IMPACT OF WIND POWER PENETRATION ON POWER SYSTEM SECURITY BY A PROBABILISTIC APPROACH

Doctoral Dissertation of:
Dinh-Duong Le

Supervisor:
Prof. Alberto Berizzi

Co-supervisor:
Dr. Diego Cirio

The Chair of the Doctoral Program:
Prof. Alberto Berizzi
NOWADAYS, in order to achieve environmental and economic benefits, renewable energy sources, such as wind and photovoltaic solar, are widely used. The integration of renewable resources into power systems is one of the major challenges in planning and operations of modern power systems. The integration has introduced additional uncertainty into various study areas of power system, together with the conventional sources of uncertainty such as the loads and the availability of resources and transmission assets; this makes clear the limitations of the conventional deterministic analysis and security assessment approaches, in which sources of uncertainty and stochastic factors affecting power system are not considered. To solve such problems, probabilistic approaches need to be used. They have been introduced and are gaining wider application in power systems with increasing levels of renewable energy sources.

The research firstly aims at developing probabilistic power flow tools which are capable of managing the wide spectrum of all possible values of the input and state variables so as to provide a complete spectrum of all possible values of outputs of interest such as nodal voltages, line power flows, etc., in terms of probability distributions which are useful for power system analysis and security assessment by probabilistic approaches.

To be taken into account in computations for power system security assessment by a probabilistic approach, modeling of various stochastic factors in power system, such as stochastic behaviour of load, wind power generation, random outages of generating units and branches, is required. Their probabilistic models are also considered in the thesis.

Among renewable resources, wind power generation is one of the most important and the most challenging ones because of its variability so that that will be focused on to stress the methodology in the research. Building a model of multi-site wind power production for power system planning and operations with large integration of wind power resources is a critical need. However, this work is very challenging, because of the stochastic features of wind speed and wind power at multiple wind farm locations.
The thesis also aims at building a model for wind speed and wind power capturing all of their stochastic characteristics. Such a model would be a very useful tool to deal with many problems in power systems involving multi-site wind power production.

In general, the analytic characterization of the random and time-varying wind power output is not available, because it is considerably more complicated than that of wind speed due to the highly non-linear mapping of wind speed into wind power output. Moreover, the spatial and temporal correlations among the wind speed and therefore the wind power output at the multi-site wind farm locations bring additional layer of complexity. In addition, when wind power data are not available due to, for example, commercial reasons or in case of new wind farms, the model for wind speed is firstly built and then wind power data are derived. For mapping wind speed to wind power for an entire wind farm or location to be used in power system studies, an approach to construct an aggregate power curve is also developed in the thesis. The procedure can be done automatically, so reducing cost and time consumption.
Acknowledgements

This work has been financed by the Research Fund for the Italian Electrical System under the Contract Agreement between RSE S.p.A. and the Ministry of Economic Development - General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency in compliance with the Decree of March 8, 2006.

First and foremost, I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Alberto Berizzi, for the invaluable direction, support, discussions as well as his kindness, patience, and understanding throughout the whole PhD study.

I am very grateful to my co-supervisor, Dr. Diego Cirio, at RSE for his advice, suggestions, and insightful discussions during my study.

I would also like to thank Prof. Cristian Bovo at the Department of Energy, Politecnico di Milano for his continuous help and support.

The support of Dr. Massimo Gallanti from the Energy System Department at RSE is gratefully acknowledged. I wish to give special thanks to Dr. Emanuele Ciapessoni and Dr. Andrea Pitto at RSE for their technical support and fruitful discussions.

I would like to express my deep gratefulness to Prof. George Gross at Electrical and Computer Engineering Department, University of Illinois at Urbana-Champaign (UIUC) for his guidance, enthusiasm, and support during the six-month period of working as a visiting scholar at UIUC under his supervision and till now.

I also wish to thank Terna (Italian TSO) and in particular Dr. Enrico Carlini for providing useful data for the research.

Of course, many thanks go to my friends and colleagues at the Department of Energy, Politecnico di Milano for making the working environment enjoyable and colourful.

From the bottom of my heart, I wish to thank my family in Vietnam and my wife for their endless love, support and understanding.
# Contents

1 Introduction ................................................................. 7
   1.1 Background and motivation ........................................... 7
   1.2 Literature review ..................................................... 8
   1.3 Contributions and outline of the thesis .............................. 9
   1.4 List of publications .................................................. 11

2 Mathematical Background .................................................. 13
   2.1 Introduction ............................................................ 13
   2.2 Probability of stochastic events ...................................... 13
   2.3 Random variable and its distribution ................................. 14
   2.4 Characteristic function ................................................ 14
   2.5 Moments and cumulants ............................................... 15
      2.5.1 Moments ............................................................ 15
      2.5.2 Cumulants ........................................................ 16
   2.6 Joint moments and joint cumulants ................................... 16
   2.7 Applying properties of cumulants to a linear combination of random variables .................................................. 17
   2.8 Probability distributions most used in probabilistic analysis of electrical power systems ............................................. 18
      2.8.1 Uniform distribution ............................................... 18
      2.8.2 Normal distribution ............................................... 19
      2.8.3 Binomial distribution .............................................. 21
      2.8.4 Weibull distribution .............................................. 22
   2.9 Approximations to probability density function and cumulative distribution function of random variables ................................. 24
      2.9.1 Approximation methods based on series expansions ............ 24
      2.9.2 Approximation method based on Von Mises function ........... 25
   2.10 Time series analysis ................................................... 27
   2.11 Conclusions ........................................................... 31
## Contents

### 3 Power System Security
- 3.1 Definitions .................................................. 33
- 3.2 Power system security assessment ......................... 36
  - 3.2.1 Deterministic security assessment ....................... 36
  - 3.2.2 Probabilistic security assessment ....................... 37
  - 3.2.3 Probabilistic vs. deterministic security assessment .... 39
- 3.3 Conclusions .................................................. 39

### 4 Wind Power Models for Security Assessment
- 4.1 Introduction .................................................. 41
- 4.2 Wind power forecast techniques and use in power system studies .... 42
- 4.3 A Multi-site model for wind speed and wind power production .... 44
  - 4.3.1 Introduction .............................................. 44
  - 4.3.2 Structural representation of wind data and Principal Component Analysis .............................................. 45
  - 4.3.3 Proposed methodology .................................... 47
  - 4.3.4 Tests and results .......................................... 49
- 4.4 Wind power curve ............................................ 66
- 4.5 Conclusions .................................................. 74

### 5 Probabilistic Security Assessment
- 5.1 Probabilistic models for security assessment of power systems under uncertainty .......... 75
  - 5.1.1 Introduction .............................................. 75
  - 5.1.2 Probabilistic model of load ................................ 75
  - 5.1.3 Probabilistic model of wind power production .......... 77
  - 5.1.4 Probabilistic models of branch outage and generating unit outage .............................................. 78
  - 5.1.5 Conclusions .............................................. 82
- 5.2 Probabilistic power flow ...................................... 82
  - 5.2.1 Introduction .............................................. 82
  - 5.2.2 Overview of probabilistic power flow methodologies .............................................. 82
  - 5.2.3 Formulation of cumulant-based probabilistic power flow methods .............................................. 83
  - 5.2.4 Tests and numerical results ................................ 90
  - 5.2.5 Final comments on the application of the cumulant-based PPF methods .............................................. 100
  - 5.2.6 Conclusions .............................................. 101
- 5.3 Distributed slack bus probabilistic power flow .................. 102
  - 5.3.1 Background and motivation ................................ 102
  - 5.3.2 Distributed slack bus in power flow calculation .......... 102
  - 5.3.3 Distributed slack bus probabilistic power flow .......... 103
  - 5.3.4 Tests and numerical results ................................ 105
  - 5.3.5 Conclusions .............................................. 124

