,j}$	Set of lines between buses $i$ and $j$
$i \& j$	Index of buses
$g$	Index of combined wind and load states
$m$	Set of all available turbines in the market, where each turbine has its own power performance curve
$m + 1$	Represents the different load level

### Variables

$x_i$	Integer variable denoting the number of wind turbines in bus $i$
$P_{lossg}$	Real power losses of the system given state $g$
$U_{i,g}$	Voltages [p.u.] at bus $i$ and state $g$
$P_{D_i}$	Active power demand at bus $i$
$P_{DG_{t,i}}$	the rated power of the $t_{th}$ wind-based DG connected at bus $i$
$V_{g,i}$	Voltage at bus $i$ during state $g$
$V_{g,j}$	Voltage at bus $j$ during state $g$
$\theta_{i,j}$	Admittance angle
$\delta_{g,j}$	Voltage angle at bus $j$ in state $g$

$\delta_{g,i}$	Voltage angle at bus $i$ in state $g$
$G_{ij}$	Conductance in line between bus $i$ & $j$
$I_{g,ij}$	Current at line between bus $i$ and $j$ in state $g$
$a_{t,i}$	The integer variable

### **Parameters of planning problem**

$Prob_g$	The probability of state $g$
$M$	Big $M$ notation
$P_r$	One-column matrix of length $m$ includes all ratings of the available wind turbines
$B$	The set of candidate buses to connect DGs
$k$	The fraction of DG penetration in the system defined by utility
$P_{G1}$	Active power generated by substation
$Q_{G1}$	Reactive power generated by substation
$Q_{Di}$	Reactive power demand at bus $i$
$P_{Di}$	Active power demand at bus $i$
$Y_{i,j}$	Admittances of buses $i$ and $j$
$V_{g,1}$	Slack bus voltage
$\delta_{g,1}$	Slack bus voltage angle
$P_{bus}$	The maximum allowable megawatt (MW) at each bus

# CHAPTER 1: INTRODUCTION

## 1.1 Preface

Because of the remarkable increase in customer demand in the past decade, the idea of developing small scale energy resources spread over the grid brought much attention. These small energy resources are known as distributed generators (DGs). Distributed generators refer to different technologies that deliver electricity at or near the demand. Distributed generators are capable of providing energy for a residential area as well as a microgrid, in which a smaller grid that can connect to the broader electricity delivery system like industries. By developing distributed generators, the transport of clean energy, reliable power, and decrease energy losses through transmission lines will be improved. The concept behind DGs is not newly discovered. However, based on the economic literature associated with the electricity market, it is a relatively new idea [1].

Among different types of renewable distributed generators, wind power is more ecologically friendly versus other energy resources. Wind power will be a strong potential if it delivers consumers energy at a reasonable price without declining the system's reliability and safety. In order to take advantage of wind power and its potentials, particularly technical and economic challenges need to be addressed. Among all types of DG technologies, the power generated by wind-based distributed generators is the most liable. Distribution system planning is one of the most critical obstacles that the system planners encountered, mainly when wind-based DG units are situated in the system. Since a few years ago, because of the advance in technologies and changes in climate and economics, attention has increased in using distributed generators.

Besides, as clean energy sources wind and solar are becoming an ideal method for generating power, the electricity infrastructure is adapting their technologies and strategies towards using these clean and unlimited resources. Nevertheless, both the radial and network systems operate without any generation in the distribution system or at customer premises. The primary reason is that, for many years, the infrastructure has been considered using the centralized generation despite decentralized generation. Thus, because of the power system's existence, the distributed generators must situate themselves in the current system to connect to the grid. The development of distributed generators can impact the

flow of power, system reliability, and voltage at customers and utility equipment. All these changes in traditional planning increase the gaps and complexity of the problem that needs to be solved. Therefore, most of the strategies taken into action are no longer valuable. The randomness and uncertainty related to the renewable energy resources brought more difficulties that have to be met to utilize the distributed generators in the shortest period.

## 1.2 Thesis objectives

The target of this thesis is to provide a planning framework that maximizes the penetration level of wind-based DGs and minimizes the system losses. The methodology of this thesis is based on generating new stochastic modeling, combining all possible operating conditions of the wind-based distributed generators and load levels with their probabilities. Subsequently, fit this model on the stochastic planning problem. Regarding the siting and sizing of the renewable DG, the primary stress of the planning problem is to optimally site and size the wind-based distributed generators in the distribution system to minimize the annual energy losses.

## 1.3 Outline

This research, classified as follows: Chapter 2, shows the literature survey of the past work done in stochastic modeling, and siting and sizing of distributed generators in the distribution system. The performance of different wind modeling versus a new mixture model will be considered in Chapter 3. Chapter 4 represents the optimal siting and sizing of wind-based DG units in the distribution system. The conclusion and a brief explanation of the future work will be covered in Chapter 5.

## CHAPTER 2: LITERATURE SURVEY

### 2.1 Introduction

The government's promises to provide clean energy have led to a growth in the investments toward more use of renewable resources to produce clean energy with less pollution and environmental impacts. The growth of the installation of DG units in the power system has economic, environmental, social, and technical impacts. Thus, this increase in DGs installation has encouraged researchers to become interested in optimal siting and sizing of DG units at the planning stage for minimizing the impacts of DGs installation and increasing the benefits of these renewable technologies. Chapter 2 will address the techniques used in literature for optimally site and size the renewable DGs and stochastic modeling of renewable energies such as wind and solar.

### 2.2 Siting and Sizing of Distributed Generators

In [2], the authors developed the mathematical model to minimize the system's power losses, considering both equality and inequality constraints such as power balance, voltage deviation limit, and reactive power compensation. The IEEE 34-bus and 85-bus radial distribution system with all possible load changes have been picked as a case study. Power balance, voltage deviation limit, and reactive power compensation have been taken into consideration as the constraints of the proposed model. In order to estimate the power loss and voltage at each branch, the load flow analysis has been used. The capacitor bank's optimal location has been categorized based on the loss sensitivity factor and voltage stability index. After placement of the capacitor, by using a curve fitting technique, the active and reactive power loss has been obtained. Finally, the simulated results have been estimated with the MINLP, PGS, PSO, HS-based methods.

The authors in [3] brought their attention to 12-bus, 15-bus, 33-bus, and 69-bus systems in order to generate different real and reactive power modes for wind and solar-based DGs. The objective function of power loss reduction and voltage stability enhancement of radial distribution systems have been considered to be optimized. Power flow and voltage limit are selected as the constraints of this work. A multi-objective particle swarm optimization (PSO) based wind turbine generation unit (WTGU) and photovoltaic

(PV) array placement approach has been taken into account to minimize power losses and more accurate voltage profile. Additionally, a voltage stability factor has been developed for increasing voltage stability.

The authors in [4] represent a new algorithm known as simulated annealing (SA) to optimize the size of a mixed PV/wind energy system considering battery storage. In order to solve the optimization problem, a heuristic approach is applied, which uses a stochastic gradient search for the exact solution. Furthermore, the objective function decided to be the minimization of the hybrid energy system's total cost. Moreover, PV size, wind turbine rotor swept area, and battery capacity are the decision variables, respectively. As a consequence, the SA algorithm returned a more satisfactory result than the response surface methodology (RSM).

The inherent uncertainty of wind speed has a huge obstacle to the successful execution of wind-based power generation technology. [5] defined a methodology by incorporating wind speed uncertainty in sizing the wind-battery system and incorporating it by considering the chance constraint programming approach. A deterministic equivalent energy balance equation has been considered in order to preannounce the reliability demand. Besides, the energy balance equation is more likely managed by chance constraint in which allows time series of the entire system. They used a sequential Monte-Carlo simulation of the system to validate the system's reliability.

It is inevitable that energy is the most crucial fundamental for all media, industries, and services across the globe. Conventional distributed generators are still supplying a majority of the energy to the demand. Because of the pollution produce by these old-fashion DGs, utility companies are working on replacing the renewable DGs with the old method since renewable DGs are clean and more cost-effective. In [6], An independent hybrid solar-wind system with battery energy storage for a remote island has been developed. In the proposed system, the effects of the PV panel sizing, wind turbine sizing, and battery bank capacity on the system's reliability and economical performance were discussed. Finally, in order to evaluate the excellent condition of economic analysis, the author used a sensitivity analysis on the load consumption and renewable energy resource.

For providing adequate capacity due to the expansion of electrical demand [7], the multi-objective optimization technique for the multistage distribution expansion planning



(MDEP) considering the three following objectives: cost of investment and operation, END minimization, total active power loss minimization, and voltage stability index enhancement. The voltage limit of the buses, power flow transmission of feeders and branches, and power output of DG have been taken into account as constraints of the objective. To obtain an optimal solution, utilizing a modified particle swarm optimization algorithm has been measured as well as a novel mutation technique to advance the ability of global search and restrain the premature convergence to local minima.

The author in [8] developed the mixed-integer linear programming algorithm (MILP) to optimally allocate the type, size, and location of distributed generators in a distribution system. The radial distribution system's steady-state operation is built up, a different type of DGs, current capacity of a short circuit, and various topographies of distribution systems have been used. In this work, they represented MILP for optimally allocate the mentioned objective. The primary focus of the objective function is to minimize the annual investment and cost of utilization.

On account of the stochastic nature of charging and discharging schedule for electric vehicles, uncertainties in generating power from renewable resources such as wind speed and solar energy, and the possibility of load extension in the future that can put the optimal siting and sizing of distributed generators (DGs) in distribution system planning in risk. The authors in [9], developed a mathematical model of chance-constrained programming (CCP) for the objective function based on minimizing the cost of investment, operation, maintenance, network loss, and capacity adequacy. Constraints of the power flow equation, the upper and lower bound of real and reactive, as well as voltage limits have been taken into consideration. a Monte Carlo simulation-embedded genetic-algorithm-based approach has been applied to resolve the CCP model. The system under study in this work is the IEEE-37-node system.

Due to the significant impact of the distributed generators on the stability of voltage margins, remarkable attention has been drawn to this section. In [10], the optimal placement and sizing of distributed generators is considered to maximize the voltage stability margins in the distributed generation system. The mixed-integer nonlinear programming algorithm has been developed to solve the mathematical problem. Constraints under study are the system voltage limits, feeders' capacity, and the DG

penetration level, the maximum penetration on each bus, power flow equations, and branch current equation. The system under examination is a rural distribution system of 41 buses.

Locating the distributed generators need some particular methods to control the DG's independent outputs. [11], using the objective function of minimizing the computational burden and sensitive analysis to minimize the operating energy loss. Different constraints regarding the number of DGs to be allocated, constraint on dispatched wind power, power flow equations, voltage magnitude at different buses, and line loading have been taken into account. For picking the location of DGs precisely, a sensitivity analysis method has been proposed as well as an evolutionary programming technique that optimally locates the distributed generators. In this work, the system under study is a rural 69-bus distribution test system.

In [12], the objective function has been represented as the minimization of the power losses in the distribution system. Some of the constraints under examination are real and reactive power balance and limitation in variables considering active and reactive power, and voltage magnitude, and real power generation limits. Firstly, one of the easiest ways to calculate overload power is to employ sensitivity factors. Thus, this method has considered to point out the best candidate busses for placing distributed generators. Furthermore, an evolutionary computation technique known as practice swarm optimization (PSO) has been used in order to decrease the search space. Finally, they used optimal power flow techniques in order to optimally size individual DGs. Similar to [2, 10], they considered the 69-bus distribution test system.

One of the proper ways to relieve the uncertainties problem caused by renewable distributed generators such as wind speed and solar energy is to apply energy storage. Properly Locating and allocating the distributed storage system has a massive effect on the cost reduction in this project [13], divided the work into three stages. In the first stage, they optimally place the storage. Secondly, they came up with the number of storage units, and finally, they mixed the first and second stages together. The objective function of this study is to minimize the sum of the generation costs of all the generators over all time periods and the daily investment cost in storage. They used constraints on the binary variables, generator output, generator minimum up and downtimes, start-up costs, ramping and storage, and transmission constraints. Moreover, a sensitivity analysis has been considered

in order to weight the impact of wind speed on the choice of storage location. The system under study is tested on the IEEE RTS 96.

Because of the limitless and cleanliness of renewable energy resources, they play an essential role in a sustainable energy supply infrastructure. In [14], they proposed the methodology based on allocating different types of renewable distributed generators (DGs) in the distribution system to minimize energy losses without violating the system's constraints. A mixed-integer nonlinear programming (MINLP) method is evaluated to find the most optimal solution for the planning problem. The objective function is defined as minimizing the system's annual energy losses. Based on the results, there is a remarkable decrease in annual energy losses for all possible scenarios.

Current changes in renewable energy technologies brought utilities attention to redevelop the distribution system infrastructures. In [1], the problem focusing on optimally allocating wind-based DG units in the distribution system in order to minimize annual energy loss. Based on the stochastic nature of the wind, they used a probabilistic generation load model that integrates all possible operating conditions on wind-based DG units and load probabilities. Similar to [14], they used mixed-integer nonlinear programming (MINLP), with an objective function for minimizing the system's annual energy losses. Finally, the results indicated that the proposed technique is more reliable than the method considering the capacity factor.

The author in [15] highlights an upgraded multi-objective optimization algorithm for the generation of renewable distributed generators to provide energy for demand optimally.

The primary aim of work in [16] is to comprehensively optimize the size and site of the distributed generation in the distribution system. So that, considering objective based on minimizing investment and operating costs, full payments of compensating for system losses during the planning period, in addition to costs associated with alternative scenarios. In order to accomplish an accurate planning decision, binary decision variables are applied in the optimization model.

A hybrid system, including solar panels, wind turbines, diesel generators, and batteries, has been created to control the Pareto optimal front [17]. The objective of this

work is to determine the technological and economic expenses concerning cost and emission.

In [18], The distributed generation is formed by single and multiple active power, reactive power, and a mixture of active-reactive power DG. The optimal placement and sizing of distributed generation solved by a multi-objective particle swarm optimization method in the radial distribution system. The first objective is to minimize the power losses, and the second is to improve the voltage profile.

The optimal placement of the biomass-thermal and wind-based distributed generators (DGs) is considered in [19]. They developed their objective based on minimizing the waste of energy, emission, the total investment, and the cost of operation as well as enhancing the productivity of the voltage regulation. As the mathematical model developed based on the mixed-integer nonlinear problem, a genetic algorithm (GA) with an integrated optimal power flow (OPF) is taken into consideration. By using the GA method and OPF, the optimal location of the distribution power and optimal thermal power generation will be handled, respectively.

The primary goal of [20] is to increase the utility owner's benefit considering optimum battery energy storage system (BESS), which is combined with small wind turbines in the distribution network. For the purpose of minimizing annual energy loss and cost of energy, they formulated a multi-objective optimization based on scheduling the charge and discharge time of BESS. Furthermore, a genetic algorithm (GA) is considered to solve the problem.

In [21], due to the advantages of the mixed renewable energy resources in environmental, economic, and technical aspects, the optimal sizing of the mixed renewable energy resources is examined. They covered several different optimization techniques in order to solve the sizing problem. Consequently, every single of these optimization techniques could have a remarkable potential to advance renewable energy systems' applicability.

Authors in [22] aimed to install discrete full cell generators in a power system to increase the efficiency of the operation. They represented an algorithm to obtain the near-optimal solution for locating these units and system losses on the network. Furthermore, resistive losses and capacity savings are the huge factors of the impacts of discrete

generation at the distribution level. The results represent the crucial impact of the placement on reducing the losses and increasing the capacity savings. Proper placement of the DGs can decrease the energy losses and boost the feeder voltage profile.

In [23], the fuzzy goal programming (FGP) is developed to solve the multi-objective DG allocation subsume the voltage characteristics of every load component and total capacities of DG units.

In [24], several attempts have been accomplished in order to reconstruct the old network into the smart grid. To begin with, they considered solar energy over other renewable resources because of the availability in higher scop. This paper aims to minimize network power losses and raise the voltage stability considering system operation and security constraints in a transmission line. They optimally sized and sited the DGs by using the following steps. In the first step, the optimal size of DGs is decided by conducting the Particle Swarm Optimization (PSO), and by using a negative load approach for reverse power flow, the possible sites of the DGs are selected. Then they found the optimal sites considering Loss Sensitivity Factor (LSF) and weak bus strategy. Besides, in the second step, they conducted a hybrid PSOGSA algorithm to locate the DGs optimally.

The authors in [25] used the ant lion optimization algorithm (ALOA) to locate and size the DG based renewable sources for different distribution systems optimally. The optimization planning problem is formulated to compute power losses, voltage profiles, and VSI. In the first step, the loss sensitivity factors (LSFs) were considered to introduce the most candidate buses for installing DGs. Then by using ALOA, the location and size of DGs were selected. Additionally, to validate the performance of the proposed algorithm, the obtained results are compared with other algorithms in this area.

In [26], the authors consider a multi-objective function to locate and size DGs and reconfigure the distribution network. He considers the expected costs but skips the investment cost. An optimal evolutionary strategy based on the Pareto optimality is used as well as a fuzzy set theory to pick the best solution among the achieved Pareto set. Based on the results, the effectiveness of the proposed method has been proved.

## 2.3 Stochastic Modeling for Load and Generation

The authors in [27] analyzed the performance of the gaussian mixture model (GMM) versus three well-known parametric models considering wind speed data for 6 locations in northwest Europe. Based on the analysis, GMM is selected as the most accurate statistical model in different aspects, such as the percentage of improvement in mean absolute error (MAE) and root mean square error (RMSE), and the goodness of fit K-S test.

In [28], the author analyses the mixture of two truncated normal distributions (MTTND) with different means and variances for each component. By using the mixture of two Normal distributions, they are aiming to represent a wind speed probability density function (PDF) in a more flexible shape. Based on the analysis, the accuracy of the MTTND was confirmed by comparing it with other well-known pdfs used in other works of literature like; weibull, rayleigh, lognormal, gamma, and inverse Gaussian and Burr. In order to estimate the parameters of the pdfs, the author used the least-squares non-linear regression method. For ensuring the accuracy of the MTTND over other pdfs, The  $R^2$  for the goodness of fit, and a statistical error measurement (RMSE) are taken into consideration.

Because of the different modes in mixture distributions, they have more accurate performance than the single distributions in the modeling of wind speed. Hence, reference [29] demonstrated the use of hierarchical mixtures of multiple distributions for the modeling of wind speed. For evaluating the parameters of each component of the mixture model, They used a nested expectation-maximization (EM) algorithm. Based on the results of the analysis, the proposed model performed better in comparison with the single and mixture distributions.

The authors in [30] propose a non-parametric kernel density estimation (KDE) model for wind speed distribution. Wind speed distribution inevitably plays a fundamental role in wind-based distributed generation placement. One of the essential ways to stop over-fitting and under-fitting is to determine the model's accurate bandwidth. The paper approach is to use a wind speed sample-based technique to estimate the optimal bandwidth considering three years of wind speed data. A goodness-of-fit test and mean square root errors are used to assess the pdfs' performance versus histograms. Based on the final results,

the better productivity of the KDE model over the conventional parametric distribution model is clearly illustrated.

The analysis in [31] considered the use of two-parameter Weibull distribution for three different sites in India. Based on the analysis, it is shown that the approximated method for evaluating the parameters of the weibull probability distribution could be a proper approach for predicting the wind speed and wind power accurately.

Manage the conversion of wind energy, evaluate the wind energy potential, and place wind farms in three crucial factors that need to be assessed in the estimation of wind speed distribution. In [32], instead of conventional wind speed planning, a new approach is addressed with more reliable wind speed distribution. This approach integrates the bayesian model averaging (BMA) and markov chain monte carlo (MCMC) sampling method. They defined BMA probability density function (PDF) of the wind speed as a mean of the model PDFs considered in the model space weighted by their posterior probabilities over the sample data.

In [33], an advanced non-parametric strategy is examined to assess the performance of wind speed probability distribution. The method includes bandwidth selection and boundary correction of kernel density estimation (KDE). For validating the proposed model, two goodness-of-fit tests are used. Consequently, the diffusion-based kernel density model (DKDM) results are compared with multimodal, weibull, and normal distributions. This outcome resulted in DKDM providing more accurate results, even when the wind speed is not present. Afterward, they concluded that a single parametric distribution was unable to model wind speed data in different sites adequately.

In [34], the maximum likelihood estimation (MLE) via particle swarm optimization (PSO) algorithm was used to calculate the parameters of the mixture of two Weibull distribution, including complete and multiply censored data. They performed a simulation study to evaluate the capability of the MLE via the PSO algorithm, quasi-newton method, and an expectation-maximization (EM) algorithm with different parameter frames and sample sizes in both complete and deleted cases. Based on the results, they represented that the PSO algorithm performs better than the quasi-newton method and the EM algorithm in most cases in terms of bias and root mean square error. In order to indicate the performance of the proposed method, two numerical examples are taken into account.

In [35], a multivariate Kolmogorov-Smirnov (K-S) goodness-fit-test was taken into account. They developed Rosenblatt's transformation and an algorithm in the goodness-of-fit test to determine the bivariate case test. Additionally, in order to simply calculate any dimension, an approximate test is used. Finally, to understand the accuracy of the multivariate test, they apply the test in a simulation scope.

Misconception and misuse of statistical tests, confidence intervals, and statistical power have been denounced for years. However, the correct use of these statistics could be significantly effective in different criteria such as medical study, finance, physics, forecasting, and many other fields. In [36], The definitions of fundamental statistics that are more general and critical have been provided. The authors stressed how violation of unstated analysis formalities could conclude to small p-values. Then, a list of 25 wrong explanations of p-values, confidence intervals, and power has been provided to enhance the statistical examination.

The author in [37] explained the evaluation of a probability density function's (PDF) problem and decision regarding the mode of a probability density function (PDF). They represented how one can build a family of approximations of the PDF and the consistency mode. The problem of calculating the mode of a probability density function and the maximum likelihood estimation problem for estimating the parameters have a bit of similarity with each other.

In [38], Athaurs considered a parametric beta and a non-parametric kernel density estimation model, which are assessed to obtain a more reliable statistical model. Additionally, the use of goodness-of-fit and four statistical error measurements have been taken into consideration. Accordingly, they demonstrated that the hybrid model of the solar irradiance represents more appropriate results in terms of the percentage of improvements versus the beta and KDE model. Furthermore, based on hypothesis testing, the hybrid model is the only one that refuses to reject the null hypothesis in all cases. Four statistical error measurements, such as mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and mean bias error (MBE), are used to represent the high performance of the hybrid model versus other distributions. Results confirm the accuracy of the hybrid model for solar irradiance modeling with percentage improvements over the beta distribution of up to 13.8% (RMSE), 11.7% (MAE), 19.3% (MAPE), and



72.5% (MBE). The K–S test results show that the proposed beta KDE hybrid is the only model for which the null hypothesis is not rejected for any of the eight data sets considered in this study

## 2.4 Application of Genetic Algorithm in Planning

In reference [39], they proposed a useful model to optimally find the size and site of the distributed generation units in the distribution system. They introduced the genetic algorithm (GA) to minimize total active and reactive power losses and bring the voltage profile to acceptable limits. Furthermore, the GA fitness function is taken into consideration, such as the active power losses, reactive power losses, and the cumulative voltage deviation variables with selecting the weight of each variable. Consequently, by minimizing the fitness function, the optimal solution has been accomplished.

The author in [40] used the combination of the analytical expressions and optimal power flow (OPF) algorithm to locate, size, and select the best mix of different DGs types optimally for minimizing the energy losses in the distribution system. Based on the simulation, the proposed work is verified for controlling the combination of different DGs types.

The use of genetic algorithm (GA) and ant colony algorithm (ACO) optimization techniques are represented in [41] to optimally size and site the distributed generators in the electrical grids. The objective function is based on the linearized model to calculate the active power losses of the generators. This strategy is based on a strong coupling between active power and power flow, taking into consideration the voltage angles with the end goal to exhibit the adequacy of the proposed method.

In [42], the use of optimal placement and sizing of dispatchable DGs and shunt capacitors is considered. Besides, the author's objective is to decline the annual energy losses and improve the voltage profile. The other considerations of this work is an average hourly variation of load demand profile. Additionally, The sensitivity analysis and GA method are used in order to gain the location and size of dispatchable DGs and shunt capacitors, respectively. The system under study in this work is the 33 bus system.

Since the concept of minimizing cost in energy management has always been an essential objective, the minimization of the emission has undoubtedly been as crucial as

the cost as well. In [43], the authors used a multi-objective real-time optimization for an islanded microgrid to solve in real-time, considering a multi-objective particle swarm optimization (MOPSO) algorithm is taken into consideration. Consequently, the results of the simulation are compared with the multi-objective genetic algorithm (GA) optimization package in Matlab. Based on the analysis, the proposed technique accomplished faster than the GA.

In [44], to site and size FCSs and wind-based DGs a new planning model is developed, considering the stochastic nature of FCSs and residential EV loads, and renewable energy generation. Different cost and revenue components associated with FCSs and wind-based DGs as well as practical constraints for installing FCSs and DGs in the area of each bus are taken into account. The proposed problem is classified as mixed-integer non-linear programming (MINLP) which requires a lot of time to be solved in deterministic method. Consequently, the genetic algorithm (GA) is considered to solve the proposed problem. The computational results illustrate that the proposed planning method is an effective way in siting and sizing FCSs and wind-based DGs while maximizing the profit, minimizing the energy losses, and deferring the need for DN upgrades.

## 2.5 Concluding Remarks

An apparent gap in the literature has been observed. First of all, Most of the research papers in stochastic modeling only used the well-known probability density functions (PDFs) to follow the wind speed pattern precisely. Furthermore, to minimize the annual energy losses in renewable distributed generators (DGs), the probabilistic generation load model considering Weibull and Rayleigh's probabilities was used in literature to fix into the deterministic optimal power flow (OPF) equations [1]. In contrast, we applied different PDF's probabilities, including unimodal and multimodal PDFs, to the planning problem to analyze the performance of each PDF in minimizing the annual energy losses.

# CHAPTER 3: MODELING OF WIND-BASED RENEWABLE RESOURCES

## 3.1 Introduction

Due to the global severe environmental pollution problem, emissions of carbon dioxide and greenhouse gases produced by traditional industries (non-renewable energy resources), most of the attention have shifted to the utilization of new energy resources, which could be an ideal solution to dealing with the limitation of non-renewable energy and climate change. Recently, the development and utilization of renewable energy, including solar, geothermal, and wind energy, have drawn much attention because of different reasons. Controlling the emission of dangerous gases in the atmosphere, decreasing the cost of energy-related to the traditional energy resources, and participation of power producers in the electricity market system are some of the reasons for developing renewable energy resources in the distribution system.

Among all other renewable energies, wind energy has had the fastest-growing source of electricity all around the globe. In Canada, the capacity of wind power has increased with a very swift slope for about 12000 megawatts from 2004 to 2018. Figure 3.1 shows the growth of Canada's installed capacity over time, based on data from [45].

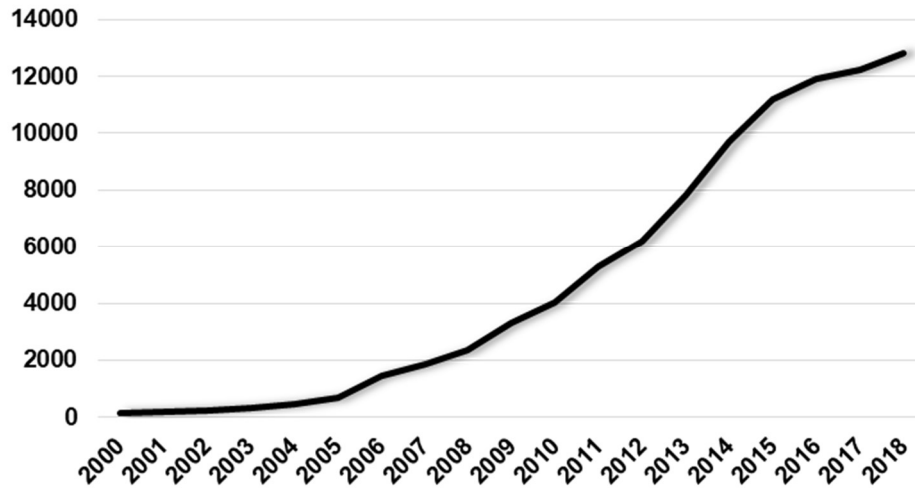


Figure 3.1: Canada's installed wind power capacity

Based on statistics, Ontario has the largest installed wind capacity with 40% of Canada's wind energy capacity, this trend followed by Quebec, Alberta, British Columbia, and finally, Nova Scotia, respectively. Figure 3.2 represents the wind capacity in different provinces in Ontario, based on data from [45].

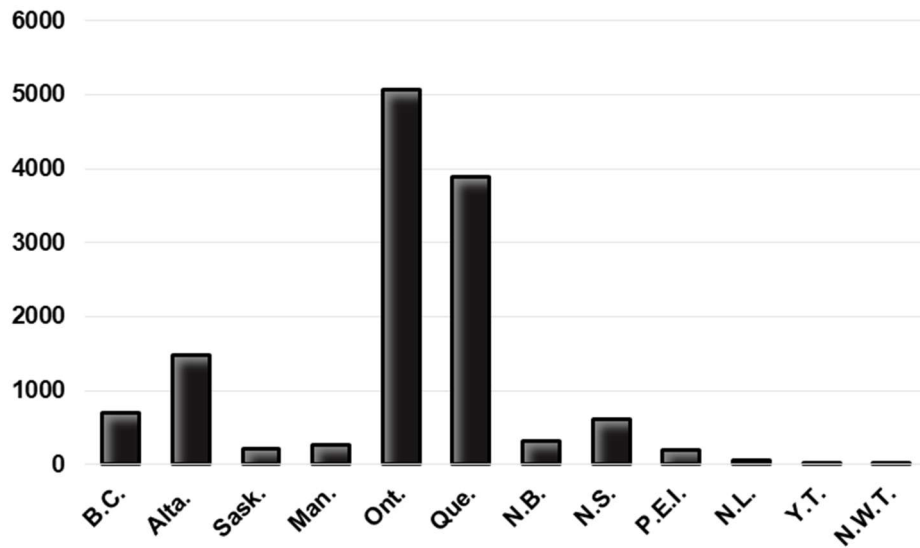


Figure 3.2: Ontario installed capacity

Wind energy remarkably relies on wind speed. Understanding the wind speed characteristic at a specified location to develop an appropriate probabilistic model that represents the wind speed pattern correctly is the initial step to place the renewable distributed generation units in the distribution system.

There are different approaches for modeling the renewable resources in the distribution system. For instance, wind speed can be modeled using different time-scales or using a specific probability density function in a certain period. The proper selection of a strategy for modeling a renewable resource such as wind speed depends on the period of the application as well as utilizing the analytical methods, specifically the Maximum Likelihood Estimation (MLE), Particle Swarm Optimization (PSO), and Mont Carlo Simulation accomplish the application [46].

Accordingly, for modeling the stochastic behavior of the wind speed, several probability density functions are examined and compared with each other. In this chapter, the most well-known probability density functions (pdfs) versus a mixture of Gamma distribution are considered and to ensure the accuracy of the mixture model different error tests are used. Consequently, after assessing the performance of different distributions, the mixture of Gamma distribution is selected as the most accurate model for the modeling of wind speed. In the following section, an analysis of the factors that might significantly impact wind speed modeling will be examined.

## 3.2 Factors Impacting Wind Speed

For modeling the wind speed in this research, the annual wind speed data is taken into consideration. However, an analysis is accomplished to observe how different factors impact the accuracy of wind speed modeling. We need to analyze the observation to identify any factor that contributes to the wind speed. If we disregard them, probably the prediction model can be inaccurate (i.e., the total error is beyond the acceptable range). Therefore, before establishing any prediction model, we need to determine wind speed's main effects using data analytic tools. Factors are split into the following category: year, season, weather, temperature, and time.

### 3.2.1 Impact of year on wind speed

We observe the wind speed in two years, 2016 and 2017. We have 8500+ observations each year. Based on the analysis the wind speed is not impacted by the year, as shown in Figure 3.3. We conducted the ANOVA test, and the p-value is 0.433646, which supports the finding statistically. Based on what observation in Figure 3.3, the prediction of the wind speed can be applied to next year.

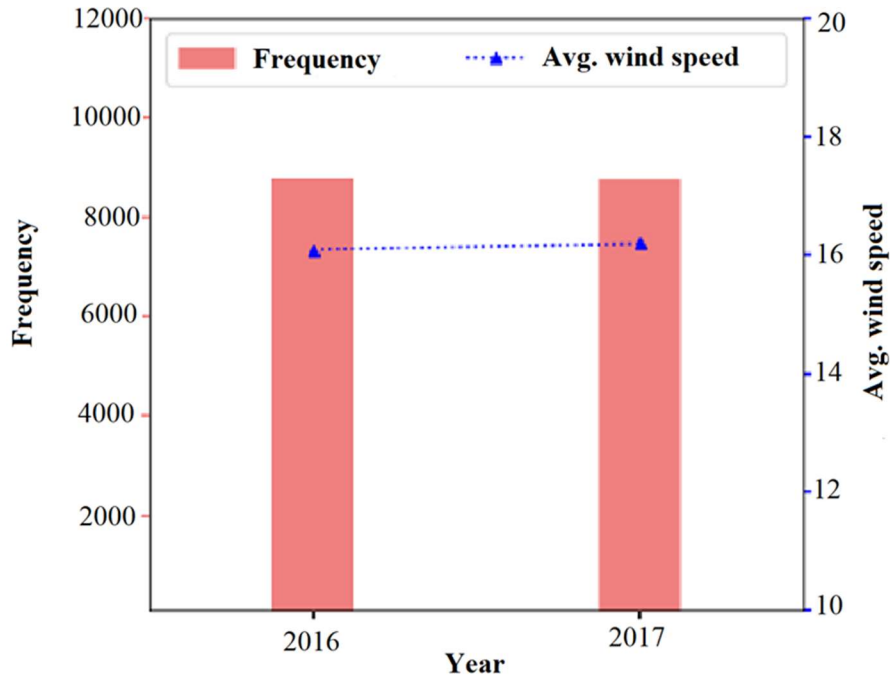


Figure 3.3: Average wind speed of years 2016 and 2017

### 3.2.2 Impact of season on wind speed

We observe the wind speed for four seasons, winter, spring, summer, and fall, in both years 2016 and 2017. We have 4000+ observations for each season. The season, as the main effect, is significant, as shown below. We conducted the ANOVA test, and the p-value is absolute zero, which supports the finding statistically. We cannot predict precisely the wind speed (i.e., fitting a probability distribution to it) unless we consider the season. If we disregard this main effect, we, in fact, consider that as a noise (uncontrollable error), and this vastly increases the total error. Of course, this does not necessarily mean we can have a precise distribution if we consider the season, as there might be other central effects. Figure 3.4 shows the average wind speed in different seasons.

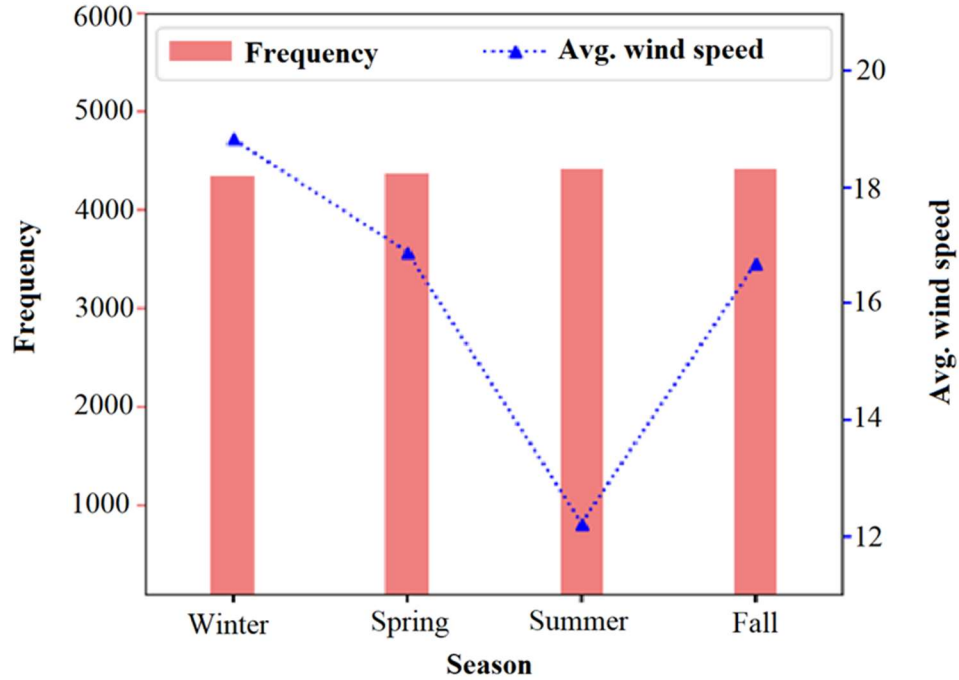


Figure 3.4: Average wind speed of different seasons

The average wind speed in different months has been plotted. There is a clear cyclic trend that iterates every year. Figure 3.5 illustrates the average wind speed for each month.

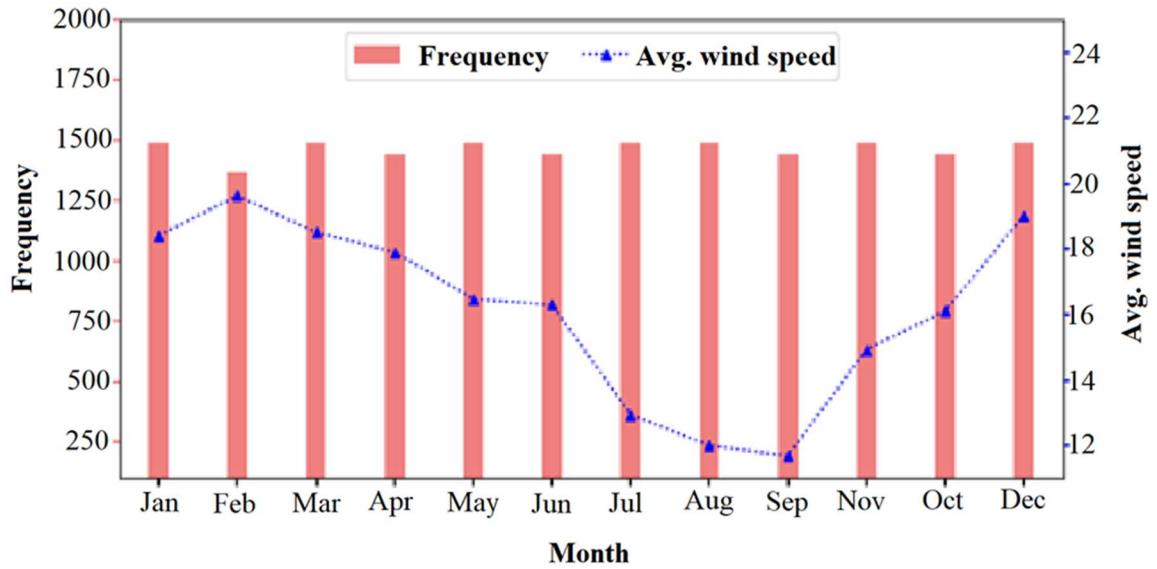


Figure 3.5: Average wind speed of different months

### 3.2.3 Impact of weather on wind speed

The wind speed for the top 10 most frequently weather types are observed. For example, the most frequent weather is mostly cloudy, with 1600+ observations. The weather, as the main effect, is significant, as shown below. We conducted the ANOVA test, and the p-value is absolute zero, which supports the finding statistically. We cannot precisely predict the wind speed (i.e., fitting a probability distribution to it) unless we consider the weather. If we disregard this main effect, we, in fact, consider that as a noise (uncontrollable error), and this vastly increases the total error. Of course, this does not necessarily mean we can have a precise distribution if we consider that as there might be other central effects. The weather is different from the season in that it's prediction. However, we can conclude that if we predict the weather type, the prediction of wind speed can be more accurate.

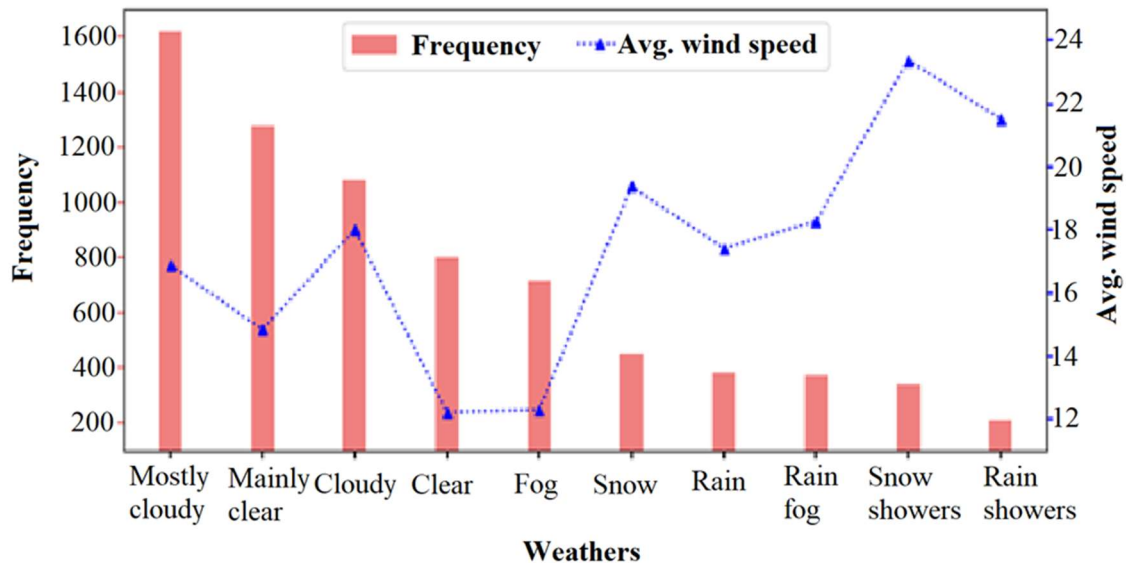


Figure 3.6: Average wind speed of different weather's type

### 3.2.4 Impact of temperature on wind speed

We plot all observations, the wind speed versus the temperature. From this scatter plot, the temperature is not a significant main effect. That is, Temperatures 10C and 20C in summer are expected to have the same impact on the wind speed. While the wind speed



in winter and summer with the same temperature is different. Figure 3.7 shows the wind speed at different temperatures.