### 6 Conclusions and Future Work
- 6.1 Conclusions .................................................. 125
- 6.2 Future work .................................................. 127
# List of Figures

2.1 *p.d.f.* of uniform distribution $U(a, b)$ ........................................... 19  
2.2 *c.d.f.* of uniform distribution $U(a, b)$ ........................................... 19  
2.3 *p.d.f.*s of normal distributions ...................................................... 20  
2.4 *c.d.f.*s of normal distributions ...................................................... 21  
2.5 *p.m.f.*s of binomial distributions .................................................... 22  
2.6 *c.d.f.*s of binomial distributions .................................................... 22  
2.7 *p.d.f.*s of Weibull distributions ..................................................... 23  
2.8 *c.d.f.*s of Weibull distributions ..................................................... 24  
2.9 Stationary time series ........................................................................ 28  
2.10 Non-stationary time series: variance changes over time ....................... 29  
2.11 Non-stationary time series with trend and seasonal pattern ................. 29  
2.12 Non-correlation between two time series .......................................... 30  
2.13 Correlation between two time series ................................................. 30  
2.14 White noise $WN(0,1)$ ....................................................................... 31  
3.1 Decision drivers of power system security .......................................... 34  
3.2 System operating states and their transitions .................................... 36  
3.3 *p.d.f.* of r.v $X$ ..................................................................................... 39  
4.1 Representation of the stochastic process .............................................. 46  
4.2 The flow diagram of the proposed approach ....................................... 48  
4.3 Wind locations in the region of Basilicata in Italy .................................. 50  
4.4 10-minute wind speed measurement from March 1, 2001 to February 28, 2002 .......................................................... 51  
4.5 Scatter plot of observed wind speed for locations $F$ and $P$ ................. 51  
4.6 Scatter plot of observed wind speed for locations $F$ and $V$ .................. 52  
4.7 Scatter plot of observed wind speed for locations $P$ and $C$ ................. 52  
4.8 Transformed stationary data of five locations ...................................... 53  
4.9 *c.d.f.*s before and after using Gaussian transform for location $F$ ....... 53  
4.10 The construction of five PCs .............................................................. 54  
4.11 Scatter plot of $z_1$ and $z_2$ ............................................................... 54
### List of Figures

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.12</td>
<td>Scatter plot of $z_1$ and $z_2$ in case of without using pre-processing and transformation techniques.</td>
</tr>
<tr>
<td>4.13</td>
<td>Residual test for time series model of $z_1$.</td>
</tr>
<tr>
<td>4.14</td>
<td>Histogram and c.d.f. of wind speed at the time step of 30 minutes ahead for location $T$.</td>
</tr>
<tr>
<td>4.15</td>
<td>Hourly wind speed measurement from September 1, 2011 to August 31, 2012.</td>
</tr>
<tr>
<td>4.16</td>
<td>Scatter plot of observed wind speed for locations $L_1$ and $L_3$.</td>
</tr>
<tr>
<td>4.17</td>
<td>Scatter plot of observed wind speed for locations $L_2$ and $L_6$.</td>
</tr>
<tr>
<td>4.18</td>
<td>Scatter plot of observed wind speed for locations $L_4$ and $L_9$.</td>
</tr>
<tr>
<td>4.19</td>
<td>Scatter plot of observed wind speed for locations $L_5$ and $L_8$.</td>
</tr>
<tr>
<td>4.20</td>
<td>Scatter plot of observed wind speed for locations $L_2$ and $L_8$.</td>
</tr>
<tr>
<td>4.21</td>
<td>c.d.f.s of transformed stationary data at nine locations.</td>
</tr>
<tr>
<td>4.22</td>
<td>c.d.f.s before and after using (4.13) for location $L_1$.</td>
</tr>
<tr>
<td>4.23</td>
<td>PC time series.</td>
</tr>
<tr>
<td>4.24</td>
<td>Residuals of dimensional approximation for location $L_5$.</td>
</tr>
<tr>
<td>4.25</td>
<td>Typical wind turbine power curve.</td>
</tr>
<tr>
<td>4.26</td>
<td>Measured wind power against measured wind speed for a real wind turbine [1].</td>
</tr>
<tr>
<td>4.27</td>
<td>Wind power versus wind speed for location $L_1$.</td>
</tr>
<tr>
<td>4.28</td>
<td>Wind power versus wind speed for location $L_3$.</td>
</tr>
<tr>
<td>4.29</td>
<td>Wind power versus wind speed for location $L_5$.</td>
</tr>
<tr>
<td>4.30</td>
<td>Wind power versus wind speed for location $L_7$.</td>
</tr>
<tr>
<td>4.31</td>
<td>Wind power versus wind speed for location $L_8$.</td>
</tr>
<tr>
<td>4.32</td>
<td>Approximate power curve for location $L_5$.</td>
</tr>
<tr>
<td>4.33</td>
<td>Approximate power curve for location $L_5$.</td>
</tr>
<tr>
<td>5.1</td>
<td>Load duration curve.</td>
</tr>
<tr>
<td>5.2</td>
<td>Example of a discrete load.</td>
</tr>
<tr>
<td>5.3</td>
<td>Wind power modeling approaches.</td>
</tr>
<tr>
<td>5.4</td>
<td>ORR vs. FOR.</td>
</tr>
<tr>
<td>5.5</td>
<td>An example of probabilistic modeling for generating unit outage.</td>
</tr>
<tr>
<td>5.6</td>
<td>Modeling of branch outage.</td>
</tr>
<tr>
<td>5.7</td>
<td>Single line diagram of the IEEE 14-bus test system [2].</td>
</tr>
<tr>
<td>5.8</td>
<td>Standard deviation of selected nodal voltage angles.</td>
</tr>
<tr>
<td>5.9</td>
<td>Standard deviation of nodal voltage magnitudes.</td>
</tr>
<tr>
<td>5.10</td>
<td>Standard deviation of selected real power flows.</td>
</tr>
<tr>
<td>5.11</td>
<td>Standard deviation of selected reactive power flows.</td>
</tr>
<tr>
<td>5.12</td>
<td>p.d.f.s of $V_{12}$.</td>
</tr>
<tr>
<td>5.13</td>
<td>c.d.f.s of $Q_{3-4}$.</td>
</tr>
<tr>
<td>5.14</td>
<td>p.d.f.s of $Q_{3-4}$.</td>
</tr>
<tr>
<td>5.15</td>
<td>c.d.f.s of $P_{3-4}$.</td>
</tr>
<tr>
<td>5.16</td>
<td>c.d.f.s of $P_{3-4}$ with random outage line 2-4.</td>
</tr>
<tr>
<td>5.17</td>
<td>p.d.f.s of $P_{126-132}$.</td>
</tr>
<tr>
<td>5.18</td>
<td>p.d.f.s of $Q_{126-132}$.</td>
</tr>
<tr>
<td>5.19</td>
<td>c.d.f.s of $Q_{126-132}$.</td>
</tr>
</tbody>
</table>
List of Figures