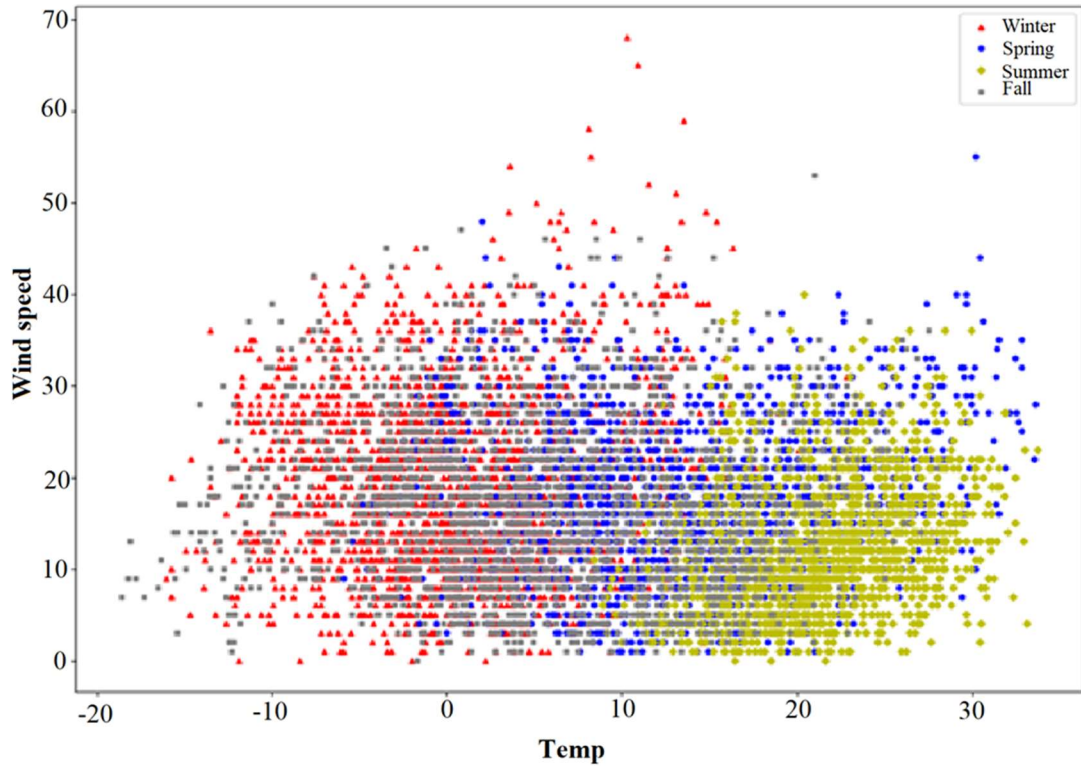


Figure 3.7: Wind speed at different temperatures

### 3.2.5 Impact of time on wind speed

We plot all observations, the wind speed versus the time in a day. From this scatter plot, time in a day is not a significant main effect. That is, the wind speed of day and night are almost similar.

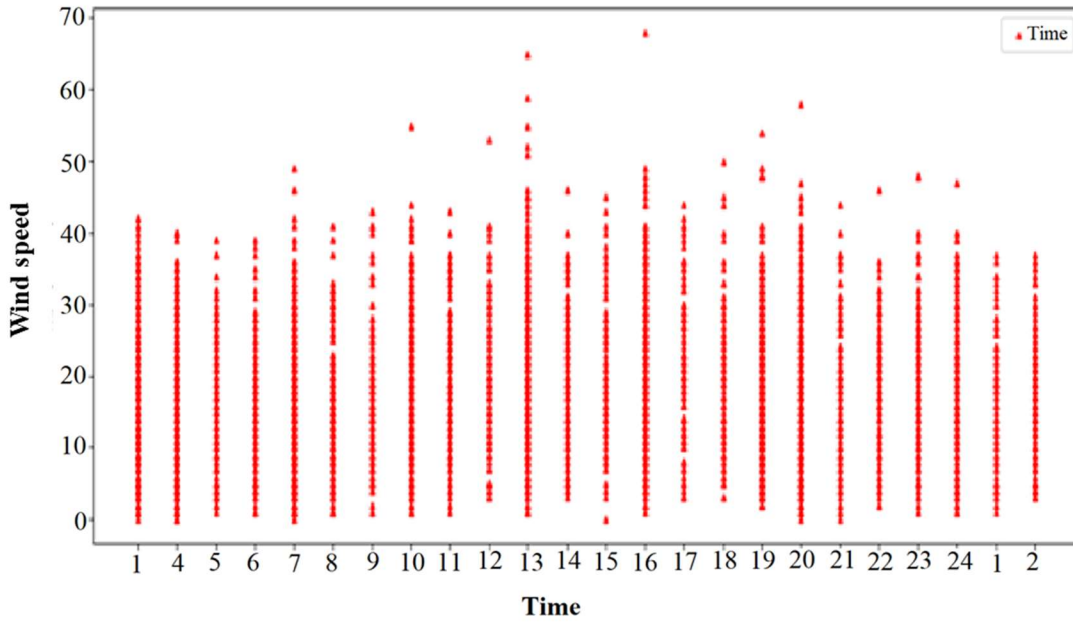


Figure 3.8: Wind speed in different time segments

Fitting probability distribution to the wind speed is a function of several main effects (like season and weather type), and we can more accurately predict the probability of wind speed if we consider season and weather as factors. The consideration, here, for example, means we develop a probability distribution for each season and weather type. There is no generic probability distribution for the wind speed that predicts the season and weather type. We can fit a probability distribution to the wind speed more precisely if we know the season and weather type.

### 3.3 Wind Speed Stochastic Modeling

Based on the new approximations, wind energy is the lowest-cost source of new electricity generations as well as being both commercially and technologically competitive. Wind energy is not only one of the most environmentally friendly sources of energy, but also it is one of the most satisfying for the requirement of energy. Nevertheless, based on the random nature of the wind speed, selecting an accurate wind speed model is a challenging problem. However, by selecting a proper probability density function (PDF) and applying the wind-based DGs in the distribution system, answers to this issue can be addressed.

### 3.3.1 Unimodal probability density functions in wind modeling

Precisely estimating wind speed frequency distribution will be beneficial for predicting the wind energy potential accurately and selecting wind energy planning optimally [29]. To begin with, a wind speed distribution mainly identifies the performance of the wind power system. Modeling wind speed, considering an accurate pdf, gives crucial parameters that uncover the characteristics of wind speed data. Based on this statement, using wind modeling is remarkably beneficial for the extended period planning problem. The four unimodal distributions (Weibull, Rayleigh, Gamma, and Johnson SB) will be explained, and the performance of each distribution will be measured.

#### A. Weibull distribution

Among various distributions, the Weibull probability distribution function (3.1) is one of the most well-known pdf to describe the stochastic behavior of the wind speed [47]. The Weibull PDF's success comes from the adjustable parameters and the flexibility in fitting distribution function to the estimated value with different patterns.

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \quad (3.1)$$

where  $k$  is the shape parameter, and  $\lambda$  is the scale parameter of the weibull distribution. In the two parameters weibull distribution, the location parameter ( $\gamma$ ) consider as 0.

Weibull functions are customarily used to describe the random behavior of the wind speed in a given location over a certain period, in most cases annually. Furthermore, the weibull function can describe wind speed distribution for a typical hour of the year. In the following paragraphs, the performance of the other distributions versus weibull distribution will be addressed.

#### B. Rayleigh distribution

A perfect expression often used to model wind speed behavior is the Rayleigh probability density function (PDF). If  $k = 2$ , then we have a particular case of the Weibull distribution called the Rayleigh distribution [1], which distribution density (3.2) is:

$$f(v) = \left(\frac{2v}{h^2}\right) \exp\left[-\left(\frac{v}{h}\right)^2\right] \quad (3.2)$$

When the shape index  $k$  equals 2, the pdf is called Rayleigh probability density function as given in (3.2), which pdf mimics most wind speed profiles. If the mean value of the wind speed for a site is known, then the scale index  $h$  can be calculated as in (3.3) and (3.4).

$$v_m = \int_0^{\infty} v f(v) dv = \int_0^{\infty} \left(\frac{2v^2}{h^2}\right) \exp\left[-\left(\frac{v}{h}\right)^2\right] dv = \frac{\sqrt{\pi}}{2} h \quad (3.3)$$

$$c \cong 1.128 v_m \quad (3.4)$$

### C. Gamma distribution

Gamma distribution is a two-parameter continuous distribution where  $\alpha$  is the shape parameter, and  $\theta$  is the scale parameter. In some cases, there is different parameterization use. In which, the shape parameter  $\alpha$  and an inverse scale parameter  $\beta = \frac{1}{\theta}$  known as rate parameter. The density function of the Gamma distribution (3.5) is:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (3.5)$$

### D. Johnson SB distribution

As stated in [48], the Johnson system of distributions is a very flexible distribution for every mean and standard deviation. The specific Johnson family that is considered in this research is the 4-parameter Johnson SB distribution, whose distribution (3.6) is:

$$z = \gamma + \delta \ln\left(\frac{X - \xi}{\xi + \lambda_o - X}\right), \quad \xi \leq X \leq \xi + \lambda_o, \quad (3.6)$$

where  $\xi$  is the location parameter,  $\delta > 0$  (by convention) and  $\gamma$  are shape parameters, and the parameter  $\lambda_o > 0$  corresponds to the range. The distribution of  $X$  is the Johnson SB distribution, which is defined as follows:

$$f(x) = \frac{\delta}{\lambda_o \sqrt{2\pi z(1-z)}} \exp \left[ -\frac{1}{2} \left( \gamma + \delta \ln \frac{z}{1-z} \right)^2 \right] \quad (3.7)$$

The Johnson SB distribution is also known as the four-parameter lognormal distribution. Johnson SB distribution is double bounded;  $\gamma = 0$  indicates symmetry. The corresponding limiting form (as  $\delta \rightarrow \infty$ ) is the Gaussian distribution. Johnson SB distribution is flexible, covers bounded distributions, as well as a wide variety of distributional shapes, including the gamma and beta distributions.

### 3.3.2 Annual wind speed modeling for unimodal distributions

In this model, Weibull, Rayleigh, Gamma, and Johnson Sb distributions are used to represent the annual wind speed distribution in the two sites under study. The historical data are collected from [49]. The annual mean and standard deviation of wind speed are calculated. Besides, the parameters of Weibull pdf (wind model 1), Rayleigh pdf (wind model 2), Gamma pdf (wind model 3), and Johnson SB pdf (wind model 4) are calculated. By using the parameters of each wind speed model, the probabilities of those distributions are predictable. For the modeling of the wind speed, 48 records of hourly wind speed are considered for each month, which means that for the whole year, 576 records are collected.

This analysis covers two different areas in Ontario, Canada (Windsor and Hamilton). The parameters of the unimodal distributions are as follows:

Table 3.1: Parameters of the unimodal distributions

Site	Gamma		Rayleigh	Weibull		Johnson SB			
	$\alpha$	$\beta$	$\sigma$	$k$	$\lambda$	$\gamma$	$\delta$	$\lambda_o$	$\xi$
<b>Windsor</b>	2.9789	5.1101	12.1460	1.9304	17.0090	2.8935	1.6728	108.4600	2.7589
<b>Hamilton</b>	3.7502	5.0530	15.1200	2.0833	21.2920	0.7740	1.0274	51.7230	0.9744

Using these parameters, we will estimate the p-values and two types of error measurements to rank the candidate distributions based on their accuracy. The following sections represent the strategies used for picking the most accurate unimodal probability density function.

### 3.3.3 Goodness-of-fit Kolmogorov-Smirnov test

The goodness of fit test examines if a particular sample data can adequately fit a distribution from a specific population. In other words, the goodness of fit seeks if a specific data set shows the data expected to find in the observed population. In this research, the well-known goodness of fit test known as the Kolmogorov-Smirnov (K-S) test is used to assess the performance of each unimodal wind speed model. The test statistic (3.8) of the K-S test is given by:

$$D = \sup_x |F_0(x) - F_{data}(x)| \quad (3.8)$$

where  $F_0(x)$  is the cumulative distribution function (CDF) of the hypothesized distribution, and  $F_{data}(x)$  is the empirical distribution function (EDF) of the observed data. The K-S test measures the immense distance between the EDF or  $F_{data}(x)$  and the theoretical function  $F_0(x)$ , measured in a vertical direction [50]. Every individual step to estimate the K-S test is as follows:

- 1) Developing an EDF for the particular data set
- 2) Specifying a specific distribution
- 3) plotting two distributions together.
- 4) Estimating the most significant vertical distance between the two graphs.

- 5) Calculating the test statistic.
- 6) Finding the critical value in the K-S table, and
- 7) Comparing the critical value.

## A. Hypothesis testing

The primary focus of this section is to challenge the new probability distribution functions (PDFs) versus the most well-known pdfs in stochastic wind modeling. Because of the randomness in the wind speed, it is crucial to select a pdf that can precisely follow the observed data set (Histogram). In other words, the closer the probability of a specific model to the observed data, the more accurate the model would be. The first step for selecting and validating a candidate model is the Kolmogorov–Smirnov (K-S) goodness-of-fit test that has been taken into account to come to a resolution that collected data comes from a specific statistical model. To conclude a shred of substantial evidence, results are picked based on p-value ( $p > 0.01$ ) analogous in strong proof that the null hypothesis is correct. Consequently, the data does not return remarkable results if a false null hypothesis is obtained at the end of the goodness-of-fit test.

### 3.3.4 Average model performance

There has been tremendous growth in the number of climate and environmental models over the last few decades. Attention also has risen in deciding which statistical strategy delivers a more appropriate and accurate approximation of the variables of interest [51]. Thus, error approximation for comparing model produced with different observations have been used in massive applications. Based on what literature commonly applied to their work, two types of error measurements are used to differentiate the productivity of the considered unimodal distribution. These two error measurements are root mean square error (RMSE) and mean absolute error (MAE). Both of these measures are dimensioned to express average model prediction error in the unit of the variable of interest [51].

#### A. Root mean square error (RMSE)

In the climate and environmental literature, primarily wind modeling, the RMSE and MAE are the most used error measures to check the performance of the statistical model. Three primary steps have to be taken in order to estimate the RMSE: The first step

is to estimate the total square error as the sum of discrete squared error. In fact, each individual error has an impact on the total error. As it is expected, the higher errors have more influence on the total square error than the smaller errors. In the second step, after obtaining the total square error, it is time to divide the total square error by  $n$  and obtain mean square error (MSE) as an output. Then in the final step, the RMSE can be calculated as the square root of the MSE. The RMSE formulation is given by (3.9):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (3.9)$$

where  $n$  is the number of bins of the specific data set.  $x_i$  Represents the number of observations,  $\hat{x}_i$  shows the probability of the wind speed and  $i$  is the calculated bin from the data set ( $i = 1, \dots, n$ ).

## B. Mean absolute error (MAE)

The calculation of MAE is relatively simple. It involves summing the magnitudes (absolute values) of the errors to obtain the 'total error' and then dividing the total error by  $n$ . The MAE (3.10) is given by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (3.10)$$

where  $n$  is the number of bins of a particular data set.  $x_i$  Represents the number of observations,  $\hat{x}_i$  shows the probability of the wind speed and  $i$  is the calculated bin from the data set ( $i = 1, \dots, n$ ).

### 3.3.5 Multimodal distribution (mixture distribution)

Because of the complex wind regimes in the real world, single distributions can not describe the wind speed distributions comprehensively. Consequently, with different components (different number of distributions), we can model the wind speed distribution more appropriately. In this work, to investigate the effectiveness of the mixture modeling,



the package `mixR` in R is considered. This package includes the maximum likelihood estimation (MLE) method for finite mixture models through the expectation-maximization (EM) algorithm. Moreover, the bootstrap likelihood ratio test is used to select the best statistical model. The mixture models used in this research are a mixture of Normal (MoN), a mixture of Weibull (MoW), and a mixture of Gamma (MoG) probability density functions [52]. The general form of the finite mixture model is given by (3.11):

$$f(x; \Phi) = \sum_{j=1}^M \pi_j f_j(x; \theta_j) \quad (3.11)$$

where  $f(x; \Phi)$  is the probability density function (PDF) of the mixture model,  $f_j(x; \theta_j)$  is the pdf of the  $j_{th}$  component of the mixture model,  $\pi_j$  is the weight of the  $j_{th}$  component and  $\theta_j$  is the parameter of the  $j_{th}$  component, and  $\Phi$  is a vector of all the parameters of the mixture model. By considering the EM algorithm, the parameters of a specific mixture model can be estimated.  $n_i$  is the number of observations that fall in the  $j_{th}$  bin, for  $i = 1, 2, \dots, r$ , and  $r$  is the total number of bins. In order to estimate the maximum likelihood of the finite mixture model for binned data, two types of variables are represented:  $x$  and  $z$ , in which  $x$  shows the value of unknown observations, and  $z$  is a vector of zeros, and one illustrates the component that  $x$  belongs to. The complete data log-likelihood is given by (3.12) [52].

$$Q(\Phi; \Phi^{(p)}) = \sum_{j=1}^M \sum_{i=1}^r n_i z^{(p)} [\log f(x^{(p)}; \theta_j) + \log \pi_j] \quad (3.12)$$

where  $z^{(p)}$  is the expected value of  $z$  given the estimated value of  $\Phi$  and expected value  $x^{(p)}$  at  $p_{th}$  iteration. In the expectation-step (E-step) of the EM algorithm, we create a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters  $\Phi$ . Then in the maximization-step (M-step), the parameters maximizing the expected log-likelihood found on the  $E$  step (calculate the expected value of the latent variables  $x$  and  $z$ ) will be calculated. These parameter-estimates are then used to determine

the distribution of the latent variables in the next E step. Then we have to repeat steps 1 and 2 until convergence. The EM algorithm is ended by the stopping rule. Sometimes the M-step of the EM algorithm does not return an appropriate solution. In that case, an iterative approach, such as Newton-Raphson (NR) or bisection method, can be used. In the case of a lack of information regarding the number of components  $g$ , its value should be calculated using data. The following steps represent the bootstrap likelihood ratio test [52]:

- 1)  $\hat{\Phi}_0$  and  $\hat{\Phi}_1$  show the null and alternative hypotheses, respectively. Estimate the observed log-likelihood  $\ell(x; \hat{\Phi}_0)$  and  $\ell(x; \hat{\Phi}_1)$ . The likelihood ratio test statistic is given by:

$$w_0 = -2(\ell(x; \hat{\Phi}_0) - \ell(x; \hat{\Phi}_1)) \quad (3.13)$$

- 2) In this step, the model produces random data similar to the original data used for the estimated parameters  $\hat{\Phi}_0$ . Then step number one starts over again. By doing this process repeatedly for  $B$  times, a vector of the simulated likelihood ratio test statistic will be estimated  $w_1^1, \dots, w_1^B$ .
- 3) As the final step, the empirical P-value will be approximated based on the following formula:

$$P = \frac{1}{B} \sum_{i=1}^B I(w_1^{(i)} > w_0) \quad (3.14)$$

Package `mixR` in R uses the bayesian information criterion (BIC) technique to analyze the appropriate number of components. The EM algorithm requires the number of mixture components  $g$  as an a priori input. By increasing the number of components or parameters, the log-likelihood function can increase. However, the model's complexity will increase and it may cause overfitting [27].

The BIC applies a penalty term to the log-likelihood function as the number of components is increasing. The BIC function is given by (3.15):

$$BIC = -2 \ln Pr(\hat{\theta} | x, g) + g \ln(n) \quad (3.15)$$

In this work, the BIC is unified with the EM algorithm to estimate the correct number of components by increasing the log-likelihood function.

## A. Expectation-maximization (EM) algorithm

Computational-wise, the EM algorithm remarkably prevalent in statistics. While numerical optimization is not easy to accomplish, the implementation of two steps (E-step & M-step) in the EM algorithm is comfortable for various statistical problems [53]. In order to estimate the parameter via the EM algorithm, the following steps need to be taken:

1. Input:
  - Wind speed data:  $x = \{x_1, x_2, \dots, x_n\}$
  - Number of the component in the mixture model (MoW, MoN, and MoG)
  - Initialize the set of parameters  $\theta$  for a specific pdf.
  - Initialize the weight  $\pi_i$  of each component.
  - Tolerance  $\varepsilon$ .
2. Process:
  - Expectation-step: In the E-step, by using the initialized parameters and weights, the expectation of the log-likelihood will be calculated.
  - Maximization-step: computes parameters maximizing the expected log-likelihood found on the  $E$  step.
  - Repeat E-step and M-step until it converged
3. Output
  - The optimal set of parameters ( $\theta$  and  $\pi_i$ )
  - BIC & log-likelihood values

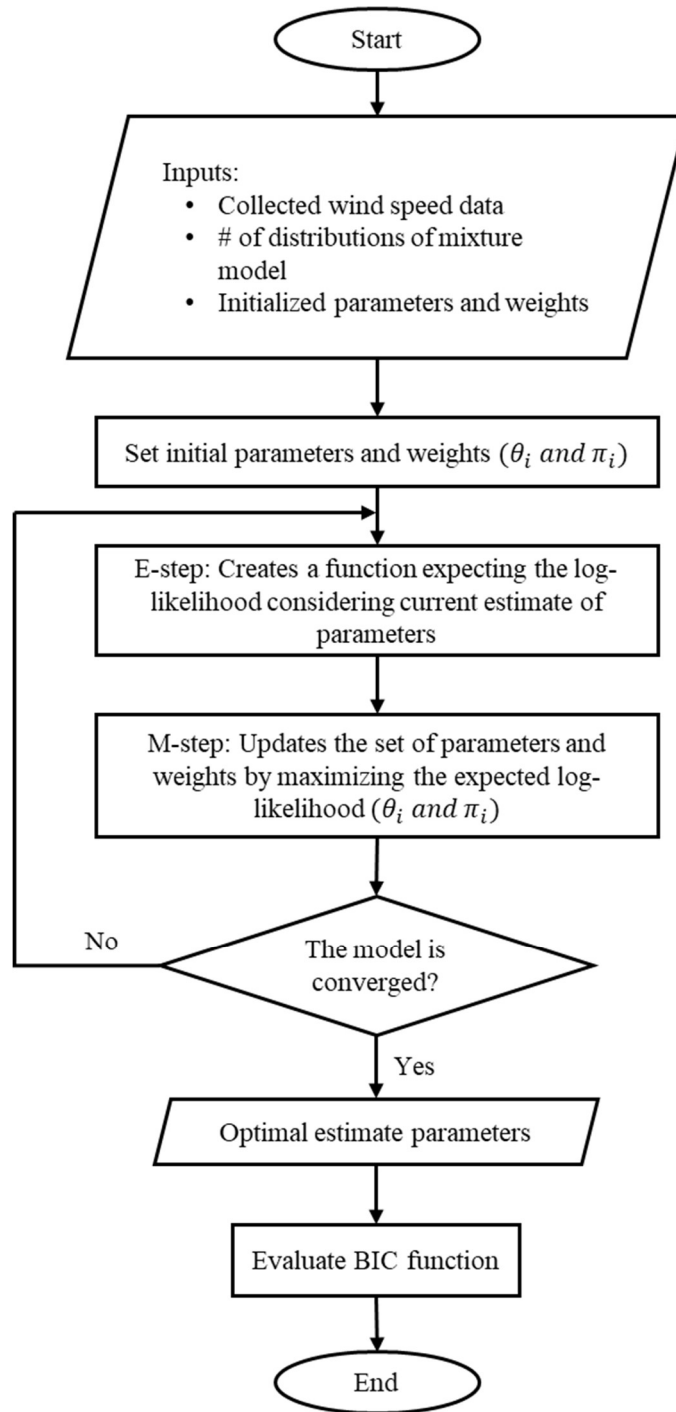


Figure 3.9: The flowchart of the EM algorithm

Sometimes M-step of the EM algorithm may not have the optimal solution. If the M-step of the EM algorithm can not assess the optimal solution, an iterative method such

as Newton-Raphson (NR) or bisection can be used [52]. In which the Newton-Raphson (NR) is the iterative root-finding procedure to find an optimal solution [54].

### 3.3.6 Proper multimodal probability density function (PDF) in wind modeling

For analyzing the performance of the mixture probability density function, three different mixture PDFs are analyzed. Each of the proposed mixture models is divided into clusters. In which each cluster represents a PDF of hourly wind speeds with different parameters. In this analysis, 576 hourly wind speed data have been collected for the whole year. The four components, Mixture of Weibull (MoW), Mixture of Normal (MoN), and Mixture of Gamma (MoG), are considered in this analysis. The Density function of MoW, MoN, and MoG are as follows:

#### A. The mixture of Weibull distribution (MoW):

Assuming that the wind speed is a random variable and follows MoW, the corresponding PDF is given in [29]:

$$f(x; \Phi) = \sum_{i=1}^M \pi_i f_i(x; k_i, \lambda_i) = \sum_{i=1}^M \pi_i \frac{k_i}{\lambda_i} \left(\frac{x}{\lambda_i}\right)^{k_i-1} e^{-\left(\frac{x}{\lambda_i}\right)^{k_i}} \quad (3.16)$$

where the  $\Phi = \{\pi_1, \pi_2, \dots, \pi_M, k_1, k_2, \dots, k_M, \lambda_1, \lambda_2, \dots, \lambda_M\}$  is the parameter set,  $\pi_i$  is the proportion of  $i_{th}$  component, and  $k_i$  and  $\lambda_i$  are parameters to set the shape and scale of the  $i_{th}$  Weibull distribution.

#### B. The mixture of Normal distribution (MoN):

Assuming that the wind speed is a random variable and follows MoN, the corresponding PDF is expressed as follows:

$$f(x; \Phi) = \sum_{i=1}^M \pi_i f_i(x; \mu_i, \sigma_i) = \sum_{i=1}^M \pi_i \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.17)$$

where the  $\Phi = \{\pi_1, \pi_2, \dots, \pi_M, \mu_1, \mu_2, \dots, \mu_M, \sigma_1, \sigma_2, \dots, \sigma_M\}$  is the parameter set,  $\pi_i$  is the proportion of  $i_{th}$  component, and  $\mu_i$  and  $\sigma_i$  are parameters to set the mean and variance of the  $i_{th}$  Normal distribution.

### C. The mixture of Gamma distribution (MoG):

Assuming that the wind speed is a random variable and follows MoG, the corresponding PDF is given in [29]:

$$f(x; \Phi) = \sum_{i=1}^M \pi_i f_i(x; \alpha_i, \theta_i) = \sum_{i=1}^M \pi_i \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} \quad (3.18)$$

where the  $\Phi = \{\pi_1, \pi_2, \dots, \pi_M, \alpha_1, \alpha_2, \dots, \alpha_M, \theta_1, \theta_2, \dots, \theta_M\}$  is the parameter set,  $\pi_i$  is the proportion of  $i_{th}$  component, and  $\alpha_i$  and  $\theta_i$  are parameters to set the shape and scale of the  $i_{th}$  Normal distribution.

Considering both Unimodal and Multimodal PDFs, we estimate wind speed probabilities for the two locations under study. As the purpose of this chapter is to generate an accurate set of probabilities using a specific probability density function, the wind turbine output power and modeling of renewable resources and load data will be represented in Section 3.4 to clarify how different wind states are differentiated.

## 3.4 Wind Turbine Output Power & Modeling of Renewable Resources and Load Data

Wind turbines convert wind energy into mechanical energy, and in the final step into electrical energy. Wind energy is not constant, and the turbine output is proportional to the cube of wind speed. The generated power of the wind turbine generator (WTG) fluctuates. If the capacity ratio of the power source for WTG is tiny, the power source does not cause the frequency to fluctuate by output fluctuation. Nevertheless, if the ratio of WTG capacity is large, the frequency fluctuation of the power system will increase [55]. The rated power of the wind turbine is the maximum power allowed for the installed generator and the control system. The fact that the rated power must not exceed the in high winds. Three primary factors control the power output of the wind turbine:

- 1) Power output curve (determined by aerodynamic power efficiency, mechanical transmission efficiency, and converting electricity efficiency) of the chosen wind turbine;
- 2) Wind speed distribution of the selected site; and
- 3) Tower height.

The provided data related to the wind turbines are the rated power, cut-in speed, rated speed, and cut-out speed. From these statements, the power output curve and output power at any wind speed can be estimated. Figure 3.10 illustrates the power output curve, including cut-in speed, rated speed, and cut-out speed.

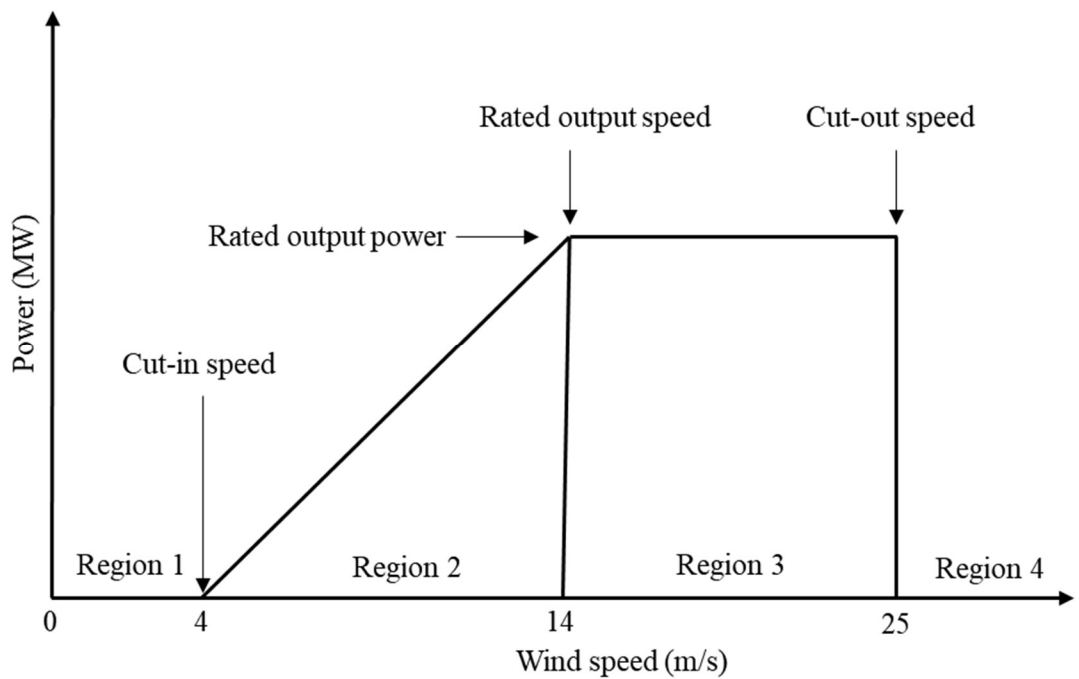


Figure 3.10: Wind turbine power curve

In the first region, when the wind speed is less than a threshold minimum, known as the cut-in speed, the power output is zero. In the second region between the cut-in and the rated speed, there is a rapid growth of power produced. In the third region, a constant (rated) output power is produced until the cut-out speed is attained. Beyond 25 m/s wind speed (region 4), the turbine is taken out of operation to protect its components from high winds; hence it produces zero power in this region.

In Chapter 4, In order to combine the output of the wind-based DGs as a multi-state variable in the planning formulation, the continuous probability density functions (Weibull, Johnson SB, MoW, MoN, and MoG) have been divided into states in which the wind speed is considered in the particular limits. This analysis is applied to the Windsor region as a case study. Table 3.2 represents the selected wind speed states:

Table 3.2: Selected wind speed states

Wind speed state (w)	Wind speed limit (m/s)
1	0-1
2	1-2
.	.
.	.
.	.
Final state	$v_{max} - 1$ to $v_{max}$

The probability of each state  $P(G_w)$  is calculated considering (3.19):

$$P(G_w) = \int_{v_{w1}}^{v_{w2}} f(v)dv \quad (3.19)$$

where  $v_{w1}$  and  $v_{w2}$  are the speed limits of state  $w$ . The wind turbine output power that corresponds to each state is estimated, considering the wind turbine power curve parameters, as calculated by:

$$P_{\frac{o}{pw}}(v) = \begin{cases} 0, & 0 \leq v_{aw} \leq v_{ci} \\ P_{rated} \times \frac{(v_{aw} - v_{ci})}{(v_r - v_{ci})}, & v_{ci} \leq v_{aw} \leq v_r \\ P_{rated}, & v_r \leq v_{aw} \leq v_{co} \\ 0, & v_{co} \leq v_{aw} \end{cases} \quad (3.20)$$



where  $v_{ci}$ ,  $v_r$  and  $v_{co}$  are the cut-in speed, rated speed, and cut-off speed of the wind turbine, respectively. For the sake of simplicity, the average value of each state is utilized to calculate the output power of that state [1]. The capacity factor (CF) of any wind turbine can be estimated by (3.21).

$$CF = \frac{P_{ave}}{P_{rated}} \quad (3.21)$$

### 3.5 Load Modeling

In order to make a proper decision for the planning problem, the system peak load will be assumed to follow the hourly load shape of the IEEE-RTS [56]. Based on this statement, the load will be divided into ten levels considering a clustering technique and utilizing the central centroid sorting process used in [51, 52]. The ten equivalent load levels (states), with different probabilities  $P(L_y)$ , provide an acceptable trade-off between the accuracy and the rapid numerical evaluation [1].

#### 3.5.1 Combined generation load model

The PDFs for the wind and load powers are merged to generate a combined generation load model. Similar to [1], the wind states and load states are considered to be uncorrelated. If a weak correlation exists between wind speed and load, the results will not be affected. Otherwise, the results will be affected because of the correlation's rule, either it is positive or negative. When there is a negative and a positive correlation between wind speed and load, the optimal penetration of DGs will be lower and higher than the value calculated based on the assumption in [1], respectively.

Table 3.3: Load model

State #	% Peak	Probability (%)
1	1	0.01
2	0.853	0.056
3	0.774	0.1057
4	0.713	0.1654
5	0.65	0.1654
6	0.585	0.163
7	0.51	0.163
8	0.451	0.0912
9	0.406	0.0473
10	0.351	0.033

The combination of wind-based DG and load probabilities is given by (3.22):

$$P(C_g) = P_w(G) \times P_y(L) \quad (3.22)$$

The complete generation load model is given by (3.23):

$$R = [\{C_g, P(C_g)\}; g = 1:N] \quad (3.23)$$

where  $m$  is a set of all available turbines in the market, where each turbine has its power performance curve;  $R$  is the entire annual generation load model of  $m$  turbines;  $C$  is combinations of the wind output power states, a matrix of  $m + 1$  columns that includes all possible corresponding to the available turbines, and the load states (i.e. columns from 1 to  $m$  represent the output power of the available  $m$  turbines as a fraction of the rated power of each turbine, whereas column  $m + 1$  represents the different load levels);  $P(C_g)$  is a one-column matrix representing the probability corresponding to matrix  $C$ ; and  $N$  is the total number of states in model  $R$ , which is equal to the product of the wind speed states and the load states.

### 3.6 Computational Results

This section is divided into three parts. In the first part, the results of the unimodal distributions, including p-value, RMSE, and MAE are represented. Then, the Multimodal results are clarified that mixture distribution is the most accurate model in all aspects. In the third part, the best models from both unimodal and multimodal distributions are picked to compare versus each other. After selecting the best model based on the proposed statistical techniques, the wind states and combination wind and load state probabilities will be represented lastly.

#### 3.6.1 Results of unimodality

The p-values obtaining from the K-S test for each unimodal distribution are given in Table 3.4:

Table 3.4: P-value of the Unimodal pdf

<b>Site</b>	<b>Gamma</b>	<b>Rayleigh</b>	<b>Weibull</b>	<b>Johnson SB</b>
<b>Windsor</b>	0.09067	0.00106	0.01144	0.09187
<b>Hamilton</b>	0.02981	0.06584	0.02466	0.37956

Based on the hypothesis testing among the unimodal distributions, it is evident that every distribution except Rayleigh indicates weak evidence against the null hypothesis, which means every distribution except Rayleigh has a P-value greater than 0.01. Then RMSE and MAE are considered to check the accuracy of each PDF. Based on the results in Table 3.5, Johnson SB has the lowest RMSE among other well-known unimodal distributions in wind speed modeling.

Table 3.5: Root mean square error of Unimodal pdfs

<b>Site</b>	<b>Gamma</b>	<b>Rayleigh</b>	<b>Weibull</b>	<b>Johnson SB</b>
<b>Windsor</b>	0.00804	0.00903	0.00872	0.00799
<b>Hamilton</b>	0.00870	0.00836	0.00857	0.00780

Figure 3.11 represents the RMSE of the unimodal distributions in the Windsor region. In the unimodality, Johnson SB, Gamma, Weibull, and Rayleigh distributions are ranked based on lower to higher RMSE, respectively.

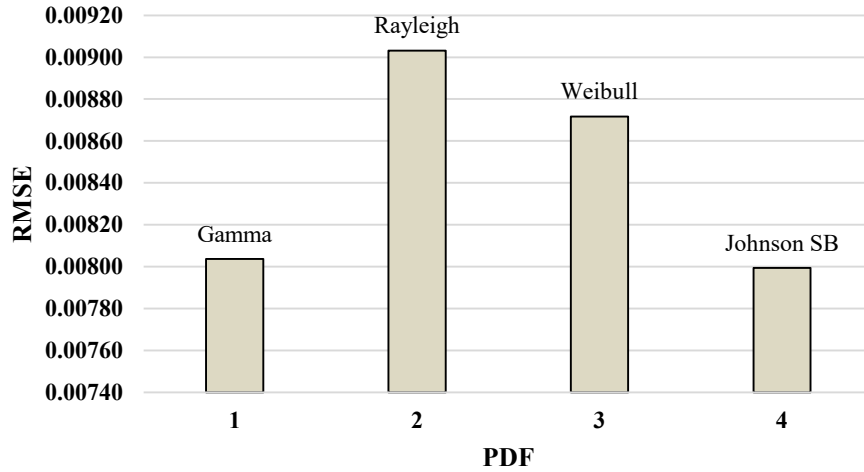


Figure 3.11: RMSE of Unimodal pdfs in Windsor

Figure 3.12 represents the RMSE of the unimodal distributions in the Hamilton region. In the unimodality, Johnson SB, Rayleigh, Weibull, and Gamma distributions are ranked based on lower to higher RMSE, respectively.

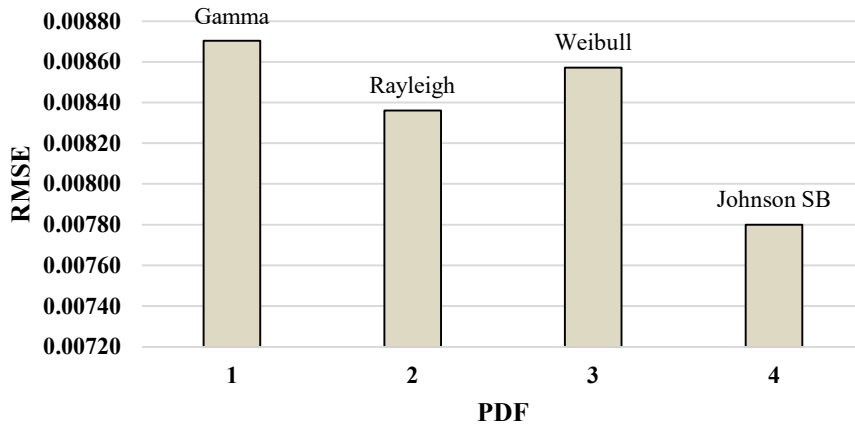


Figure 3.12: RMSE of Unimodal pdfs in Hamilton

Based on the results in Table 3.6, Johnson SB has the lowest MAE among other well-known unimodal distributions in wind speed modeling.

Table 3.6: Mean absolute error of Unimodal pdfs

Site	Gamma	Rayleigh	Weibull	Johnson SB
<b>Windsor</b>	0.00518	0.00572	0.00552	0.00516
<b>Hamilton</b>	0.00654	0.00620	0.00628	0.00585

Figure 3.13 represents the MAE of the unimodal distributions in the Windsor region. In the unimodality, Johnson SB, Gamma, Weibull, and Rayleigh distributions are ranked based on lower to higher MAE, respectively.

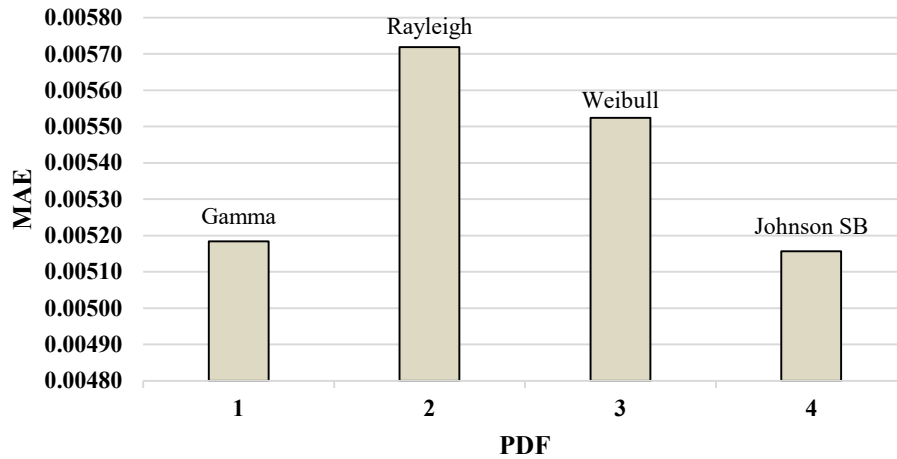


Figure 3.13: MAE of Unimodal pdfs in Windsor

Figure 3.14 represents the MAE of the unimodal distributions in the Hamilton region. In the unimodality, Johnson SB, Rayleigh, Weibull, and Gamma distributions are ranked based on lower to higher MAE, respectively.

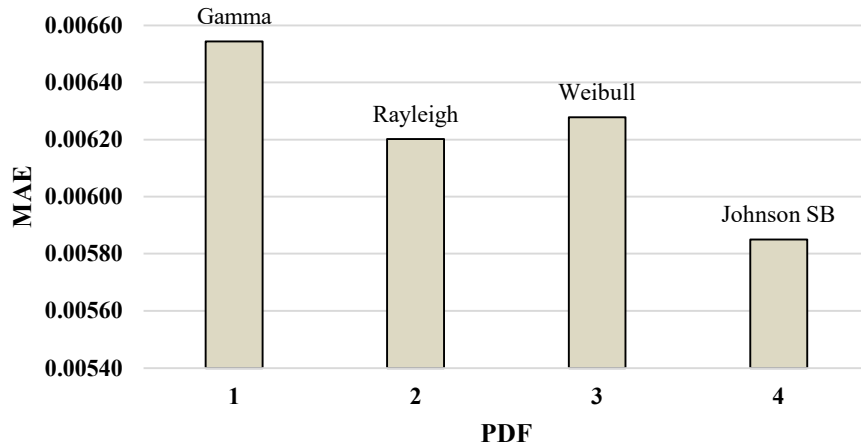


Figure 3.14: MAE of Unimodal pdfs in Hamilton

Comparing the distribution’s RMSE and MAE, Johnson SB demonstrates more accurate results.

As Figure 3.15 demonstrates, the Johnson SB distribution has a slightly better PDF, visually. It means that the estimated probabilities considering the Johnson SB distribution are closer to the observed probabilities than the other distributions.