5.20 SSBPPF vs. DSBPPF .................................................. 104
5.21 Single line diagram of the modified IEEE 14-bus test system ... 106
5.22 p.d.f.s of $P_{g2}$ at time step $t_k$ .................................. 108
5.23 p.d.f.s of $P_{g2}$ at time step $t_{k+1}$ .............................. 108
5.24 p.d.f. of ramping $R_{g2}$ of generator $G_2$ ....................... 109
5.25 c.d.f.s of $V_9$ at time step $t_{k+1}$ ................................. 110
5.26 p.d.f.s of $P_{2-3}$ at time step $t_{k+1}$ .............................. 110
5.27 c.d.f.s of $P_{2-3}$ at time step $t_{k+1}$ .............................. 111
5.28 Impacts of explicit representation of correlations on $\tilde{P}_{g2}$ at $t_{k+1}$ .............................. 111
5.29 Impacts of explicit representation of correlations on $P_{2-3}$ at $t_{k+1}$ .............................. 112
5.30 Impacts of explicit representation of correlations on $Q_{2-3}$ at $t_{k+1}$ .............................. 112
5.31 Impacts of explicit representation of correlations on $V_9$ at $t_{k+1}$ .............................. 113
5.32 Impacts of contingencies on $P_{g2}$ at $t_{k+1}$ .............................. 114
5.33 Impacts of contingencies on ramping $R_{g2}$ of generator $G_2$ .............................. 114
5.34 c.d.f. curves of $P_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies .............................. 115
5.35 p.d.f. curves of $Q_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies .............................. 115
5.36 c.d.f. curves of $Q_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies .............................. 116
5.37 Impacts of contingencies on $V_9$ at $t_{k+1}$ .............................. 116
5.38 Impacts of contingencies on $P_{2-3}$ at $t_{k+1}$ .............................. 117
5.39 p.d.f.s of $V_{12}$ (voltage level: 150kV) .............................. 119
5.40 p.d.f.s of $P_{110-66}$ ................................................. 120
5.41 c.d.f.s of $P_{110-66}$ ................................................. 121
5.42 p.d.f.s of $Q_{110-66}$ ................................................. 122
5.43 p.d.f.s of $P_{g468}$ ................................................. 123

A.1 p.m.f. of $P_{l9}$ .................................................. 131
A.2 p.m.f. of $Q_{l9}$ .................................................. 132
A.3 p.m.f. of $P_{g1}$ .................................................. 133
A.4 p.m.f. of $P_{g2}$ .................................................. 134
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Security-related decisions in power system security assessment</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Probabilistic vs. deterministic security assessment</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>Covariance matrix of observed wind speed data from five locations in Basilicata</td>
<td>50</td>
</tr>
<tr>
<td>4.2</td>
<td>The contribution of five PCs</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Covariance matrix of observed wind speed data from nine locations in Italy</td>
<td>57</td>
</tr>
<tr>
<td>4.4</td>
<td>The contribution of nine PCs</td>
<td>58</td>
</tr>
<tr>
<td>5.1</td>
<td>ARMS for 3 selected output r.v.s</td>
<td>96</td>
</tr>
<tr>
<td>5.2</td>
<td>ARMS of $\tilde{P}_{3-4}$ with random outage line 2-4</td>
<td>96</td>
</tr>
<tr>
<td>5.3</td>
<td>Computation time comparison for IEEE 300-bus test system</td>
<td>99</td>
</tr>
<tr>
<td>5.4</td>
<td>Computation time of method M2 with different thresholds</td>
<td>99</td>
</tr>
<tr>
<td>5.5</td>
<td>ARMS (%) of IEEE 300-bus test system (large errors in bold)</td>
<td>100</td>
</tr>
<tr>
<td>5.6</td>
<td>Indications for the application of methods</td>
<td>101</td>
</tr>
<tr>
<td>5.7</td>
<td>Wind power forecasts at time step $t_k$</td>
<td>105</td>
</tr>
<tr>
<td>5.8</td>
<td>Load forecast at time step $t_k$</td>
<td>106</td>
</tr>
<tr>
<td>5.9</td>
<td>Correlation coefficients among loads</td>
<td>107</td>
</tr>
<tr>
<td>5.10</td>
<td>Wind power forecasts at time step $t_{k+1}$</td>
<td>107</td>
</tr>
<tr>
<td>5.11</td>
<td>Real power schedules (MW) at the considered time steps</td>
<td>112</td>
</tr>
<tr>
<td>5.12</td>
<td>Outage replacement rate</td>
<td>118</td>
</tr>
<tr>
<td>5.13</td>
<td>Computation time comparison</td>
<td>118</td>
</tr>
<tr>
<td>A.1</td>
<td>Branch data for IEEE 14-bus test system</td>
<td>130</td>
</tr>
<tr>
<td>A.2</td>
<td>Normally distributed loads for IEEE 14-bus test system</td>
<td>130</td>
</tr>
<tr>
<td>A.3</td>
<td>Discretely distributed load at bus 9 for IEEE 14-bus test system</td>
<td>131</td>
</tr>
<tr>
<td>A.4</td>
<td>Binomial distributions for IEEE 14-bus test system</td>
<td>131</td>
</tr>
<tr>
<td>B.1</td>
<td>Discrete loads for IEEE 300-bus test system</td>
<td>136</td>
</tr>
<tr>
<td>C.1</td>
<td>Nominal power of wind farms</td>
<td>137</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

1.1 Background and motivation

Planning and operations of modern power systems are challenging due to the integration of many sources of uncertainty especially from Renewable Energy Source (RES). Among RES, wind generation is one of the most important sources, and the most challenging one; that is why it is chosen to focus on in the thesis.

In power systems with significant integration of RES, building a model which captures all salient features of wind power production such as high variability, temporal and spatial correlations, diurnal and seasonal patterns, non-normality, non-stationarity, etc., is one of the most challenging issues. Wind power is known as a time-varying, intermittent, uncertain, and non-dispatchable resource. Wind speed is a non-Gaussian and non-stationary stochastic process with distinct diurnal and seasonal patterns. Wind speed at a specific location is temporally correlated, while wind speeds at different locations are correlated both in space and in time. These features make the modeling for wind very challenging, especially in cases of large number of wind farms requiring high-dimensional multivariate representation. Their modeling is necessary for the security assessment, where congestions, voltage profiles, security violations, generator ramping limits, etc., must be assessed. Such a model brings many benefits to attendees in the electricity world: for Transmission System Operators (TSOs) to improve the accuracy of the estimation of the operation for the next day, to increase security margins, to reduce the needs (and the related costs) for spinning and non-spinning reserve requirements, to compute more accurately Total Transfer Capacity (TTC) values, to increase the quality of information given to market operators, to evaluate more accurately possible needs for new transmission investments, etc.; for regulators to better estimate the behaviour of TSO and to identify possible discriminatory behaviour, etc.; for market operators to build more convenient bid structures for the day-ahead market,
to improve their risk assessment procedures and methodologies and so on. A number of models of wind power generation have been introduced and a few applications already exists. Nevertheless, significant improvements are still needed to obtain more accurate results, and especially a "realistic" model for wind is critical need for realistic-sized power systems.