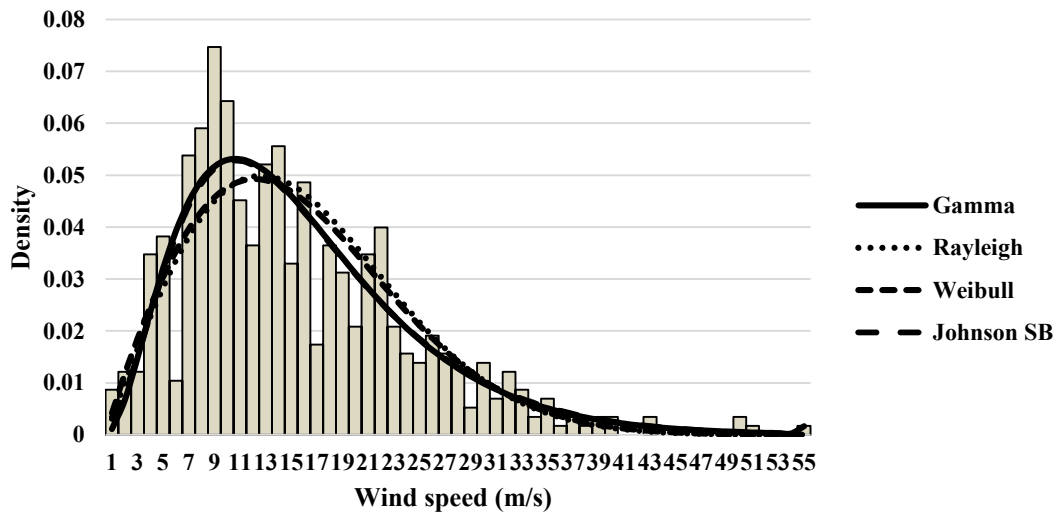


Figure 3.15: Pdfs of unimodal distributions in Windsor

As Figure 3.16 demonstrates, the Johnson SB distribution has a slightly better PDF, visually. It means that the estimated probabilities considering the Johnson SB distribution

are closer to the observed probabilities than the other distributions because of the lower error in RMSE and MAE.

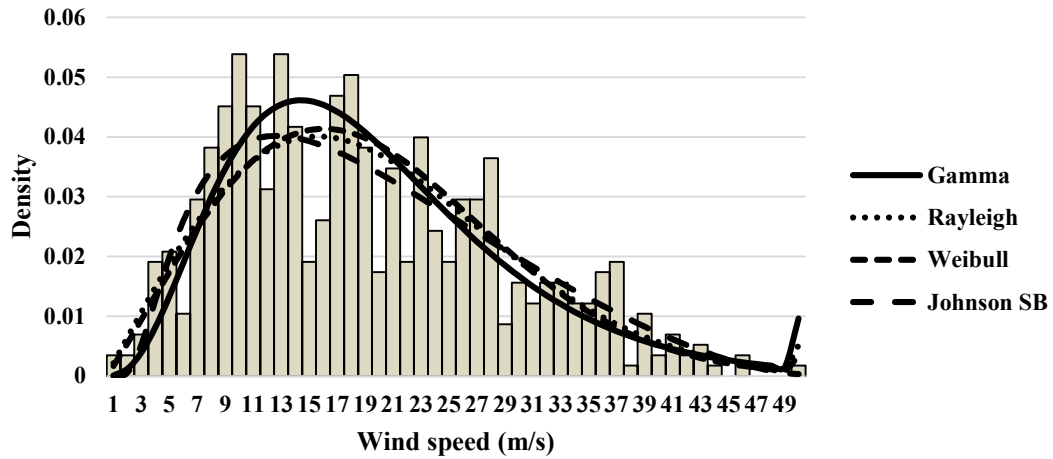


Figure 3.16: Pdfs of unimodal distributions in Hamilton

Based on the different statistical tests such as the goodness of fit test and two types of error measurements we can confidently conclude that the Johnson SB distribution is very proper in the modeling wind speed in all aspects in comparison with a well-known distribution such as Weibull, which is repeatedly used in literature.

### 3.6.2 Results of multimodality

Tables 3.7 and 3.8 illustrate the results of BIC and log-likelihood regarding a different number of components. Based on the analysis, the four components of mixture distributions are taken into consideration as the optimal mixture model because the log-likelihood of the four components distribution is larger than other scenarios.

Table 3.7: BIC & Log-likelihood of mixture distributions with different # of parameters

Windsor	1 Comp	2 Comp	3 Comp	4 Comp
<b>Gamma, Log-Likelihood</b>	-1978.717	-1978.07	-1977.742	-1971.19
<b>Gamma, BIC</b>	3970.128	3987.878	4006.262	4012.202
<b>Normal, Log-Likelihood</b>	-2048.679	-1992.953	-1981.845	-1977.737
<b>Normal, BIC</b>	4110.054	4017.643	4014.469	4025.296
<b>Weibull, Log-Likelihood</b>	-1985.597	-1979.361	-1976.445	-1974.19
<b>Weibull, BIC</b>	3983.889	3990.459	4003.67	4018.2

Table 3.8: BIC & Log-likelihood of mixture distributions with different # of parameters

Hamilton	1 Comp	2 Comp	3 Comp	4 Comp
<b>Gamma, Log-Likelihood</b>	-2087.709	-2077.264	-2075.169	-2073.65
<b>Gamma, BIC</b>	4188.123	4186.29	4201.16	4217.18
<b>Normal, Log-Likelihood</b>	-2120.806	-2081.98	-2076.489	-2076.15
<b>Normal, BIC</b>	4254.317	4195.724	4203.798	4222.18
<b>Weibull, Log-Likelihood</b>	-2084.307	-2076.477	-2074.981	-2074.593
<b>Weibull, BIC</b>	4181.32	4184.716	4200.783	4219.064

Using the simulation in R, the optimal parameters and weights of the MoW distribution, as well as the number of the iteration, are represented in Tables 3.9:

Table 3.9: Parameters and weights of MoW distributions

Windsor	comp1	comp2	comp3	comp4
$\pi_i$	0.597019	0.336274	0.059702	0.007005
$k_i$	2.490784	3.987939	8.056641	160.0008
$\lambda_i$	10.89286	22.03286	34.69444	51.79767
<b>EM iterations</b>	500			
Hamilton	comp1	comp2	comp3	comp4
$\pi_i$	0.513365	0.335622	0.149274	0.001739
$k_i$	2.627759	5.097876	7.658252	160.0008
$\lambda_i$	12.68394	24.45766	36.1803	49.3872
<b>EM iterations</b>	500			

Table 3.10 shows the optimal parameters and weights and the number of iteration of the MoN distribution.

Table 3.10: Parameters and weights of MoN distributions

Windsor	comp1	comp2	comp3	comp4
$\pi_i$	0.623315	0.285651	0.083654	0.00738
$\mu_i$	9.516896	20.35534	31.89631	50.73769
$\sigma_i$	3.908009	3.908009	3.908009	3.908009
<b>EM iterations</b>	236			
Hamilton	comp1	comp2	comp3	comp4
$\pi_i$	0.403646	0.436466	0.149074	0.010814
$\mu_i$	9.725939	20.89811	33.76235	38.18064
$\sigma_i$	3.593326	5.182815	4.866548	7.043399
<b>EM iterations</b>	500			



Table 3.11 shows the optimal parameters and weights and the number of iteration of the MoG distribution.

Table 3.11: Parameters and weights of MoG distributions

<b>Windsor</b>	<b>comp1</b>	<b>comp2</b>	<b>comp3</b>	<b>comp4</b>
$\pi_i$	0.09584	0.433713	0.464291	0.006157
$\alpha_i$	6.365918	10.1588	8.950586	320.0004
$\theta_i$	0.613894	0.981825	2.366035	0.159431
<b>EM iterations</b>		500		
<b>Hamilton</b>	<b>comp1</b>	<b>comp2</b>	<b>comp3</b>	<b>comp4</b>
$\pi_i$	0.556488	0.08774	0.204782	0.15099
$\alpha_i$	4.418604	320.0004	59.37661	53.75205
$\theta_i$	2.721127	0.054354	0.414624	0.647341
<b>EM iterations</b>		247		

Using the parameters of the mixture model, we will be able to estimate each model's probabilities and two types of error measurements. The results, including p-value, root mean square error (RMSE), and mean absolute error (MAE), are evaluated as follows: Table 3.12 demonstrates the p-value of the multimodal distributions.

Table 3.12: P-value of the Multimodal pdfs

<b>Site</b>	<b>MoN</b>	<b>MoW</b>	<b>MoG</b>
<b>Windsor</b>	0	0.02000	0.78
<b>Hamilton</b>	0	0.01000	0.01

Based on the hypothesis testing among the multimodal distributions, it is evident that every distribution except MoN indicates weak evidence against the null hypothesis, which means every distribution except MoN has a P-value greater than 0.01. then RMSE and MAE are considered to check the accuracy of each PDF. Based on the results in Table 3.13, MoG has the lowest RMSE and MAE among other well-known multimodal distributions used in this research.

Table 3.13: Root mean square error of Multimodal pdfs

<b>Site</b>	<b>MoN</b>	<b>MoW</b>	<b>MoG</b>
<b>Windsor</b>	0.007637	0.007701	0.007296
<b>Hamilton</b>	0.007582	0.007556	0.006679

Figure 3.17 represents the RMSE of the multimodal distributions in the Windsor region. In the multimodality, MoG, MoN, and MoW distributions are ranked based on lower to higher RMSE, respectively.

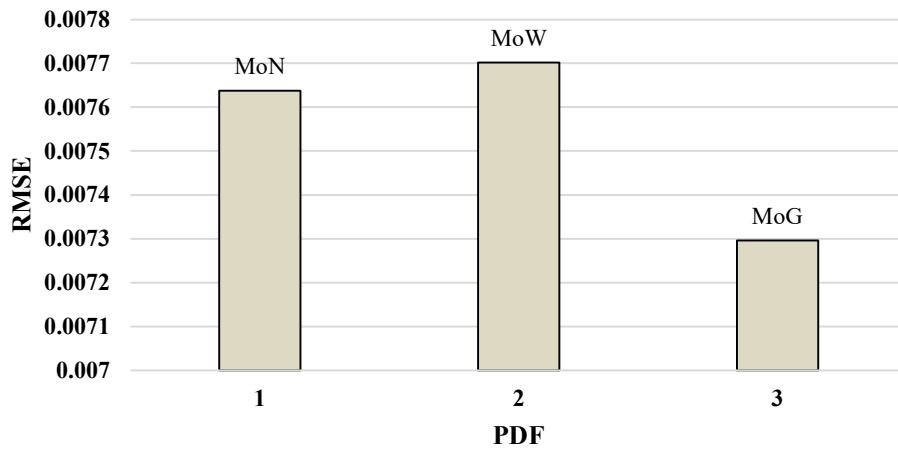


Figure 3.17: RMSE of Multimodal pdfs in Windsor

Figure 3.18 represents the RMSE of the multimodal distributions in the Hamilton region. In the multimodality, MoG, MoW, and MoN distributions are ranked based on lower to higher RMSE, respectively.

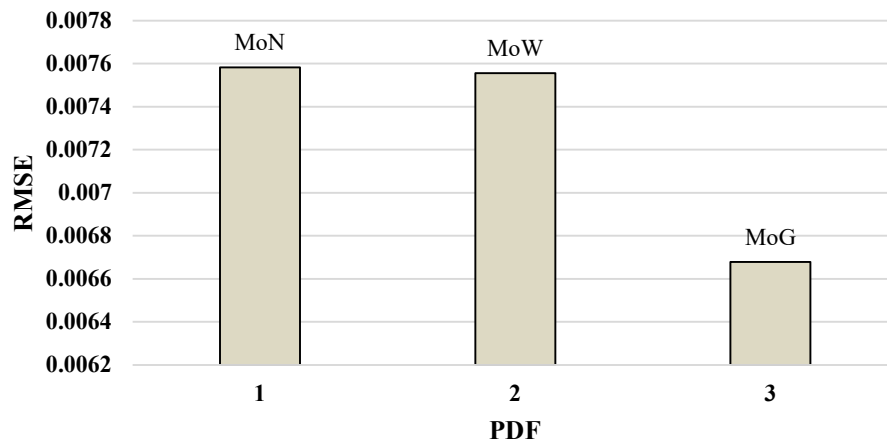


Figure 3.18: RMSE of Multimodal pdfs in Hamilton

Based on the results in Table 3.14, MoG has the lowest MAE among other well-known multimodal distributions in wind speed modeling.

Table 3.14: Mean absolute error of Multimodal pdfs

Site	MoN	MoW	MoG
<b>Windsor</b>	0.004988	0.005011	0.004972
<b>Hamilton</b>	0.005418	0.005412	0.004837

Figure 3.19 represents the MAE of the multimodal distributions in the Windsor region. In the multimodality, MoG, MoN, and MoW distributions are ranked based on lower to higher MAE, respectively.

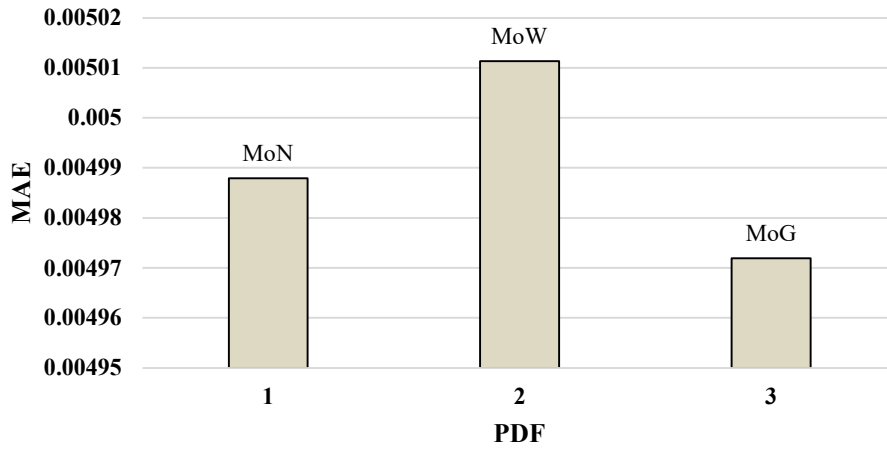


Figure 3.19: MAE of Multimodal pdfs in Windsor

Figure 3.20 represents the MAE of the multimodal distributions in the Hamilton region. In the multimodality, MoG, MoW, and MoN distributions are ranked based on lower to higher MAE, respectively.

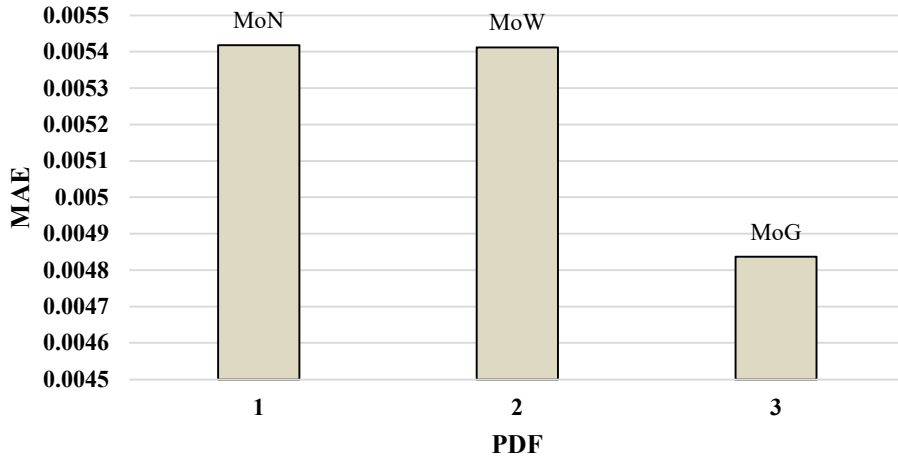


Figure 3.20: MAE of Multimodal pdfs in Hamilton

Comparing the distribution’s RMSE and MAE, MoG demonstrates more accurate results. As Figure 3.21 demonstrates, the MoG distribution has a more accurate PDF, visually. It means that the estimated probabilities considering the MoG distribution are closer to the observed probabilities than the other distributions.

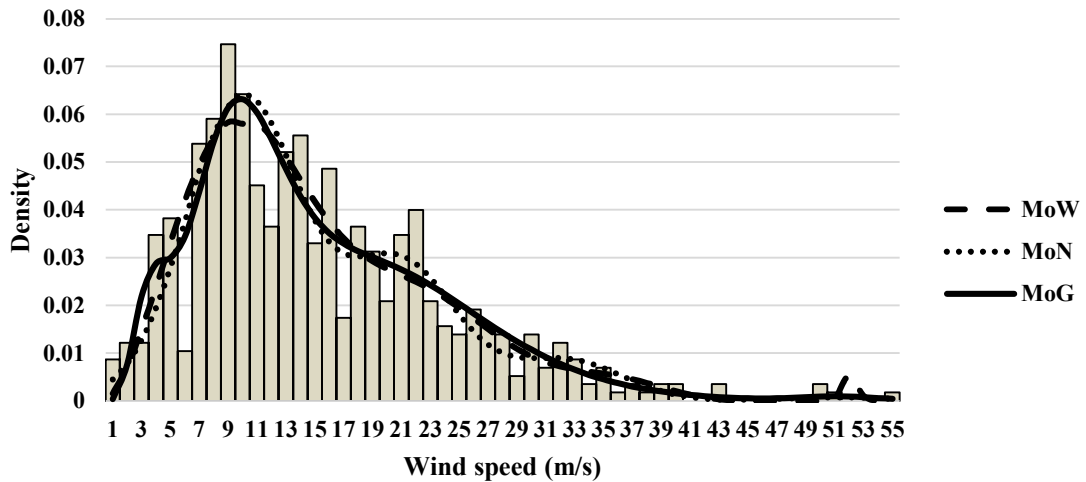


Figure 3.21: Pdfs of Multimodal distributions in Windsor

As Figure 3.22 demonstrates, the MoG distribution has a more accurate PDF, visually. It means that the estimated probabilities considering the MoG distribution are closer to the observed probabilities than the other distributions because of the lower error in RMSE and MAE.

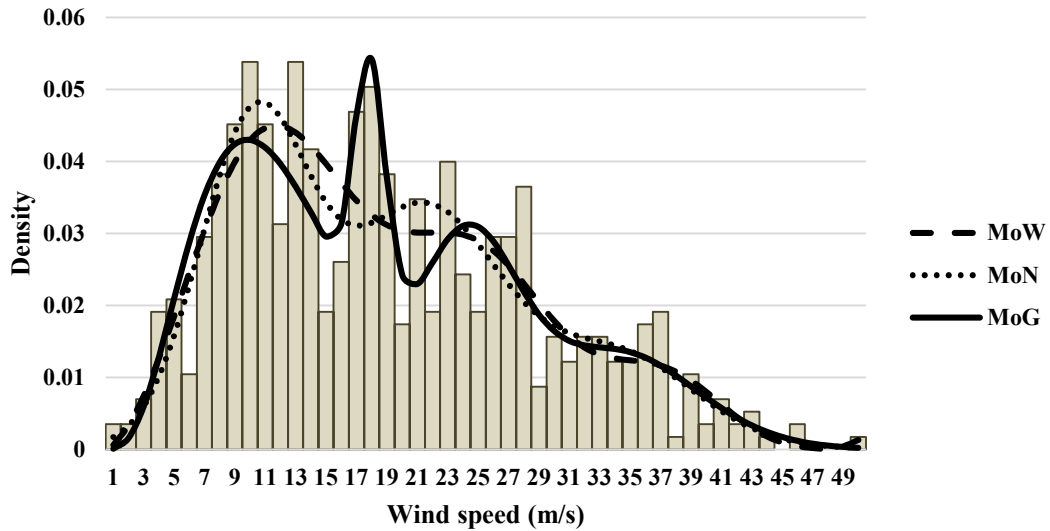


Figure 3.22: Pdfs of Multimodal distributions in Hamilton

### 3.6.3 Best candidate distributions

Based on the results, the Johnson Sb is selected as the most accurate unimodal distribution. On the other hand, the MoG is selected from the multimodal distribution as the best multimodal distribution. Table 3.15 demonstrates the RMSE for MoG and Johnson SB distributions. Based on the computational results in Table 3.15, We can confidently conclude that the MoG represents the most accurate wind speed model in terms of RMSE.

Table 3.15: RMSE of two candidate distributions

Site	Johnson SB	MoG
<b>Windsor</b>	0.00799	0.007296
<b>Hamilton</b>	0.00780	0.006679

Figure 3.23 represents the RMSE of the Johnson SB and MoG distributions in the Windsor region. Results demonstrate the better performance of the MoG versus Johnson SB distribution.

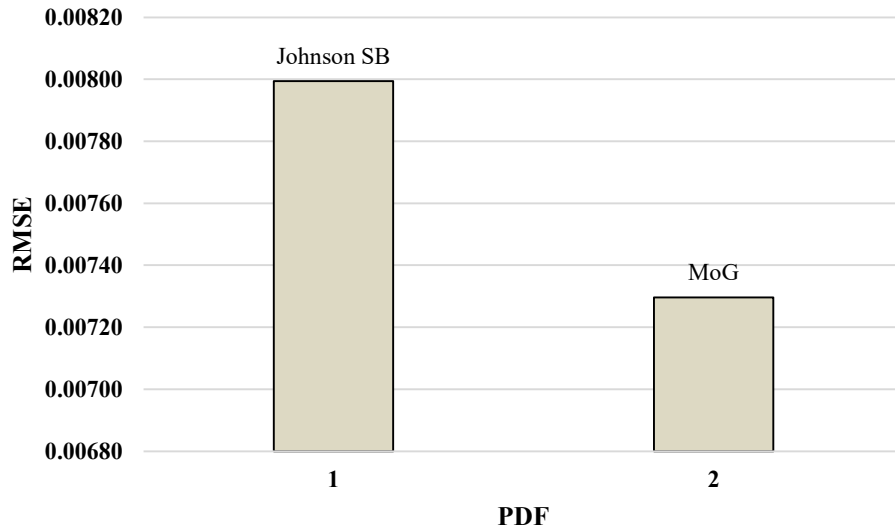


Figure 3.23: RMSE of candidate pdfs in Windsor

Figure 3.24 represents the RMSE of the Johnson SB and MoG distributions in the Hamilton region. Results demonstrate the better performance of the MoG versus Johnson SB distribution.

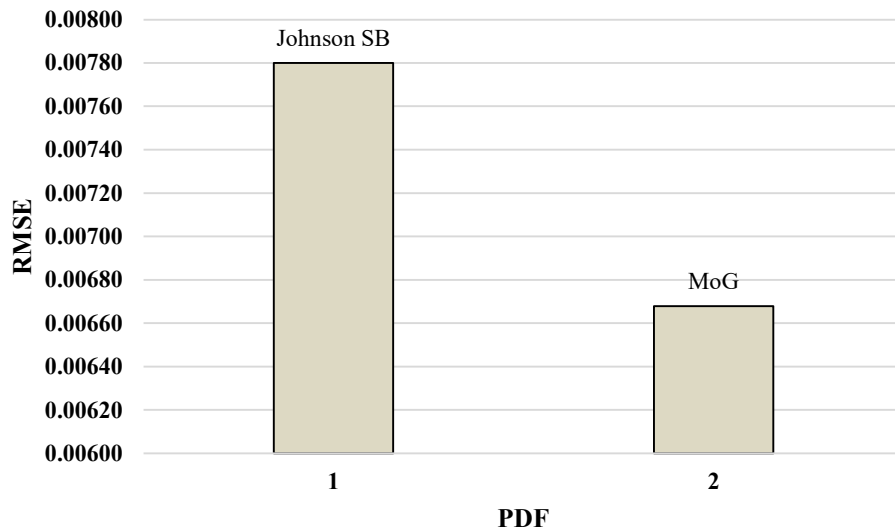


Figure 3.24: RMSE of candidate pdfs in Hamilton

Based on the results, the Johnson Sb is selected as the most accurate unimodal distribution. On the other hand, the MoG is selected from the multimodal distribution as the best multimodal distribution. Table 3.16 demonstrates the MAE for MoG and Johnson

SB distributions. Based on the computational results in Table 3.16, We can confidently conclude that the MoG represents the most accurate wind speed model in terms of MAE.

Table 3.16: MAE of two candidate distributions

Site	Johnson SB	MoG
Windsor	0.00516	0.004972
Hamilton	0.00585	0.004837

Figure 3.25 represents the MAE of the Johnson SB and MoG distributions in the Windsor region. Results demonstrate the better performance of the MoG versus Johnson SB distribution.

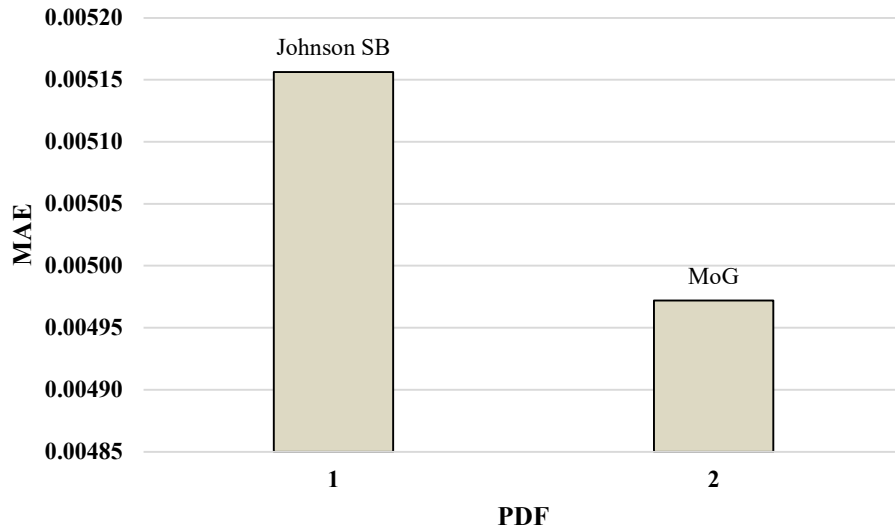


Figure 3.25: MAE of candidate pdfs in Windsor

Figure 3.26 represents the MAE of the Johnson SB and MoG distributions in the Hamilton region. Results demonstrate the better performance of the MoG versus Johnson SB distribution.

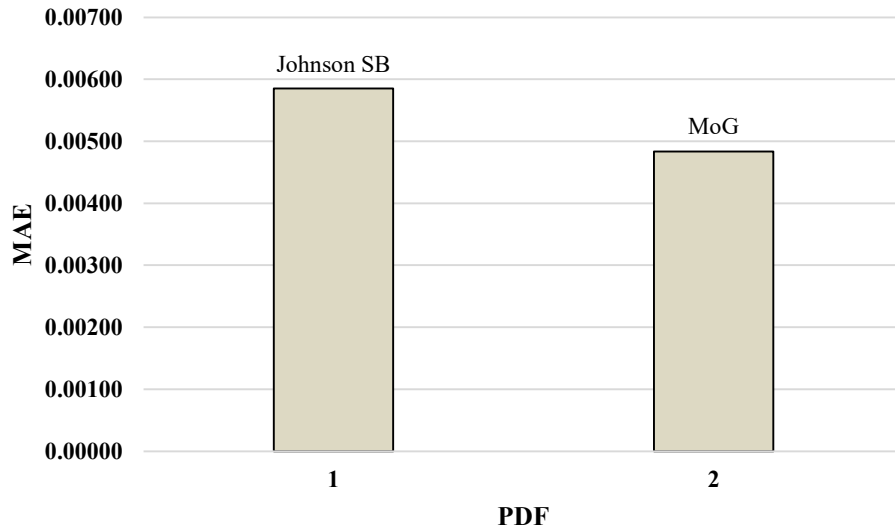


Figure 3.26: MAE of candidate pdfs in Hamilton

Comparing the distributions RMSE and MAE, MoG demonstrates more accurate results. As Figure 3.27 demonstrates, the MoG distribution has a more accurate PDF, visually. It means that the estimated probabilities considering the MoG distribution are closer to the observed probabilities than the other distributions.

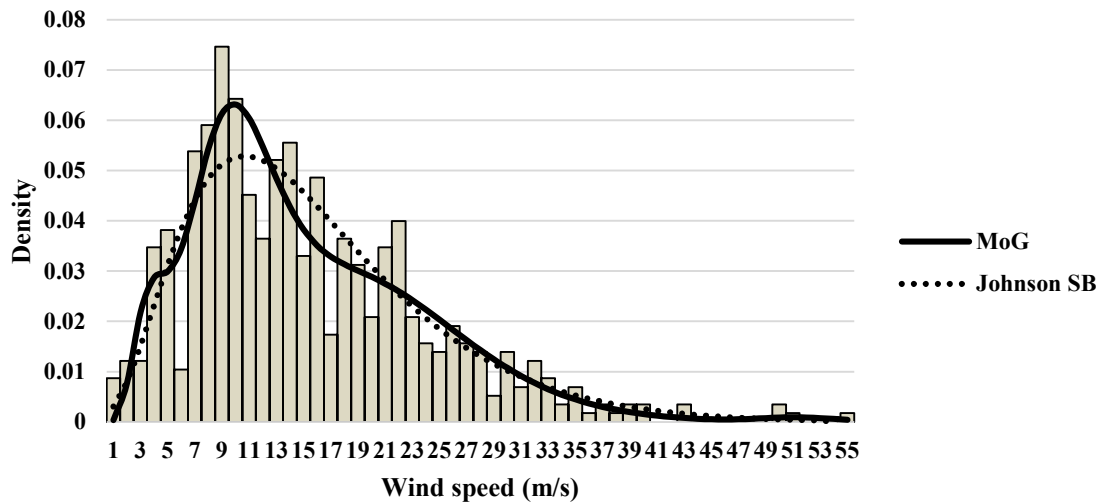


Figure 3.27: Pdfs of candidate distributions in Windsor



As Figure 3.28 demonstrates, the MoG distribution has a more accurate PDF, visually. It means that the estimated probabilities considering the MoG distribution are closer to the observed probabilities than the other distributions because of the lower error in RMSE and MAE.

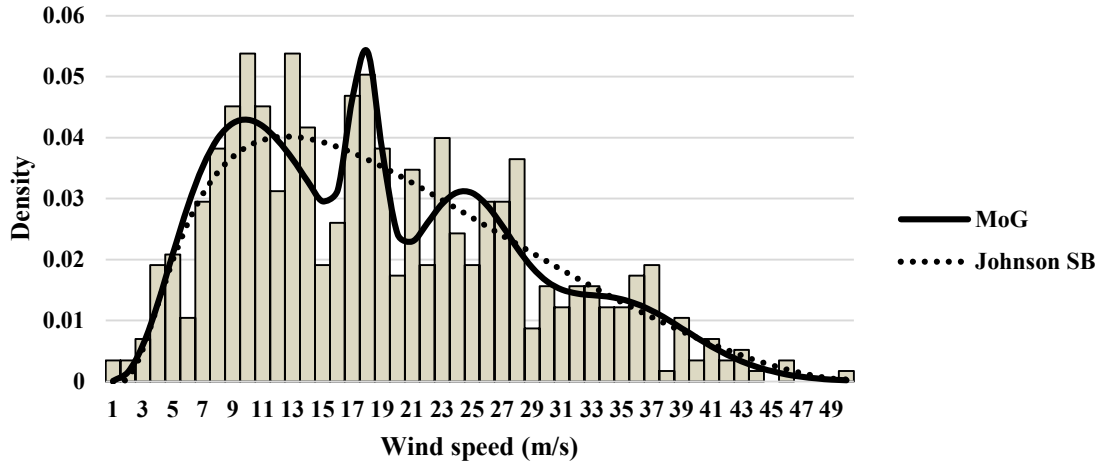


Figure 3.28: Pdfs of candidate distributions in Hamilton

In conclusion, we can confidently conclude that the MoG distribution is very proper in modeling the wind speed in all aspects compared to other candidate distributions in this research.

### 3.7 Combination of the Wind and Load State Probabilities

Based on the results from Section 3.6, the Johnson SB and MoG distributions are represented more significant results in the modeling of wind speed. In this section, the results, including the probabilities of wind states, and the combination of wind and load states for the Johnson SB, Weibull, MoW, MoN, and MoG distributions in the Windsor region, will be illustrated as follows. In Chapter 4, these probabilities will be applied in the wind-based planning problem to compare each distribution's performance.

Table 3.17 demonstrates the wind state probabilities of the Johnson SB distribution. Wind speed range shows the wind speed between a particular range used to estimate each state's probability. For every individual state, there is a probability evaluated using Johnson

SB distribution. Furthermore, the percentage of peak power generated for each state is given.

Table 3.17: The wind state probabilities of Johnson SB distribution

Wind speed range	Johnson SB	% of peak power generated
0-4	0.04965	0
4-5	0.03128	54.9824
5-6	0.03842	164.9472
6-7	0.04422	219.9296
7-8	0.0485	384.8768
8-9	0.05126	494.8416
9-10	0.05263	604.8064
10-11	0.05281	714.7712
11-12	0.052	824.736
12-13	0.05042	934.7008
13-14	0.04827	1044.6656
14-25	0.34603	1100
25-INF	0.13451	0

Table 3.18 demonstrates the wind state probabilities of the Weibull distribution. Wind speed range shows the wind speed between a particular range used to estimate each state's probability. For every individual state, there is a probability evaluated using Weibull distribution. Furthermore, the percentage of peak power generated for each state is given.

Table 3.18: The wind state probabilities of Weibull distribution

Wind speed range	Weibull	% of peak power generated
0-4	0.05934	0
4-5	0.03048	54.9824
5-6	0.03542	164.9472
6-7	0.03964	219.9296
7-8	0.04309	384.8768
8-9	0.04577	494.8416
9-10	0.04767	604.8064
10-11	0.04883	714.7712
11-12	0.04927	824.736
12-13	0.04905	934.7008
13-14	0.04824	1044.6656
14-25	0.38115	1100
25-INF	0.12205	0

Table 3.19 demonstrates the wind state probabilities of the MoW distribution. Wind speed range shows the wind speed between a particular range used to estimate each state's probability. For every individual state, there is a probability evaluated using MoW distribution. Furthermore, the percentage of peak power generated for each state is given.

Table 3.19: The wind state probabilities of MoW distribution

Wind speed range	MoW	% of peak power generated
0-4	0.04773269	0
4-5	0.03325095	54.9824
5-6	0.04200385	164.9472
6-7	0.04947644	219.9296
7-8	0.05506471	384.8768
8-9	0.05839305	494.8416
9-10	0.05805954	604.8064
10-11	0.05814849	714.7712
11-12	0.05519651	824.736
12-13	0.05105974	934.7008
13-14	0.046382	1044.6656
14-25	0.31745926	1100
25-INF	0.12777277	0

Table 3.20 demonstrates the wind state probabilities of the MoN distribution. Wind speed range shows the wind speed between a particular range used to estimate each state's probability. For every individual state, there is a probability evaluated using MoN distribution. Furthermore, the percentage of peak power generated for each state is given.

Table 3.20: The wind state probabilities of MoN distribution

Wind speed range	MoN	% of peak power generated
0-4	0.04495796	0
4-5	0.02805424	54.9824
5-6	0.0376305	164.9472
6-7	0.0483385	219.9296
7-8	0.05580299	384.8768
8-9	0.0617395	494.8416
9-10	0.06416761	604.8064
10-11	0.062806	714.7712
11-12	0.05817748	824.736
12-13	0.05147829	934.7008
13-14	0.04423611	1044.6656
14-25	0.31977455	1100
25-INF	0.12283627	0

Table 3.21 demonstrates the wind state probabilities of the MoG distribution. Wind speed range shows the wind speed between a particular range used to estimate each state's probability. For every individual state, there is a probability evaluated using MoG distribution. Furthermore, the percentage of peak power generated for each state is given.

Table 3.21: The wind state probabilities of MoG distribution

<b>Wind speed range</b>	<b>MoG</b>	<b>% of peak power generated</b>
0-4	0.05865764	0
4-5	0.03000198	54.9824
5-6	0.03449242	164.9472
6-7	0.0439791	219.9296
7-8	0.05441987	384.8768
8-9	0.06143488	494.8416
9-10	0.06318297	604.8064
10-11	0.06033005	714.7712
11-12	0.05478325	824.736
12-13	0.04849029	934.7008
13-14	0.04279699	1044.6656
14-25	0.31732506	1100
25-INF	0.1301055	0

After estimating wind state probabilities, we can evaluate the combination of wind and load state probabilities. Table 3.22 demonstrates the combination of the wind and load state probabilities of the Johnson SB probability density function (PDF). Wind states and Johnson SB show the percentage of peak power generated and the wind state probabilities, respectively. The load state probabilities are assumed to be constant so that the combination of wind and load states is expressed by multiplying the probabilities of Johnson SB PDF with load probabilities.

Table 3.22: The combination of wind and load state probabilities of Johnson SB distribution

<b>Wind states</b>	<b>Johnson SB</b>	<b>Load probabilities</b>	<b>Combined probabilities wind and load</b>
0	0.13451	0.01	0.0013451
1100	0.34603	0.01	0.0034603
1044.6656	0.04827	0.01	0.0004827
934.7008	0.05042	0.01	0.0005042
824.736	0.052	0.01	0.00052
714.7712	0.05281	0.01	0.0005281
604.8064	0.05263	0.01	0.0005263
494.8416	0.05126	0.01	0.0005126
384.8768	0.0485	0.01	0.000485
219.9296	0.04422	0.01	0.0004422
164.9472	0.03842	0.01	0.0003842
54.9824	0.03128	0.01	0.0003128
0	0.04965	0.01	0.0004965

Table 3.23 demonstrates the combination of the wind and load state probabilities of the Weibull probability density function (PDF). Wind states and Weibull show the percentage of peak power generated and the wind state probabilities, respectively. The load state probabilities are assumed to be constant so that the combination of wind and load states is expressed by multiplying the probabilities of Weibull PDF with load probabilities.

Table 3.23: The combination of wind and load state probabilities of Weibull distribution

<b>Wind states</b>	<b>Weibull</b>	<b>Load probabilities</b>	<b>Combined probabilities wind and load</b>
0	0.12205	0.01	0.0012205
1100	0.38115	0.01	0.0038115
1044.6656	0.04824	0.01	0.0004824
934.7008	0.04905	0.01	0.0004905
824.736	0.04927	0.01	0.0004927
714.7712	0.04883	0.01	0.0004883
604.8064	0.04767	0.01	0.0004767
494.8416	0.04577	0.01	0.0004577
384.8768	0.04309	0.01	0.0004309
219.9296	0.03964	0.01	0.0003964
164.9472	0.03542	0.01	0.0003542
54.9824	0.03048	0.01	0.0003048
0	0.05934	0.01	0.0005934

Table 3.24 demonstrates the combination of the wind and load state probabilities of the MoW probability density function (PDF). Wind states and MoW show the

percentage of peak power generated and the wind state probabilities, respectively. The load state probabilities are assumed to be constant so that the combination of wind and load states is expressed by multiplying the probabilities of MoW PDF with load probabilities.

Table 3.24: The combination of wind and load state probabilities of MoW distribution

<b>Wind states</b>	<b>MoW</b>	<b>Load probabilities</b>	<b>Combined probabilities wind and load</b>
0	0.12777277	0.01	0.001277728
1100	0.31745926	0.01	0.003174593
1044.6656	0.046382	0.01	0.00046382
934.7008	0.05105974	0.01	0.000510597
824.736	0.05519651	0.01	0.000551965
714.7712	0.05814849	0.01	0.000581485
604.8064	0.05805954	0.01	0.000580595
494.8416	0.05839305	0.01	0.00058393
384.8768	0.05506471	0.01	0.000550647
219.9296	0.04947644	0.01	0.000494764
164.9472	0.04200385	0.01	0.000420038
54.9824	0.03325095	0.01	0.000332509
0	0.04773269	0.01	0.000477327

Table 3.25 demonstrates the combination of the wind and load state probabilities of the MoN probability density function (PDF). Wind states and MoN show the percentage of peak power generated and the wind state probabilities, respectively. the combination of wind and load states is expressed by multiplying the probabilities of MoN PDF with load.

Table 3.25: The combination of wind and load state probabilities of MoN distribution

<b>Wind states</b>	<b>MoN</b>	<b>Load probabilities</b>	<b>Combined probabilities wind and load</b>
0	0.12283627	0.01	0.001228363
1100	0.31977455	0.01	0.003197746
1044.6656	0.04423611	0.01	0.000442361
934.7008	0.05147829	0.01	0.000514783
824.736	0.05817748	0.01	0.000581775
714.7712	0.062806	0.01	0.00062806
604.8064	0.06416761	0.01	0.000641676
494.8416	0.0617395	0.01	0.000617395
384.8768	0.05580299	0.01	0.00055803
219.9296	0.0483385	0.01	0.000483385
164.9472	0.0376305	0.01	0.000376305
54.9824	0.02805424	0.01	0.000280542
0	0.04495796	0.01	0.00044958

Table 3.26 demonstrates the combination of the wind and load state probabilities of the MoG probability density function (PDF). Wind states and MoG show the percentage of peak power generated and the wind state probabilities, respectively. The load state probabilities are assumed to be constant so that the combination of wind and load states is expressed by multiplying the probabilities of MoG PDF with load probabilities.

Table 3.26: The combination of wind and load state probabilities of MoG distribution

<b>Wind states</b>	<b>MoG</b>	<b>Load probabilities</b>	<b>Combined probabilities wind and load</b>
0	0.1301055	0.01	0.001301055
1100	0.31732506	0.01	0.003173251
1044.6656	0.04279699	0.01	0.00042797
934.7008	0.04849029	0.01	0.000484903
824.736	0.05478325	0.01	0.000547832
714.7712	0.06033005	0.01	0.000603301
604.8064	0.06318297	0.01	0.00063183
494.8416	0.06143488	0.01	0.000614349
384.8768	0.05441987	0.01	0.000544199
219.9296	0.0439791	0.01	0.000439791
164.9472	0.03449242	0.01	0.000344924
54.9824	0.03000198	0.01	0.00030002
0	0.05865764	0.01	0.000586576

### 3.8 Discussion

In Chapter 3, different approaches for modeling the random behavior of wind speed are presented. In modeling wind speed, both unimodal and multimodal distributions, including Weibull, Rayleigh, Gamma, Johnson SB, MoW, MoN, and MoG, are considered. Different techniques are used to evaluate the probabilities of unimodal and multimodal distributions, including the goodness of fit K-S test and maximum likelihood estimation (MLE) method via the expectation-maximization (EM) algorithm to evaluate the optimal parameters, respectively. The most important advantage of using multimodal distributions is high flexibility in fitting to the the random data, which is highly applicable in wind speed modeling. Based on the investigation, the mixture of Gamma (MoG) distribution is selected as the most accurate statistical model for modeling the annual hourly wind speed profile.

# CHAPTER 4: AN APPLICATION OF THE PROPOSED MIXTURE MODEL IN THE PLANNING OF DISTRIBUTION SYSTEMS

## 4.1 Introduction

It is utterly inevitable that renewable energy resources are a critical part of supplying sustainable energy since it is non-polluting and limitless. One of the cleanest and the most available renewable energy is wind energy, which will be used in this work. The inaccurate siting and sizing of wind-based DG units in the distribution system can have a negative impact on system performance. Some of these negative impacts can occur on the acceptable voltage limit, the capacity of the distribution feeder, and the logical amount of the reverse power flow. Based on these statements, an appropriate allocation of renewable DG units in the distribution system is one of the most influential aspects of siting and sizing the DG units.

This chapter takes an in-depth look into the optimal allocation of the wind-based DG units in distribution generation so that it minimizes the annual energy losses. The methodology is based on a stochastic generation load model that combines all possible operating conditions of wind-based renewable DG units via probabilities and then adapting the stochastic model in the planning problem. A mixed-integer nonlinear programming (MINLP) approach is used to formulate the planning problem, with an objective function that minimizes the system's annual energy losses. The constraints, including the voltage limits, the feeder's capacity, the maximum penetration limit, and the discrete size of the available DG units, are used. This proposed methodology is applied to a typical rural distribution system with different scenarios, including all possible combinations of renewable DG units [1].