In general, the analytic characterization of the random and time-varying wind power is not available. Indeed, the modeling of wind power is considerably more complicated than that of wind speed due to the highly non-linear mapping of wind speed into wind power production. Moreover, the correlation among the wind speed and therefore the wind output at the multi-site wind farm locations brings additional layer of complexity. In addition, modeling wind power production requires characterizing wind speed when wind power data are not available due to, for example, commercial reasons or in case of planning for new wind farms. For such cases, the model for wind speed is firstly built and then wind power data can be constructed from wind speed data.

The integration of RES has introduced additional uncertainty into power system analysis and security assessment. This source of uncertainty combined with the conventional sources such as the loads and the availability of resources and transmission assets makes clear the limitations of the conventional deterministic analysis and security assessment approaches in which stochastic factors and sources of uncertainty affecting power systems are not considered. Similarly to the very important role of the conventional deterministic power flow in power system analysis, security assessment, operations, planning, and control, probabilistic power flow tool is a crucial need to deal with high levels of uncertainty and to account for stochastic factors in power systems. Probabilistic approaches have introduced and can serve as effective tools for power system planning, operations, and security assessment under uncertainty. However, significant improvements are still possible in order to overcome some evident drawbacks of the techniques already implemented, to obtain a significant increase in the accuracy of results and the computational efficiency, to provide more useful security-related information as well as to give suitable suggestions toward applications.

1.2 Literature review

In this section, a general overview of the previous related work is given. However, a detailed literature review on specific work can be found in the corresponding chapters.

In spite of the fact that deterministic analysis and security assessment are widely used approach in practice, the probabilistic approach also has a solid theoretical background [3,4] and is gaining wider application in power systems with the significant integration of RES. Probabilistic approach is based on probabilistic point of view and adopts probability theory and statistics [5–11], in order to be capable of taking into account all stochastic factors in power systems to provide realistic results.

Probabilistic approach has been applied to various areas in power system; it has been investigated deeply in the present work. The detailed literature review on probabilistic power flow methodologies as well as their developments are presented in Chapter 5.

The thesis also focuses on building a model for wind speed and wind power production from multiple sites. A number of techniques have been introduced to capture the wind characteristics from available data and reproduce wind data for a desired horizon.
in future. The future wind data can be derived from a set of quantiles that represents
the discrete probability distribution of the wind generation random variable \[12,13\].
Another form of representation of the data is the set of intervals that provides a range
of possible values \[14–16\]. Both interval and quantile representations can provide the
marginal distribution for each time step in future but they do not characterize the inter-
temporal dependence between different time steps. On the other way, there are some
approaches to build a model capturing salient features from wind data and then the
model built is used to generate a set of scenarios – the so-called sample paths over the
predefined time horizon – which is very useful for probabilistic security assessment
and solving decision-making problem under uncertainty. Regarding this approach, a
number of methods are proposed in the literature \[17\]. In recent work, the authors
in \[18\] build the model for multi-site wind speed using a noise vector that drives a vec-
tor autoregressive process. This model accounts both spatial and temporal correlations
of wind speed; it assumes joint Gaussian distribution and stationarity. However, these
assumptions are not always satisfied in reality. This model is improved in \[19,20\] by
using some techniques to obtain approximate actual wind speed data to stationary and
Gaussian data. In \[18–20\], the problem is simplified by the assumption that the matrix
of autoregressive coefficients is diagonal. This implies that spatial correlations of wind
speed at different locations are modeled fully by the underlying noise vector so that the
multivariate time series model can be decoupled into different univariate time series
model for each wind site. To be more realistic and capable of applying to real-size
power systems, the model for wind power production needs to be developed.

## 1.3 Contributions and outline of the thesis

The main contributions of the thesis is summarized as follows:

1) The investigation of various probabilistic power flow approaches is presented.
   An in-depth analysis and comparison of cumulant-based probabilistic power flow
   methodologies are given. Moreover, suitable suggestions to exploit their features
   in the different power system applications are provided. Further researches on
   these methods are also being developed to increase the accuracy and the compu-
tational efficiency.

2) Probabilistic power flow with distributed slack bus is proposed. Beside explicitly
taking into account the correlation of input nodal power injections and contin-
gencies due to branch and generating unit outages, it can also explicitly represent
the steady-state behaviour of the frequency control so as to provide more use-
ful information for security assessment such as probability of line overloading,
probability of over-/under-voltage and, in particular, the probability of not meet-
ing ramping requirements and probability of violation of over-/under-regulation
limits of conventional generators. The distributed slack bus probabilistic power
flow is developed so that any power imbalance in power system can be charged to
a set of dispatchable generators, so reflecting the actual power system operation.
The formulation is fully general and it can be applied to primary, secondary, and
tertiary regulation.

3) A multi-site model for wind speed and wind power production is proposed with
Chapter 1. Introduction

the main attractive features:

– it is capable of capturing the salient features of wind power generation from multiple wind farm locations, e.g., high variability, temporal and spatial correlations, diurnal and seasonal patterns, etc.;
– it is realistic because it does not require any assumption to develop the model: owing to this features, it can be used for any real wind data;
– it is a powerful tool for both decorrelating data and reducing computational burden by adopting principal component analysis technique, thus being suitable for dealing with high-dimensional data like wind data from multiple wind sites;
– it provides a very good direction for dealing with many problems in power systems involving wind power production from multiple locations, especially providing a very appropriate input for probabilistic security assessment and solving decision-making problems under uncertainty, for both operating and planning horizons.

4) An approach to construct an aggregate power curve for an entire wind farm or location is also developed for mapping wind speed to wind power to be used in power system studies. The procedure can be done automatically, so reducing time consumption.

The thesis is organized as follows:

Chapter 1: the background and motivation of the thesis are presented, together with the literature review. Then, the main contributions, the outline of the thesis, and also the list of publications are given.

Chapter 2: this chapter briefly recalls fundamental concepts and methods in probability theory and statistics which are used in the research.

Chapter 3: deterministic and probabilistic approaches for security assessment are presented and compared in this chapter.

Chapter 4: wind power forecast techniques as well as their use in power system studies are discussed. A multi-site model for wind speed and wind power production is proposed. An approach to construct aggregate power curve for an entire wind farm or location is developed.

Chapter 5: various probabilistic power flow methods are investigated and compared. An in-depth analysis and comparison of cumulant-based probabilistic power flow methodologies are given together with a number of developments to increase the accuracy and the computational efficiency. Probabilistic power flow with distributed slack bus is proposed.