This chapter is ordered as follows: The first Section 4.1 explains the introduction to the optimal siting and sizing of the renewable DG units in the distribution system. The next four sections represent the planning problem objective, site matching, the planning problem formulation, and the case study, respectively. In Section 4.6, the data regarding load data will be explained. The wind speed and wind turbine data will be covered in



Section 4.7. Finally, in Section 4.8 and 4.9, optimal allocation of wind-based DGs using genetic algorithm (GA) and the planning problem computational results will be represented.

## 4.2 Planning Problem Objective

Because of the environmental concerns and fuel cost uncertainties related to the use of conventional energy sources, attention has been concentrated on implementing renewable DG units in distribution systems. Therefore, in Canada, based on Ontario's Standard Offer Program (SOP), local distribution companies (LDCs) are required to receive a given percentage of customer-owned wind-based DG units in their system. Consequently, LDCs can use the proposed strategy to select the allocations that would maximize benefits [46].

In general, the benefits of maximization in any common planning problem means minimizing cost while maintaining the performance of the system within acceptable limits. Costs include the following [1]:

Capital cost: Is the cost of the wind turbine's installation in the distribution system. Based on the assumption, the wind-based DGs' capital cost is the customer's responsibility because they are the smaller LDCs that buy their power from larger LDCs.

Running cost (operation and maintenance cost): Similar to the capital cost, operation, and maintenance are the pure responsibilities of the consumers.

Cost of unserved energy due to interruption: Based on the current practice of deploying DG units in distribution systems, this cost shows the impact of renewable DG units on the reliability of such distribution systems. In this regard, the following should be noted:

1. A distribution network is fed from a transmission network, and when the connection to the transmission system is lost, i.e., the distribution network is islanded, all DG units are required to shut down for loss-of-main protection. So, DGs cannot improve the reliability of the supply [59].
2. If islanding is allowed, the system can not rely solely on renewable DG units to supply the island's load. Renewable DG units are characterized by high random

power fluctuation levels that result in power mismatch issues, causing stability problems for the system voltage and frequency [60].

3. In opposition, renewable DG units have the prospective to recover the distribution system supply adequacy by increasing the amount of generated power in the system. Moreover, renewable DG units can correspondingly improve the system's reliability from the viewpoint of relieving substation transformers and main feeders during peak load periods. This relief may spread the usable lifetime of the transformer and decrease the probability of premature failure by reason of overloading. On the other hand, these possible improvements in reliability and capability do not depend on the DG units' location on the feeders and are consequently outside the scope of this research. Thus, for the purposes of this study, it is assumed that the location of renewable DG units on a given feeder has no direct influence on the reliability of the distribution system [46].

Feeder power losses: Network losses are a vital element through the planning horizon for the following reasons:

1. While DG units may unload lines and reduce losses if they are improperly allocated, the reverse power flows from larger DG units can give rise to excessive losses and overheat feeders.
2. Minimizing system power losses can positively impact relieving the feeders, reducing the voltage drop, and improving the voltage profile other environmental and economic benefits.

Hence, based on these considerations, the proposed planning problem's objective is, for all possible operating conditions, to minimize the annual energy losses of the system without violating the system constraints.

### 4.3 Site Matching

Every individual step for choosing the optimum wind turbine for a particular location will be the selection which is based on the capacity factor (CF) of the available wind turbines. The ratio between the average output power and the rated power needs to be considered for calculating the capacity factor. The hourly average output power of a wind turbine is a summation of the power generated at all feasible states for this hour

multiplied by the corresponding probability of each state. As soon as the average output power is estimated for each time slot, the average output power is calculated for the ordinary day in each season. Therefore, the annual average output power [46].

Based on the work in [46], the optimal wind turbine is picked based on the higher CF measures value. On the other hand, it is undeniable that choosing the candidate wind turbine's location by considering the highest CF value may not produce optimal loss minimization from a planning standpoint. Occasionally applying the highest CF limits the search area to a multiple of picked wind turbine rating.

#### 4.4 Planning Problem Formulation

The planning problem formulation developed in [1] is considered. However, a different approach is used to solve the proposed probabilistic formulation. Several PDFs are evaluated and compared in the proposed probabilistic generation load model and separately will be fixed into the deterministic optimal power flow (OPF) equations to analyze the performance of each PDF in annual energy loss minimization. Therefore, for each state, there is an active/reactive power flow equation. In other words, the number of states is equal to the number of active/reactive power flow equations. To calculate and then weight the energy losses, we need the probability of each state happening in the whole year. Although, the penetration of each state changes based on the generation load model. Thus, optimally locating and sizing the DG units will minimize the total energy losses without violating the system constraints.

In order to control the optimum allocation of the wind-based DG units, six scenarios are dictated.

1. In the first scenario, no DG units are connected to the system (base case).
2. The proposed planning problem is formulated based on the wind-based DG units resulted from Johnson SB distribution.
3. The proposed planning problem is formulated based on the wind-based DG units resulted from Weibull distribution.
4. The proposed planning problem is formulated based on the wind-based DG units resulted from a mixture of Normal (MoN) distribution.

5. The proposed planning problem is formulated based on the wind-based DG units resulted from a mixture of Weibull (MoW) distribution.
6. The proposed planning problem is formulated based on the wind-based DG units resulted from a mixture of Gamma (MoG) distribution.

The output power from DG units in Scenarios 2 to 6 depending on the wind probabilities.

#### 4.4.1 Objective function

The objective function is considered to minimize the annual energy losses in the distribution system for all probable combinations of load and wind-based DG output power. Equation (4.1) represents the proposed objective function of the planning problem. The planning problem is formulated as a mixed-integer nonlinear programming (MINLP) on the Matlab environment considering the Genetic Algorithm (GA) toolbox.

$$\min \sum_{g=1}^N p_{lossg} \times p(c_g) \times 8760 \quad (4.1)$$

where  $P_{lossg}$  is the total power losses in the system during state  $g$ , and  $p(c_g)$  is the combination of the wind and load state probabilities.

#### 4.4.2 Constraints

A. Power flow equations:

$$P_{G_1} + \sum_{t=1}^m C(g, t) \times P_{DG_{t,i}} - C(g, m + 1) \times P_{D_i} = \sum_{j=1}^n V_{g,i} \times V_{g,j} \times Y_{ij} \times \cos(\theta_{ij} + \delta_{g,j} - \delta_{g,i}), \quad \forall i \quad (4.2)$$

$$Q_{G_1} - C(g, m + 1) \times Q_{D_i} = - \sum_{j=1}^n V_{g,i} \times V_{g,j} \times Y_{ij} \times \sin(\theta_{ij} + \delta_{g,j} - \delta_{g,i}), \quad \forall i \quad (4.3)$$

where  $P_{DG_{t,i}}$  is the rated power of the  $t$ th wind-based DG connected at bus  $i$ ;  $P_{G_{g,1}}$  is the substation active power injected at bus  $i$ ;  $Q_{G_{g,1}}$  is the substation reactive power injected at

bus  $i$ ;  $P_{Di}$  is the peak active load at bus  $i$ ;  $Q_{Di}$  is the peak reactive load at bus  $i$ ;  $V_{g,i}$  is the voltage at bus  $i$  during state  $g$ ; and  $n$  is the total number of buses in the system.

B. Power losses constraints:

$$P_{loss_g} = 0.5 \times \sum_{i=1}^n \sum_{j=1}^n G_{ij} \times [(V_{g,i})^2 + (V_{g,j})^2 - 2 \times V_{g,i} \times V_{g,j} \times \cos(\delta_{g,j} - \delta_{g,i})], \quad \forall i \quad (4.4)$$

The power loss constraint in (4.4) estimates the power losses in the system caused by transmission through electric conductors where the parameter  $G$  represents the conductance in each branch [1].

C. Branch current equation:

$$I_{g,ij} = |Y_{ij}| * [(V_{g,i})^2 + (V_{g,j})^2 - 2 * V_{g,i} * V_{g,j} * \cos(\delta_{g,j} - \delta_{g,i})]^{\frac{1}{2}} \quad \forall g, i, j \quad (4.5)$$

The feeder constraint estimates the current flowing through the conductors between each pair of buses  $i$  and  $j$  in each state [1].

D. Slack bus voltage and angle (assumed to be bus one)

$$V_{g,1} = 1.025 \quad \& \quad \delta_{g,1} = 0.0 \quad (4.6)$$

E. Voltage limits at load buses

$$V_{min} \leq V_{g,i} \leq V_{max} \quad \forall i \notin \text{substation bus} \quad (4.7)$$

The voltage limits guarantee that the voltages remain within acceptable limits ( $0.95 \leq V_i \leq 1.05$ ) to keep appliances connected to the grid safe from damage [1].

F. Feeder capacity limits:

$$0 \leq I_{g,ij} \leq I_{ijmax} \quad \forall i, j, g \quad (4.8)$$

It keeps electrical conductors like transmission and distribution lines harmless from overheating [1].

G. Discrete size of wind-based DG

$$P_{DG_{t,i}} = a_{t,i} \times P_r \quad \forall i \in B \quad t \in m \quad (4.9)$$

where  $a_{t,i}$  is the integer variable;  $P_r$  is the one-column matrix of length  $m$  includes all ratings of the available wind turbines; and  $B$  is the set of candidate buses to connect DGs.

H. Maximum penetration on each bus:

$$\sum_{t=1}^m P_{DG_{t,i}} \leq P_{bus} \quad \forall i \quad (4.10)$$

where,  $P_{bus}$  is the maximum allowable megawatt (MW) at each bus, which is limited to 10 MW; This is based on the current practice in LDC Ontario, Canada [1].

I. Maximum penetration of DG units in the system: The maximum penetration limit will be calculated based on the wind-based DG's average penetration.

$$\sum_{i=1}^n \sum_{t=1}^m CF \times P_{DG_{t,i}} \leq x \times \sum_{i=1}^n P_{Di} \quad (4.11)$$

where  $x$  is the maximum penetration limit that will be 30% of the peak load.

## 4.5 Case Study

This section shows collected wind speed data used for analyzing different PDFs, and wind turbine data for the system under study.

### 4.5.1 System under study

Figure 4.1 illustrates the usual rural distribution system with a peak load of 16.18 MVA. To deliver energy to the rural area, the substation at bus one is considered [1]. There

are candidate buses to connect to the DG units based on the wind pattern and land accessibility in the following set B: { 19, 23, 33, 35, 37, 38, 39, 40 }. Besides, based on the local distribution company's recent decision in Ontario, Canada, the maximum permissible MW at each bus is restricted to 10 MW, and the maximum penetration limit is 30% of the peak load [1].

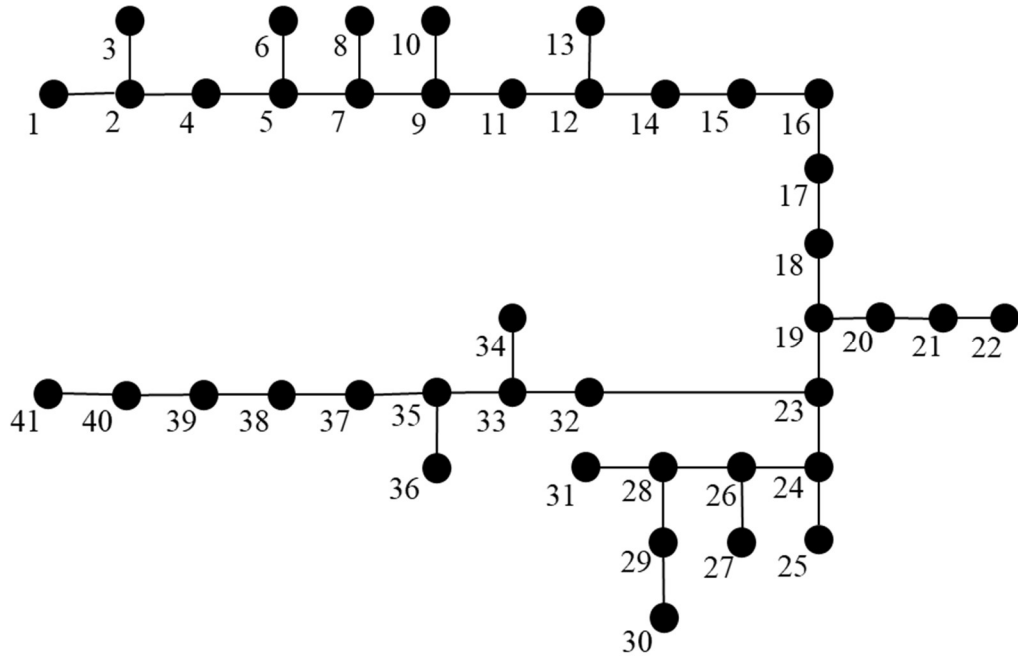


Figure 4.1: System understudy

## 4.6 Load Data

Table 4.1 represents the probabilistic load model, including the percentage of peak load and probability for each state.

Table 4.1: Probabilistic load model

State #	% Peak	Probability (%)
1	1	0.01
2	0.853	0.056
3	0.774	0.1057
4	0.713	0.1654
5	0.65	0.1654
6	0.585	0.163
7	0.51	0.163
8	0.451	0.0912
9	0.406	0.0473
10	0.351	0.033

## 4.7 Wind Speed and Wind Turbine Data

Based on the data gathered from [49], the minimum, mean, and maximum wind speeds are 1 m/s, 15.22 m/s, and 55 m/s, respectively. Furthermore, based on the assumption, only one size of wind turbine is available. The wind turbine data, including rated power, cut-in speed, nominal speed, and cut-off speed, are 1.1 MW, 4 m/s, 14 m/s, and 25 m/s, respectively.

Using the wind turbine data, the probabilistic wind output power model and CF are calculated. The value of the CF is estimated to be around 22.09%. Based on the analysis, the most appropriate probabilistic model designed for the wind speed modeling in this work is the Mixture of Gamma (MoG) probability density function. Tables 4.2 and 4.3 represent the wind speed and wind power probabilities.



Table 4.2: Wind speed probabilities

Wind speed limits, m/s	Probability
0-4	0.058657642
4-5	0.030001983
5-6	0.034492422
6-7	0.043979099
7-8	0.054419874
8-9	0.061434879
9-10	0.063182968
10-11	0.060330052
11-12	0.054783246
12-13	0.048490289
13-14	0.042796988
14-25	0.317325058
> 25	0.1301055

Based on the power curve parameters of the available wind turbine (Figure 3. 10) in Chapter 3, some states are merged together; for instance, wind speed states from 1 m/s to 4 m/s are gathered in one single state. In terms of the number of wind states, 12 states are taken into account, similar to Table 4.2. The last state, which is state 13, shows the wind speed greater than the cut-out wind speed; therefore, no electrical energy will be generated from the wind turbine. Table 4.3 represents the combination of the wind and load model, respectively.

Table 4.3: Wind power probabilities

Wind states	Wind Prob	% of Load states	Probabilities	Combined probabilities wind and load
0	0.1301055	1	0.01	0.001301055
1100	0.31732506	1	0.01	0.003173251
1044.6656	0.04279699	1	0.01	0.00042797
934.7008	0.04849029	1	0.01	0.000484903
824.736	0.05478325	1	0.01	0.000547832
714.7712	0.06033005	1	0.01	0.000603301
604.8064	0.06318297	1	0.01	0.00063183
494.8416	0.06143488	1	0.01	0.000614349
384.8768	0.05441987	1	0.01	0.000544199
219.9296	0.0439791	1	0.01	0.000439791
164.9472	0.03449242	1	0.01	0.000344924
54.9824	0.03000198	1	0.01	0.00030002
0	0.05865764	1	0.01	0.000586576

Combining the 12 states of wind and 10 states of load, we have 120 states in total.

#### 4.8 Optimal Allocation of Wind-based DGs Using Genetic Algorithm (GA)

The Genetic Algorithm (GA) is a metaheuristic search algorithm motivated by natural selection in living organization populations. The GA aims to maintain a population of solutions to a problem as individuals' programmed information gradually progresses [61]. The GA encompasses three different search segments starting by creating an initial population, evaluating a fitness function, and producing a new population [62]:

The GA begins with generating the initial population (chromosomes) randomly in which each individual will be weighted through a fitness function. Based on the fitness value, individual generations are repeated or removed. Then, all the calculated solutions that satisfy the constraints are listed and compared. The solution with the lowest value is

compared with the base case. If the selected solution is lower than the base case, then it will be considered as the optimum solution for the proposed location of DGs. Then GA repeats similar steps to assess the next suggested locations. By comparing all solutions, the best locations can be selected. By applying the GA operators, more generations will be created [39].

The Genetic Algorithm (GA) is designed by three Operators:

- Crossover: the individuals are structured randomly in pairs so that they have their location together. In that way, each pair of individuals produce a new pair.
- Mutation: In this step, some individuals are changed randomly to measure the research space's other area (points).
- Selection: After the crossover and mutation steps, the individuals will be estimated. Moreover, for being inserted in a new population, the greater probability will be given to better individuals based on a probabilistic rule.

The problem at hand is a mixed-integer non-linear programming (MINLP). Therefore, no exact optimal solutions are guaranteed. Moreover, it is computationally intractable. The genetic algorithm (GA) is utilized to solve larger instances of this planning problem. Genetic algorithm (GA) can provide a sub-optimal solution when its parameters are tuned properly. In comparison with other meta-heuristic techniques in terms of solution error and performance time, GA is represented more appropriate performance [44]. The mathematical model in the planning is solved considering the GA toolbox in Matlab. The Matpower power flow simulation toolbox is run within the GA runs and generations to solve for the optimal location and number of DGs, and the energy losses in the distribution system.

Figure 4.2 presents the formation of a chromosome's 8 variables. The flowchart in Figure 4.3 depicts the logical flow of the code developed in Matlab using the GA solver. It shows how the mathematical model is applied and how it is combined with the solution algorithm to provide the results.

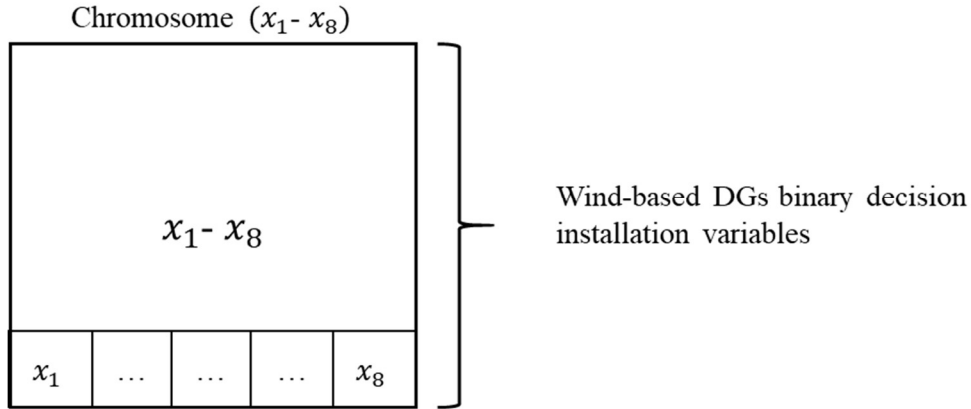


Figure 4.2: The chromosome of the eight decision variables

Before GA starts simulating, several data sets are required. Those data sets are consist of the information related to the distribution network under study, the stochastic components of the combination of the wind and load. Moreover, the constraints consist of the power flow constraints and the wind power distributed generation. In step two, all stochastic components and all 120 states associated with the combination of those states' probabilities are taken into action. The initial values of the decision variables, such as the location and number of wind-based DGs is set in step three.

In step four, the power flow starts simulating with the initial values set in step three to approximate the current set up. In step five, a series of estimations utilize based on the computational results of the simulation. The voltage of the buses and current through lines are saved in matrices. The reactive and active power of each candidate bus is calculated, and also the candidate locations are recorded. Even the transmission losses are taken into account. As the final step, the total generated power from the wind-based DGs is calculated. For each scenario, steps 4 and 5 are repeated 120 times.

In step seven, the objective function is estimated using the gathered data from the 120 states and each scenario's probability. The outcome would be the minimized losses over the entire planning problem. If the simulation does not meet the termination criteria, the decision variables will be adjusted using crossover and mutation rules before step 4 starts over.

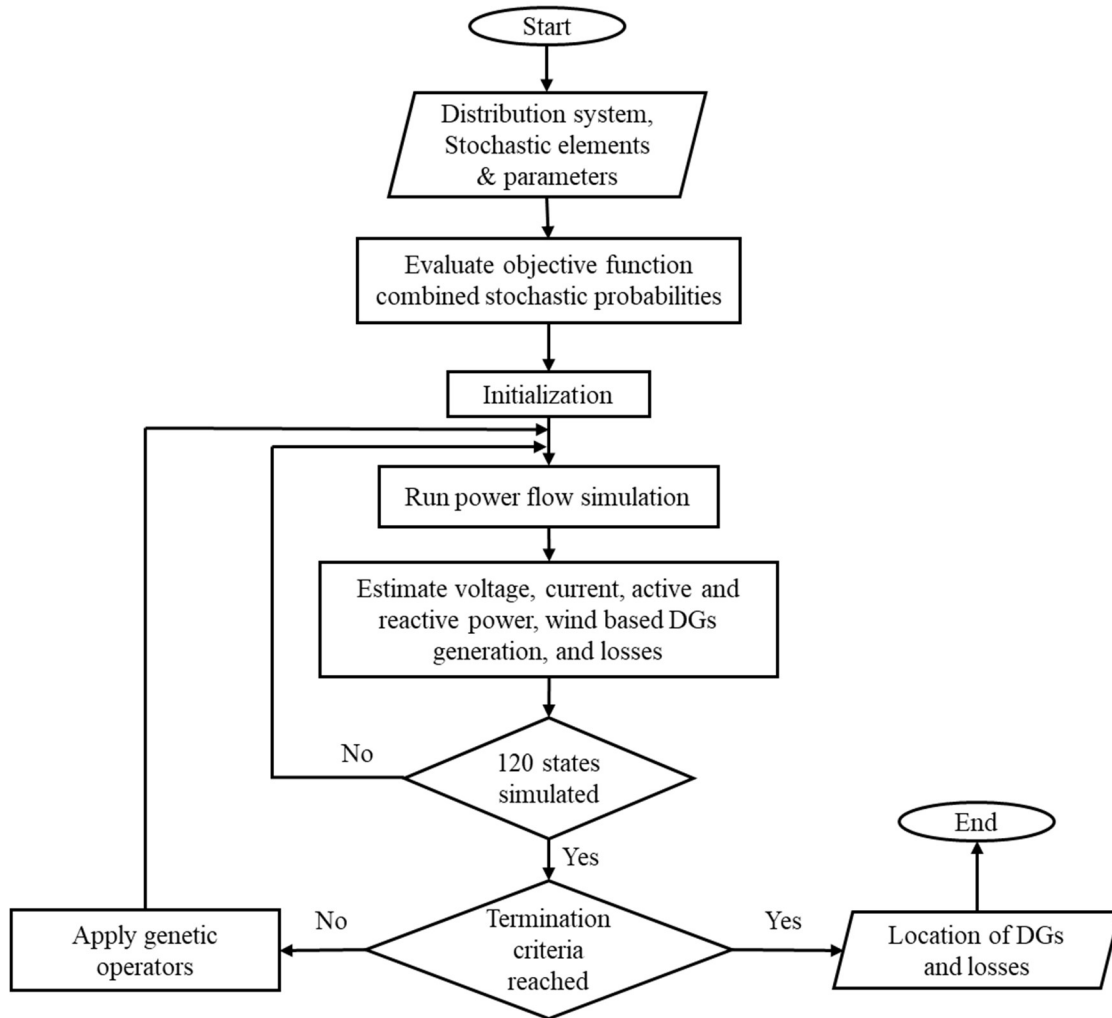


Figure 4.3: Adopted hybridized solution approach

## 4.9 Computational Results

This section analyzes all six scenarios to determine the optimum allocation of wind-based DG units that minimize the system energy losses. The scenario with the estimated loss closer to the actual loss is selected as the most accurate probability modeling. The accurate estimation of wind probability is an essential tool in the planning of the system. The primary purpose of this research is to demonstrate the major difference in results when using a single PDFs versus the newly developed mixture PDFs. Results reveal that using probabilities of MoG in Scenario 6, have a significant improvement in the annual energy losses than other scenarios. In other words, MoG is selected as the most accurate model with a closer estimated loss to its actual loss. Additionally, the actual annual energy losses

of the system are estimated to validate the proposed planning technique. The actual annual energy loss can be calculated for each scenario separately. Table 4.4 shows all 6 scenario's estimated and actual losses.

Table 4.4: Losses and installed wind DGs for each scenario

PDF model	# of wind DGs	Estimated Loss	Actual Loss	Installed DGs
-	0	1177.3	1175.9	DG=[0 0 0 0 0 0 0]
Johnson	5	743.46	763.99	DG=[3 1 0 0 0 1 0]
Weibull	5	737.8	763.99	DG=[3 1 0 0 0 1 0]
MoWeibull	5	743.5	763.88	DG=[3 1 0 0 0 0 1]
MoNormal	5	734.67	763.99	DG=[3 1 0 0 0 1 0]
MoGamma	5	745.55	763.88	DG=[3 1 0 0 0 0 1]

## 4.10 Discussion

In this chapter, a probabilistic planning technique is proposed to optimally allocate wind-based DG units into the distribution system to minimize the annual energy losses. Specifically, this technique is based on generating a stochastic generation-load model that contains all possible operating conditions; hence, this model can be accommodated into a deterministic optimal power flow (OPF) formulation. The random behavior of wind speed is modeled utilizing various distributions (i.e., Gamma, Weibull, Rayleigh, Johnson SB, MoN, MoW, and MoG), respectively. The optimization problem is formulated as mixed-integer non-linear programming (MINLP) under the Matlab environment using the GA toolbox. The system's technical constraints include the voltage limits, the DG's discrete size, and the maximum penetration limit of the DG units. Different scenarios are considered in the proposed planning technique. Based on the results obtained, we can observe that the most energy loss reduction was obtained in the last scenario (scenario 6).

## CHAPTER 5: CONCLUSION AND FUTURE WORK

### 5.1 Conclusion

The stochastic modeling of the wind speed (i.e., comparing the performance of different PDFs versus each other) and optimal allocation of wind-based DG units in the distribution system considering the most accurate PDF from the probabilistic wind modeling, are the two primary objectives in this thesis.

The thesis starts with comparing various unimodal probability density functions (PDFs) using the goodness-of-fit K-S test to model the wind speed. The proposed strategy utilized one-year of wind speed data collected from [49]. In the second step, two well-known error measurement techniques MAE and RMSE are introduced to find the most accurate PDF over other statistical models. In the third step, for modeling of wind speed more accurately, the maximum likelihood estimation (MLE) method for finite mixture modeling via the Expectation-Maximization (EM) algorithm is introduced. The results from the proposed model compared to the rest of the well-known distributions represented that the proposed MoG distribution provides a better fit for the actual data in terms of two types of error measurements. Thus, the MoG distribution should be a new alternative when compared to unimodal distributions such as Weibull, Rayleigh, Gamma, or Johnson SB. Besides, the MoG PDF represents the wind probabilities that provide accurate estimated loss when compared to the actual loss during the planning stages. In conclusion, MoG pdf provides more accurate probabilities for use in stochastic programs and so should be used instead of unimodal PDFs.

### 5.2 Future Work

As the future work, one interesting direction in the modeling of wind speed is to analyze the factors impacting the wind speed such as year, season, weather, temperature, and time. We need to analyze the observation to identify any factor that contributes to the wind speed. Analyzing the different factors of wind speed, we can understand which factors have high fluctuations. For instance, high differences in average wind speed in different seasons demonstrate that if we disregard this main effect, we consider that as an uncontrollable error, which can remarkably increase the total error. By considering factors

such as season in case of fluctuations, we can evaluate both unimodal and multimodal probability density functions based on the influential factors. So there is a high chance to have a more precise wind model if we consider the wind speed factors. Applying these new pdfs to the planning problem might be effective in minimizing the energy losses of renewable Distributed generators (DGs). Another interesting direction is to optimally site and size the mixed renewable (DGs) in the distribution system using multimodal distributions, evaluated based on the influential factors on wind speed.



## APPENDIX A: DISTRIBUTION SYSTEM DATA

Table A- 1: 41-BUS SYSTEM REAL AND REACTIVE LOADS AT EACH BUS

<b>Bus No.</b>	<b>Pd</b>	<b>Qd</b>	<b>Bus No.</b>	<b>Pd</b>	<b>Qd</b>
1	0	0	22	47.5	15.61
2	0	0	23	9.5	3.12
3	0	0	24	0	0
4	6413.46	2108	25	289.75	95.24
5	0	0	26	0	0
6	903.06	511.79	27	152	49.96
7	0	0	28	0	0
8	3187.25	1047.6	29	0	0
9	0	0	30	194.75	64.01
10	576	507.98	31	517.75	170.18
11	0	0	32	0	0
12	0	0	33	0	0
13	19	0	34	204.25	67.13
14	346.75	113.97	35	0	0
15	0	0	36	80.75	26.54
16	0	0	37	104.5	34.34
17	0	0	38	0	0
18	0	0	39	0	0
19	0	0	40	0	0
20	0	0	41	1000	320
21	0	0			

## BIBLIOGRAPHY

- [1] Y. M. Atwa and E. F. El-Saadany, "Probabilistic approach for optimal allocation of wind-based distributed generation in distribution systems," *IET Renew. Power Gener.*, vol. 5, no. 1, pp. 79–88, 2011.
- [2] K. R. Devabalaji, K. Ravi, and D. P. Kothari, "Optimal location and sizing of capacitor placement in radial distribution system using Bacterial Foraging Optimization Algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 71, pp. 383–390, 2015.
- [3] P. Kayal and C. K. Chanda, "Placement of wind and solar based DGs in distribution system for power loss minimization and voltage stability improvement," *Int. J. Electr. Power Energy Syst.*, vol. 53, pp. 795–809, 2013.
- [4] O. Ekren and B. Y. Ekren, "Size optimization of a PV/wind hybrid energy conversion system with battery storage using simulated annealing," *Appl. Energy*, vol. 87, no. 2, pp. 592–598, 2010.
- [5] A. Roy, S. B. Kedare, and S. Bandyopadhyay, "Optimum sizing of wind-battery systems incorporating resource uncertainty," *Appl. Energy*, vol. 87, no. 8, pp. 2712–2727, 2010.
- [6] T. Ma, H. Yang, and L. Lu, "A feasibility study of a stand-alone hybrid solar-wind-battery system for a remote island," *Appl. Energy*, vol. 121, pp. 149–158, 2014.
- [7] J. Aghaei, K. M. Muttaqi, A. Azizivahed, and M. Gitizadeh, "Distribution expansion planning considering reliability and security of energy using modified PSO (Particle Swarm Optimization) algorithm," *Energy*, vol. 65, pp. 398–411, 2014.
- [8] A. C. Rueda-Medina, J. F. Franco, M. J. Rider, A. Padilha-Feltrin, and R. Romero, "A mixed-integer linear programming approach for optimal type, size and allocation of distributed generation in radial distribution systems," *Electr. Power Syst. Res.*, vol. 97, pp. 133–143, 2013.
- [9] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *IEEE Trans. Power Deliv.*, vol. 26, no. 4, pp. 2541–2551, 2011.
- [10] R. S. Al Abri, E. F. El-Saadany, and Y. M. Atwa, "Optimal placement and sizing method

- to improve the voltage stability margin in a distribution system using distributed generation,” *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 326–334, 2013.
- [11] D. K. Khatod, V. Pant, and J. Sharma, “Evolutionary programming based optimal placement of renewable distributed generators,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 683–695, 2013.
- [12] B. H. Dias, L. W. Oliveira, F. V. Gomes, I. C. Silva, and E. J. Oliveira, “Hybrid heuristic optimization approach for optimal Distributed Generation placement and sizing,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–6, 2012.
- [13] H. Pandzic, Y. Wang, T. Qiu, Y. Dvorkin, and D. S. Kirschen, “Near-Optimal Method for Siting and Sizing of Distributed Storage in a Transmission Network,” *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2288–2300, 2015.
- [14] Y. M. Atwa, E. F. El-Saadany, M. M. A. Salama, and R. Seethapathy, “Optimal renewable resources mix for distribution system energy loss minimization,” *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 360–370, 2010.
- [15] P. S. Acharya, M. M. Wagh, and V. V. Kulkarni, “Intelligent algorithmic multi-objective optimization for renewable energy system generation and integration problems: A review,” *Int. J. Renew. Energy Res.*, vol. 9, no. 1, pp. 271–280, 2019.
- [16] W. El-Khattam, Y. Hegazy, and M. Salama, “An integrated distributed generation optimization model for distribution system planning,” *2005 IEEE Power Eng. Soc. Gen. Meet.*, vol. 3, no. 2, p. 2392, 2005.
- [17] A. Saif, K. G. Elrab, H. H. Zeineldin, S. Kennedy, and J. L. Kirtley, “Multi-objective capacity planning of a PV-wind-diesel-battery hybrid power system,” *2010 IEEE Int. Energy Conf. Exhib. EnergyCon 2010*, pp. 217–222, 2010.
- [18] K. Mahesh, P. Nallagownden, and I. Elamvazuthi, “Advanced Pareto front non-dominated sorting multi-objective particle swarm optimization for optimal placement and sizing of distributed generation,” *Energies*, vol. 9, no. 12, 2016.
- [19] L. W. Oliveira, F. V. Gomes, E. J. Oliveira, A. R. Oliveira, A. M. Variz, and H. A. Silva, “Power distribution systems planning with distributed thermal and wind generation,” *2015 IEEE Eindhoven PowerTech, PowerTech 2015*, 2015.

- [20] N. T. Bazargani, S. M. T. Bathaee, and M. Hosseina, "Optimal Sizing of Battery Energy Storage and penetration degree of wind turbines using NSGA-II," *2017 25th Iran. Conf. Electr. Eng. ICEE 2017*, pp. 973–979, 2017.
- [21] O. Erdinc and M. Uzunoglu, "Optimum design of hybrid renewable energy systems: Overview of different approaches," *Renew. Sustain. Energy Rev.*, vol. 16, no. 3, pp. 1412–1425, 2012.
- [22] T. Griffin, K. Tomsovic, D. S. A. Law, and A. Corp, "Placement of Dispersed Generations Systems for Reduced Losses," vol. 00, no. c, pp. 1–9, 2000.
- [23] S. Xie, X. Wang, C. Qu, X. Wang, and J. Guo, "Impacts of different wind speed simulation methods on conditional reliability indices," *Int. Trans. Electr. energy Syst.*, vol. 20, no. March, pp. 1–6, 2013.
- [24] A. Ramamoorthy and R. Ramachandran, "Optimal Siting and Sizing of Multiple DG Units for the Enhancement of Voltage Profile and Loss Minimization in Transmission Systems Using Nature Inspired Algorithms," *Sci. World J.*, vol. 2016, 2016.
- [25] E. S. Ali, S. M. A. Elazim, and A. Y. Abdelaziz, "Ant Lion Optimization Algorithm for optimal location and sizing of renewable distributed generations," *Renew. Energy*, vol. 101, pp. 1311–1324, 2017.
- [26] I. Ben Hamida, S. B. Salah, F. Msahli, and M. F. Mimouni, "Optimal network reconfiguration and renewable DG integration considering time sequence variation in load and DGs," *Renew. Energy*, vol. 121, pp. 66–80, 2018.
- [27] M. Wahbah, O. Alhussein, T. H. M. El-Fouly, B. Zahawi, and S. Muhaidat, "Evaluation of Parametric Statistical Models for Wind Speed Probability Density Estimation," *2018 IEEE Electr. Power Energy Conf. EPEC 2018*, pp. 1–6, 2018.
- [28] D. Mazzeo, G. Oliveti, and E. Labonia, "Estimation of wind speed probability density function using a mixture of two truncated normal distributions," *Renew. Energy*, vol. 115, pp. 1260–1280, 2018.
- [29] Q. Hu, Y. Wang, Z. Xie, P. Zhu, and D. Yu, "On estimating uncertainty of wind energy with mixture of distributions," *Energy*, vol. 112, pp. 935–962, 2016.
- [30] Z. Qin, W. Li, and X. Xiong, "Estimating wind speed probability distribution using kernel

- density method,” *Electr. Power Syst. Res.*, vol. 81, no. 12, pp. 2139–2146, 2011.
- [31] B. G. Kumaraswamy, B. K. Keshavan, and S. H. Jangamshetti, “A statistical analysis of wind speed data in west central part of karnataka based on weibull distribution function,” *2009 IEEE Electr. Power Energy Conf. EPEC 2009*, pp. 1–4, 2009.
- [32] G. Li and J. Shi, “Application of Bayesian model averaging in modeling long-term wind speed distributions,” *Renew. Energy*, vol. 35, no. 6, pp. 1192–1202, 2010.
- [33] X. Xu, Z. Yan, and S. Xu, “Estimating wind speed probability distribution by diffusion-based kernel density method,” *Electr. Power Syst. Res.*, vol. 121, pp. 28–37, 2015.
- [34] F. K. Wang and P. R. Huang, “Implementing particle swarm optimization algorithm to estimate the mixture of two Weibull parameters with censored data,” *J. Stat. Comput. Simul.*, vol. 84, no. 9, pp. 1975–1989, 2014.
- [35] A. Justel, D. Pefia, and R. Zamar, “~!. STATISTICS& PROBABILITY LETTERS A multivariate Kolmogorov-Smimov test of goodness,” *Stat. Probab. Lett.*, vol. 35, pp. 251–259, 1997.
- [36] S. Greenland *et al.*, “Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations,” *Eur. J. Epidemiol.*, vol. 31, no. 4, pp. 337–350, 2016.
- [37] E. Parzen, “On Estimation of a Probability Density Function and Mode,” *Institute of Mathematical Statistics.*, vol. 33, no. 3, pp. 1065–1076, 2020.
- [38] M. Wahbah, T. H. M. El-Fouly, B. Zahawi, and S. Feng, “Hybrid Beta-KDE Model for Solar Irradiance Probability Density Estimation,” *IEEE Trans. Sustain. Energy*, vol. 3029, no. c, pp. 1–1, 2019.
- [39] M.F.Kotb, K.M.Shebl, M. El Khazendar, and A. El Hussein, “Genetic Algorithm for Optimum Siting and Sizing of Distributed Generation,” *14th Int. Middle East Power Syst. Conf.*, no. July, pp. 433–440, 2010.
- [40] K. Mahmoud and Y. Naoto, “Optimal Siting and Sizing of Distributed Generations,” *Electric Distribution Network Planning.*, 2018.
- [41] Y. Y. Zakaria, R. A. Swief, N. H. El-Amary, and A. M. Ibrahim, “Optimal Distributed Generation Allocation and Sizing Using Genetic and Ant Colony Algorithms,” *J. Phys.*

*Conf. Ser.*, vol. 1447, no. 1, 2020.

- [42] S. Das, D. Das, and A. Patra, "Operation of distribution network with optimal placement and sizing of dispatchable DGs and shunt capacitors," *Renew. Sustain. Energy Rev.*, vol. 113, no. January, p. 109219, 2019.
- [43] A. J. Litchy and M. H. Nehrir, "Real-time energy management of an islanded microgrid using multi-objective Particle Swarm Optimization," *IEEE Power Energy Soc. Gen. Meet.*, vol. 2014-Octob, no. October, pp. 1–5, 2014.
- [44] A. Amer, M. A. Azzouz, A. Azab, and A. S. A. Awad, "Stochastic Planning for Optimal Allocation of Fast Charging Stations and Wind-Based DGs," *IEEE Syst. J.*, pp. 1–11, 2020.
- [45] "<https://www.nrcan.gc.ca/science-data/data-analysis/energy-data-analysis/energy-facts/renewable-energy-facts/20069> | Mendeley." <https://www.mendeley.com/reference-management/web-importer> (accessed Nov. 16, 2020).
- [46] Y. M. Atwa, "Distribution System Planning and Reliability Assessment under High DG Penetration," *Distribution*, 2010.
- [47] J. V Seguro and T. W. Lambert, "Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis," vol. 85, 2000.
- [48] T. Soukissian, "Use of multi-parameter distributions for offshore wind speed modeling: The Johnson SB distribution," *Appl. Energy*, vol. 111, pp. 982–1000, 2013.
- [49] "Historical Climate Data - Climate - Environment and Climate Change Canada." <https://climate.weather.gc.ca/> (accessed Nov. 17, 2020).
- [50] M. A. Stephens, "Introduction to Kolmogorov ( 1933 ) On the Empirical Determination of a Distribution," pp. 93-105, 1933.
- [51] C. Res, C. J. Willmott, and K. Matsuura, "Advantages of the mean absolute error ( MAE ) over the root mean square error ( RMSE ) in assessing average model performance," vol. 30, pp. 79–82, 2005.
- [52] Y. Yu, "Finite mixture modeling for raw and binned data," Package 'mixR ,' Jun. 05, 2018.

- [53] C. F. Jeff Wu, "On the Convergence Properties of the EM Algorithm," *Institute of mathematical statistics*, vol. 11, no. 1, pp. 95–103, 1983.
- [54] U. L. Approximations and S. Equa-, "The Newton-Raphson Method Using Linear Approximations to Solve Equations," pp. 1–11.
- [55] T. Senjyu, R. Sakamoto, S. Member, and N. Urasaki, "Output Power Leveling of Wind Turbine Generator for All Operating Regions by Pitch Angle Control," *IEEE*, vol. 21, no. 2, pp. 467–475, 2006.
- [56] J. M. S. Pinheiro, C. R. R. Dornellas, R. De Janeiro, and A. C. G. Melo, "Probing the New IEEE Reliability Test System (RTS-96): HL-II Assessment," *IEEE*, vol. 13, no. 1, pp. 171–176, 1998.
- [57] C. Singh, Y. Kim, "An Efficient Technique for Reliability Analysis of Power Systems Including Time Dependent Sources," *IEEE Transactions on Power Systems*, vol. 3, no. 3, pp. 1090–1096, 1988.
- [58] Ch. Singh, A. Lago-gonzalez, "Improved Algorithms for Multi-Area Reliability Evaluation Using The Decomposition-Simulation Approach," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 321–328, 1989.
- [59] IEEE, "IEEE Application Guide for IEEE Std 1547, IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems," *IEEE-SA Standards Board*, no. April. 2008.
- [60] R. N. Allan, R. Billinton, A. M. Breipohl, C. H. Grigg, "Bibliography on The Application of Probability Methods in Power System Reliability Evaluation," *IEEE Transaction Power Systems*, vol. 14, no. 1, pp. 51–57, 1999.
- [61] M. Gandomkar, M. Vakilian, M. Ehsan, "A Combination of Genetic Algorithm and Simulated Annealing for Optimal DG Allocation in Distribution Networks," *Canadian Conference on Electrical and Computer Engineering*, no. May, 2005.
- [62] Y. Alinejad-Beromi, M. Sedighzadeh, M. R. Bayat, and M. E. Khodayar, "Using Genetic Algorithm for Distributed Generation Allocation to Reduce Losses and Improve Voltage Profile," *2007 42<sup>nd</sup> International Universities Power Engineering Conference*, no. 1, pp. 954–959, 2007.

## VITA AUCTORIS

NAME: Mohammad Borooshan

PLACE OF BIRTH: Tehran, Iran

YEAR OF BIRTH: 1995

EDUCATION: Islamic Azad University of Firuzkuh, B.Sc.,  
Tehran, Iran, 2017

University of Windsor, M.Sc., Windsor, ON,  
2021