Chapter 6: the conclusions of the thesis are presented in this chapter. A number of future areas of research are also provided in the end.
1.4 List of publications


2.1 Introduction

In power system, there are many phenomena occurring as stochastic events and processes such as the fluctuation of load, the random failures of generation, transmission, and distribution components, and the stochastic nature of renewable resources (e.g., wind, photovoltaic solar, etc.) and so on. These stochastic factors make clear the limitations of deterministic approaches for analysis and security assessment of power system. Therefore, probabilistic approaches, based on probabilistic point of view and adopting probability theory and statistics, have been introduced to deal with the above problem. The present chapter briefly recalls fundamental concepts and methods in probability theory and statistics which are used in this thesis.

2.2 Probability of stochastic events

In probability theory and statistics, a stochastic event is an event that may or may not occur. Probability is a measure for the possibility of occurrence of stochastic events. The probability of event $A$ is usually denoted as $P\{A\}$:

$$0 \leq P\{A\} \leq 1$$ (2.1)

When $P\{A\} = p$, the stochastic event $A$ is likely to occur with the probability $p$. In particular, if:

- $P\{A\} = 1$: occurrence of event $A$ is certain;
- $P\{A\} = 0$: non-occurrence of event $A$ is certain.

The set of all possible outcomes of a random phenomenon is called sample space, denoted as $\Omega$. 
Chapter 2. Mathematical Background

The sum of probabilities of all possible outcomes of an event is equal to 1: \( P(\Omega) = 1 \).

Example 2.1: Tossing a coin, there are two possible outcomes, i.e., head \((H)\) and tail \((T)\), in the sample space: \( \Omega = \{H, T\} \). If the coin is balanced, the probabilities for getting either head or tail will be: \( P\{H\} = P\{T\} = 0.5 \).

2.3 Random variable and its distribution

A random variable (r.v.) is a function that assigns a number to each point in its sample space. There are basically two types of r.v.s: continuous and discrete r.v.s.

Generally, if a r.v. can take on any value in the sample space, it is called a continuous r.v.; otherwise, called a discrete r.v.

For every real number \( x \), the cumulative distribution function (c.d.f.) \( F_{\tilde{X}}(x) \) of a r.v. \( \tilde{X} \) is given by:

\[
F_{\tilde{X}}(x) = P\{\tilde{X} \leq x\}
\]  

and if \( \tilde{X} \) is continuous, \( F_{\tilde{X}}(x) \) can be defined in terms of its probability density function (p.d.f.) \( f_{\tilde{X}}(x) \) as follows:

\[
F_{\tilde{X}}(x) = \int_{-\infty}^{x} f_{\tilde{X}}(x) \, dx
\]

If \( \tilde{X} \) is discrete with possible values \( x_i \), its c.d.f. is defined as:

\[
F_{\tilde{X}}(x) = \sum_{x_i \leq x} p_i
\]

where \( p_i \) is the probability corresponding to the possible value \( x_i \): \( p_i = P\{\tilde{X} = x_i\} \).

It is noted that, in probability theory and statistics, probability mass function (p.m.f.) is usually used for discrete r.v.s instead of p.d.f.

The p.m.f. is:

\[
f_{\tilde{X}}(x) = \begin{cases} P\{\tilde{X} = x_i\} & \text{if } x = x_i \\ 0 & \text{if } x \neq x_i \end{cases}
\]

Example 2.2: In power systems, load can vary and get any value between its minimum and maximum values so that it can be modeled as a continuous r.v. Similarly, nodal voltages and power flows are also continuous r.v.s. On the contrary, components in power system such as transmission lines, transformers, and generators have random phenomena with two, i.e., running and outage, or more states, so that their operating states can be modeled as discrete r.v.s.

2.4 Characteristic function

For a r.v. \( \tilde{X} \) with the c.d.f. \( F_{\tilde{X}}(\cdot) \), its characteristic function \( \psi_{\tilde{X}}(t) \) is given as follows [5]:

\[
\psi_{\tilde{X}}(t) = \mathbb{E}(e^{it\tilde{X}}) = \int_{-\infty}^{+\infty} e^{itx} dF_{\tilde{X}}(x)
\]
2.5. Moments and cumulants

2.5.1 Moments

Given a continuous r.v. $\tilde{X}$, the $r^{th}$ ($r \in \{1, 2, 3, \ldots\}$) order moment is defined as [5]:

$$m_{\tilde{X}}^r = \mathbb{E}(\tilde{X}^r) = \int_{-\infty}^{+\infty} x^r f_{\tilde{X}}(x) dx$$ \hspace{1cm} (2.7)

The quantity $m_{\tilde{X}} = m_{\tilde{X}}^1$ is called the expectation (mean) of $\tilde{X}$:

$$m_{\tilde{X}} = m_{\tilde{X}}^1 = \mathbb{E}(\tilde{X}) = \int_{-\infty}^{+\infty} x f_{\tilde{X}}(x) dx$$ \hspace{1cm} (2.8)

From the mean $m_{\tilde{X}}$, the $r^{th}$ order central moment is defined as:

$$\mu_{\tilde{X}}^r = \mathbb{E}[(\tilde{X} - m_{\tilde{X}})^r] = \int_{-\infty}^{+\infty} (x - m_{\tilde{X}})^r f_{\tilde{X}}(x) dx$$ \hspace{1cm} (2.9)

The second central moment $\mu_{\tilde{X}}^2$ is called the variance and it is usually denoted by $\sigma_{\tilde{X}}^2$, where $\sigma_{\tilde{X}}$ is the standard deviation of $\tilde{X}$.

If $\tilde{X}$ is discrete with $\nu$ possible values, the $r^{th}$ order moment is:

$$m_{\tilde{X}}^r = \mathbb{E}(\tilde{X}^r) = \sum_{i=1}^{\nu} p_i x_i^r$$ \hspace{1cm} (2.10)

where $p_i$ is the probability corresponding to value $x_i$ of $\tilde{X}$. The mean is consequently defined as in (2.11):

$$m_{\tilde{X}} = m_{\tilde{X}}^1 = \mathbb{E}(\tilde{X}) = \sum_{i=1}^{\nu} p_i x_i$$ \hspace{1cm} (2.11)

Analogously, the $r^{th}$ order central moment is:

$$\mu_{\tilde{X}}^r = \sum_{i=1}^{\nu} p_i (x_i - m_{\tilde{X}})^r$$ \hspace{1cm} (2.12)

Central moments are used more often than moments, because they relate only to the spread and shape of the distribution, rather than to its location.
2.5.2 Cumulants

The cumulants of a probability distribution are an alternative to the moments in representing r.v.s. Their algebraic properties are very interesting and allow easier calculations.

Using McLaurin’s series to expand $\ln \psi(t)$ of $\tilde{X}$, the result is:

$$\ln \psi(t) = \sum_{r=1}^{n} \frac{k_{\tilde{X}}}{r!}(jt)^r + e_n$$  \hspace{1cm} (2.13)

where $k_{\tilde{X}}$ is defined as the $r^{th}$ ($r \in \{1, 2, 3, \ldots \}$) order cumulant of $\tilde{X}$, and $e_n$ is the error of the order expansion.

Cumulants $k_{\tilde{X}}$ can be obtained from moments $m_{\tilde{X}}$ and vice versa \[21\]:

\[
\begin{align*}
\{k_{\tilde{X}} &= m_{\tilde{X}} \\
\} k_{\tilde{X}_{r+1}} &= m_{\tilde{X}_{r+1}} - \sum_{i=1}^{r} C_i r_i m_{\tilde{X}_i} k_{\tilde{X}_{r-i+1}}
\end{align*}
\]  \hspace{1cm} (2.14)

and

\[
\begin{align*}
\{m_{\tilde{X}} &= k_{\tilde{X}} \\
\} m_{\tilde{X}_{r+1}} &= k_{\tilde{X}_{r+1}} + \sum_{i=1}^{r} C_i r_i m_{\tilde{X}_i} k_{\tilde{X}_{r-i+1}}
\end{align*}
\]  \hspace{1cm} (2.15)

where $C_i = \frac{r!}{i!(r - i)!}$.

2.6 Joint moments and joint cumulants

If $N$ r.v.s $\tilde{X}_i$ ($i = 1, 2, ..., N$) are correlated, the joint p.d.f. $f_{\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N}(x_1, x_2, \ldots, x_N)$ and the joint characteristic function $\psi_{\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N}(t_1, t_2, ..., t_N)$ are presented as \[6\]:

$$\psi_{\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N}(t_1, t_2, ..., t_N) = E(e^{jt^T\tilde{X}})$$

$$= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} e^{jt^T \tilde{x}} f_{\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N}(x_1, x_2, \ldots, x_N) \, dx_1 \, dx_2 \ldots dx_N$$  \hspace{1cm} (2.16)

where $t = [t_1, t_2, ..., t_N]^T$ and $\tilde{x} = [x_1, x_2, ..., x_N]^T$.

The $r^{th}$ order joint moment is computed as \[6\]:

$$m_{\tilde{X}_1^{r_1}, \tilde{X}_2^{r_2}, ..., \tilde{X}_N^{r_N}} = E(\tilde{X}_1^{r_1} \tilde{X}_2^{r_2} \ldots \tilde{X}_N^{r_N})$$

$$= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} x_1^{r_1} x_2^{r_2} \ldots x_N^{r_N} f_{\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_N}(x_1, x_2, \ldots, x_N) \, dx_1 \, dx_2 \ldots dx_N$$  \hspace{1cm} (2.17)

where $r_1 + r_2 + \cdots + r_N = r$. 

16
2.7. Applying properties of cumulants to a linear combination of random variables

Analogously, \( \ln \psi_{\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_N}(t_1, t_2, \ldots, t_N) \) is expanded using McLaurin’s series as:

\[
\ln \psi_{\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_N}(t_1, t_2, \ldots, t_N) = \sum_{r_1, r_2, \ldots, r_N=0}^{\infty} k_{\tilde{X}_1^{r_1}, \tilde{X}_2^{r_2}, \ldots, \tilde{X}_N^{r_N}} \frac{(jt_1)^{r_1}}{r_1!}\frac{(jt_2)^{r_2}}{r_2!}\ldots \frac{(jt_N)^{r_N}}{r_N!}
\]

(2.18)

where \( k_{\tilde{X}_1^{r_1}, \tilde{X}_2^{r_2}, \ldots, \tilde{X}_N^{r_N}} \) is the \( r^{th} \) order joint cumulant of the \( N \) r.v.s.

In particular, taken as an example in case of two r.v.s \( \tilde{X}_1 \) and \( \tilde{X}_2 \), the second order joint cumulants are calculated as:

\[
\begin{align*}
k_{\tilde{X}_1^{r_1}, \tilde{X}_2^{r_2}} &= \sigma_{X_1}^{2} \\
k_{\tilde{X}_1^{r_1}, \tilde{X}_2^{0}} &= \sigma_{X_2}^{2} \\
k_{\tilde{X}_1, \tilde{X}_2} &= \rho_{\tilde{X}_1, \tilde{X}_2} \sigma_{\tilde{X}_1} \sigma_{\tilde{X}_2}
\end{align*}
\]

(2.19)

where \( \rho_{\tilde{X}_1, \tilde{X}_2} \) is the correlation coefficient between \( \tilde{X}_1 \) and \( \tilde{X}_2 \), calculated as:

\[
\rho_{\tilde{X}_1, \tilde{X}_2} = \frac{\text{Cov}(\tilde{X}_1, \tilde{X}_2)}{\sigma_{\tilde{X}_1} \sigma_{\tilde{X}_2}} = \frac{\mathbb{E}[(\tilde{X}_1 - m_{\tilde{X}_1})(\tilde{X}_2 - m_{\tilde{X}_2})]}{\sigma_{\tilde{X}_1} \sigma_{\tilde{X}_2}}
\]

(2.20)

where:

- \( \text{Cov}(\tilde{X}_1, \tilde{X}_2) \) is the covariance between \( \tilde{X}_1 \) and \( \tilde{X}_2 \),
- \( m_{\tilde{X}_1} \) and \( m_{\tilde{X}_2} \) are the means (expected values) of \( \tilde{X}_1 \) and \( \tilde{X}_2 \), respectively,
- \( |\rho_{\tilde{X}_1, \tilde{X}_2}| \leq 1 \); in particular, if \( \rho_{\tilde{X}_1, \tilde{X}_2} = 0 \), \( \tilde{X}_1 \) and \( \tilde{X}_2 \) are independent; if \( \rho_{\tilde{X}_1, \tilde{X}_2} = 1 \), \( \tilde{X}_1 \) and \( \tilde{X}_2 \) are perfectly positively dependent; if \( \rho_{\tilde{X}_1, \tilde{X}_2} = -1 \), \( \tilde{X}_1 \) and \( \tilde{X}_2 \) are perfectly negatively dependent.

2.7 Applying properties of cumulants to a linear combination of random variables

If \( \tilde{Y} \) is a r.v., a linear combination of \( N \) r.v.s \( \tilde{X}_i \) (\( i = 1, 2, \ldots, N \)):

\[
\tilde{Y} = a_1 \tilde{X}_1 + a_2 \tilde{X}_2 + \cdots + a_N \tilde{X}_N
\]

(2.21)

then the \( r^{th} \) order cumulant of \( \tilde{Y} \) is calculated, depending on the relationship of the \( N \) r.v.s \( \tilde{X}_i \), as [6]:

- when the \( N \) r.v.s \( \tilde{X}_i \) are independent:
  \[
k_{\tilde{Y},r} = a_1^r k_{\tilde{X}_1} + a_2^r k_{\tilde{X}_2} + \cdots + a_N^r k_{\tilde{X}_N}
\]
  (2.22)

- when the \( N \) r.v.s \( \tilde{X}_i \) are dependent, \( k_{\tilde{Y},r} \) can be computed by some formulas; for example, the first three cumulants are:
Chapter 2. Mathematical Background

\[ k_{\tilde{Y}_1} = \mathbb{E}(\sum_{i=1}^{N} a_i \tilde{X}_i) = \sum_{i=1}^{N} a_i k_{\tilde{X}_i} \]

\[ k_{\tilde{Y}_2} = \mathbb{E}[(\sum_{i=1}^{N} a_i \tilde{X}_i)^2] - [\mathbb{E}(\sum_{i=1}^{N} a_i \tilde{X}_i)]^2 \]

\[ = \sum_{i=1}^{N} a_i^2 k_{\tilde{X}_i} + 2 \sum_{i=1, i \neq j}^{N} a_i a_j k_{\tilde{X}_i, \tilde{X}_j} \]  

(2.23)

\[ k_{\tilde{Y}_3} = \mathbb{E}[(\sum_{i=1}^{N} a_i \tilde{X}_i)^3] - 3\mathbb{E}(\sum_{i=1}^{N} a_i \tilde{X}_i)\mathbb{E}[(\sum_{i=1}^{N} a_i \tilde{X}_i)^2] + 2[\mathbb{E}(\sum_{i=1}^{N} a_i \tilde{X}_i)]^3 \]

\[ = \sum_{i=1}^{N} a_i^3 k_{\tilde{X}_i} + 3 \sum_{i=1, i \neq j}^{N} a_i^2 a_j k_{\tilde{X}_i, \tilde{X}_j} + 6 \sum_{i=1, i < j < l}^{N} a_i a_j a_l k_{\tilde{X}_i, \tilde{X}_j, \tilde{X}_l} \cdots \]

2.8 Probability distributions most used in probabilistic analysis of electrical power systems

2.8.1 Uniform distribution

A uniform distribution is a distribution that has constant probability in a certain interval, e.g., \([a, b]\). The general formula for the p.d.f. of a uniform r.v. \(\tilde{X}\) is:

\[ f_{\tilde{X}}(x) = \begin{cases} \frac{1}{b - a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \]  

(2.24)

where:

- \(a\) is the location parameter,
- \((b - a)\) is the scale parameter.

A uniform distribution with interval \([a, b]\) is usually denoted as \(U(a, b)\). In particular, when \(a = 0\) and \(b = 1\), it is called standard uniform distribution.

The c.d.f. is:

\[ F_{\tilde{X}}(x) = \begin{cases} 0 & x < a \\ \frac{x - a}{b - a} & a \leq x \leq b \\ 1 & x > b \end{cases} \]  

(2.25)

The p.d.f. and c.d.f. of uniform distribution \(U(a, b)\) are shown in Fig. 2.1 and 2.2 respectively.
2.8. Probability distributions most used in probabilistic analysis of electrical power systems

2.8.2 Normal distribution

The normal distribution (or Gaussian distribution) is a continuous probability distribution. It is one of the most widely known and used distributions.

The p.d.f. of a normal r.v. $\bar{X}$ is:

$$f_{\bar{X}}(x) = \frac{1}{\sqrt{2\pi}\sigma_{\bar{X}}} \exp\left(-\frac{(x - m_{\bar{X}})^2}{2\sigma_{\bar{X}}^2}\right)$$  \hspace{1cm} (2.26)

where:

- $m_{\bar{X}}$ is the mean,
- $\sigma_{\bar{X}}^2$ is the variance,
- $\sigma_{\bar{X}}$ is the standard deviation.

In probability theory and statistics, a normal distribution is usually denoted by $N(m, \sigma^2)$ for short.
The c.d.f. of a normal distribution is:

\[
F_{\tilde{X}}(x) = \frac{1}{\sqrt{2\pi}\sigma_{\tilde{X}}} \int_{-\infty}^{x} \exp\left(-\frac{(t-m_{\tilde{X}})^2}{2\sigma^2_{\tilde{X}}}\right)dt
\] (2.27)

In particular, when \(m_{\tilde{X}} = 0\) and \(\sigma^2_{\tilde{X}} = 1\), the normal distribution is called standard normal distribution with its p.d.f. and c.d.f. as follows:

\[
\phi_{\tilde{X}}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)
\] (2.28)

\[
\Phi_{\tilde{X}}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}t^2\right)dt
\] (2.29)

The p.d.f.s and c.d.f.s of normal distributions, for example, with different values of \(m\) and \(\sigma\) are shown in Fig. 2.3 and 2.4, respectively.

![Figure 2.3: p.d.f.s of normal distributions](image-url)
2.8. Probability distributions most used in probabilistic analysis of electrical power systems

2.8.3 Binomial distribution

The binomial distribution is a discrete probability distribution. A binomial r.v. \( \tilde{X} \) is the number of successes \( x \) in \( n \) repeated trials of the so-called Bernoulli experiment.

Let perform a Bernoulli experiment. Assume that there are only two possible outcomes for a trial: success with probability \( p \) (the same on every trial) and failure with probability \( (1 - p) \). The trial is repeated \( n \) times. Also assume that all trials are independent, i.e., the outcome of any trial has no effect on others. If the number of occurrences of event success is \( x \), the corresponding probabilities will be \( p^x \) for event success and \( (1 - p)^{n-x} \) for event failure.

The p.m.f. of a binomial distribution is:

\[
f_{\tilde{X}}(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}
\]  

(2.30)

where \( \binom{n}{x} = \frac{n!}{x!(n - x)!} \), \( x = 0, 1, 2, ..., n \), and the c.d.f. is:

\[
F_{\tilde{X}}(x; n, p) = \sum_{i=0}^{x} \binom{n}{i} p^i (1 - p)^{n-i}
\]  

(2.31)

The mean and variance of a binomial distribution are as follows:

\[
m_{\tilde{X}} = np
\]  

(2.32)

\[
\sigma_{\tilde{X}}^2 = np(1 - p)
\]  

(2.33)

The p.m.f.s and c.d.f.s of binomial distributions, for example, with different values of \( n \) and \( p \) are depicted in Fig. 2.5 and 2.6 respectively.
Chapter 2. Mathematical Background

2.8.4 Weibull distribution

Weibull distribution is a continuous probability distribution. The \( p.d.f. \) and \( c.d.f. \) of a Weibull \( r.v. \) \( X \) are:
2.8. Probability distributions most used in probabilistic analysis of electrical power systems

\[ f_{\tilde{X}}(x) = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \exp\left[ -\left( \frac{x}{\beta} \right)^\alpha \right] \]  

(2.34)

\[ F_{\tilde{X}}(x) = 1 - \exp\left[ -\left( \frac{x}{\beta} \right)^\alpha \right] \]  

(2.35)

where \( \alpha > 0 \) and \( \beta > 0 \) are the shape and the scale parameters of the distribution, respectively.

It should be noted that the distribution is defined on a semi-bounded range \( x \geq 0 \), then \( f_{\tilde{X}}(x) = 0 \) and \( F_{\tilde{X}}(x) = 0 \) when \( x < 0 \).

The mean and variance of the Weibull distribution are as follows:

\[ m_{\tilde{X}} = \beta \Gamma(1 + \frac{1}{\alpha}) \]  

(2.36)

\[ \sigma^2_{\tilde{X}} = \beta \Gamma(1 + \frac{2}{\alpha}) - \left[ \beta \Gamma(1 + \frac{1}{\alpha}) \right]^2 \]  

(2.37)

where \( \Gamma(\cdot) \) is the well-known Gamma function.

The p.d.f.s and c.d.f.s of Weibull distributions with different values of \( \alpha \) and \( \beta \) are plotted in Fig. 2.7 and 2.8 respectively.

![Figure 2.7: p.d.f.s of Weibull distributions](image)

Figure 2.7: p.d.f.s of Weibull distributions
Chapter 2. Mathematical Background

2.9 Approximations to probability density function and cumulative distribution function of random variables

2.9.1 Approximation methods based on series expansions

The Gram-Charlier series expansion is a popular method to obtain the p.d.f. and c.d.f. of desired r.v.s from their moments or cumulants [6]. In particular, this expansion gives an approximation of a p.d.f. and c.d.f. of a continuous r.v. around the normal distribution.

Considering an arbitrary continuous r.v. \( \tilde{Y} \) with mean value \( \tilde{m}_Y \) and standard deviation \( \tilde{\sigma}_Y \), expressed in standard normalized form as \( \tilde{X} = (\tilde{Y} - \tilde{m}_Y) / \tilde{\sigma}_Y \), the c.d.f. of the standardized variable \( \tilde{X} \) can be written as:

\[
F_{\tilde{X}}(x) = \Phi_{\tilde{X}}(x) + \frac{C_1}{1!} \frac{d\Phi_{\tilde{X}}(x)}{dx} + \frac{C_2}{2!} \frac{d^2\Phi_{\tilde{X}}(x)}{dx^2} + \cdots \tag{2.38}
\]

and its p.d.f. \( f_{\tilde{X}}(\cdot) \) can be written as:

\[
f_{\tilde{X}}(x) = \phi_{\tilde{X}}(x) + \frac{C_1}{1!} \frac{d\phi_{\tilde{X}}(x)}{dx} + \frac{C_2}{2!} \frac{d^2\phi_{\tilde{X}}(x)}{dx^2} + \cdots \tag{2.39}
\]

where \( \Phi_{\tilde{X}}(\cdot) \) and \( \phi_{\tilde{X}}(\cdot) \) represent the c.d.f. and p.d.f. of standard normal distribution, respectively.

The coefficients \( C_i \) can be calculated from its central moments [6][22], for example, the first eight coefficients are calculated as follows:
2.9. Approximations to probability density function and cumulative distribution function of random variables

\[ C_1 = 0 \quad C_5 = \frac{\mu_{X_5}}{\sigma_X^5} + 10 \frac{\mu_{X_3}}{\sigma_X^3} \]
\[ C_2 = 0 \quad C_6 = \frac{\mu_{X_6}}{\sigma_X^6} - 15 \frac{\mu_{X_4}}{\sigma_X^4} + 30 \]
\[ C_3 = -\frac{\mu_{X_3}}{\sigma_X^3} \quad C_7 = \frac{\mu_{X_7}}{\sigma_X^7} + 21 \frac{\mu_{X_5}}{\sigma_X^5} - 105 \frac{\mu_{X_3}}{\sigma_X^3} \]
\[ C_4 = \frac{\mu_{X_4}}{\sigma_X^4} - 3 \quad C_8 = \frac{\mu_{X_8}}{\sigma_X^8} - 28 \frac{\mu_{X_6}}{\sigma_X^6} + 210 \frac{\mu_{X_4}}{\sigma_X^4} - 315 \]

(2.40)

Central moments can be calculated as functions of cumulants [7]:

\[
\begin{align*}
\mu_{X_2} &= k_{X_2} \\
\mu_{X_3} &= k_{X_3} \\
\mu_{X_4} &= k_{X_4} + 3k_{X_2}^2 \\
\mu_{X_5} &= k_{X_5} + 10k_{X_3}k_{X_2} \\
\mu_{X_6} &= k_{X_6} + 15k_{X_4}k_{X_2} + 10k_{X_2}^3 + 15k_3^2 \\
\mu_{X_7} &= k_{X_7} + 21k_{X_5}k_{X_2} + 25k_{X_4}k_{X_3} + 105k_{X_3}k_{X_2}^2 \\
\mu_{X_8} &= k_{X_8} + 28k_{X_6}k_{X_2} + 56k_{X_5}k_{X_3} + 35k_{X_4}^2 \\
&\quad + 210k_{X_3}k_{X_2}^2 + 280k_{X_2}k_{X_2}^2 + 105k_{X_2}^4
\end{align*}
\]

(2.41)

Therefore, if cumulants of a r.v. are known, the Gram-Charlier expansion makes it possible to approximate its c.d.f. and p.d.f. using (2.38) and (2.39). However, it is worth noticing that the necessary condition for the convergence of Gram-Charlier series is that the above c.d.f. and p.d.f. functions are continuously differentiable everywhere [8]. If this condition is not satisfied, larger errors will occur.

Beside Gram-Charlier expansion, there are two other expansions, i.e., Edgeworth and Cornish-Fisher, which are also popularly used to obtain c.d.f. and p.d.f. from moments or cumulants of a r.v. However, all these expansions give approximation around the normal distribution so they are expected to provide a good result when the input distributions are normal or close to normal. There are some common features of the three expansions: their accuracy and convergence do not necessarily increase using higher order of cumulants and moments; they do not guarantee the convergence.

In particular, the Gram-Charlier and Edgeworth expansions can give an approximate c.d.f. which does not necessarily within the range [0, 1]. Moreover, Cornish-Fisher expansion can give a better performance when working with non-normal distribution [23]. However, both Cornish-Fisher and Edgeworth expansions might have convergence problems and, especially, have larger error in the left and right tail regions of the distribution curve. The advantages and disadvantages of these expansions can also be found in [6–8]. In general, the choice of an expansion should be made depending on the distributions used.

2.9.2 Approximation method based on Von Mises function

The Von Mises function allows the definition of a discrete distribution characterized by \( \nu \) impulses starting from the first \((2\nu-1)\) moments \( m_i \) [9,24]. The method needs the
Chapter 2. Mathematical Background

following steps:
First, determinants \( D_0, D_1, \ldots, D_{\nu-1} \) are calculated:

\[
D_0 = |m_0| 
\]

\[
D_1 = \begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix} 
\]

\[
D_{\nu-1} = \begin{vmatrix} m_0 & m_1 & \cdots & m_{\nu-1} \\ m_1 & m_2 & \cdots & m_{\nu} \\ \vdots & \vdots & \ddots & \vdots \\ m_{\nu-1} & m_{\nu} & \cdots & m_{2\nu-2} \end{vmatrix} 
\]

A necessary condition for the existence of the searched discrete distribution is that all the above determinants are positive. After that step, the vector \( \mathbf{m} \) is built:

\[
\mathbf{m} = -[m_0, m_{\nu+1}, \ldots, m_{2\nu-1}]^T 
\]

and (2.46) is solved for \( c \):

\[
[D_{\nu-1}] c = \mathbf{m} 
\]

where \([D_{\nu-1}]\) is the matrix corresponding to determinant \( D_{\nu-1} \).

The solution \( c \) is then used to solve (2.47) for \( \mathbf{x} \):

\[
x_{\nu} + c_{\nu-1}x_{\nu-1} + \cdots + c_1x + c_0 = 0 
\]

Finally, the solution \( \mathbf{x} \) is necessary to compute the probabilities \( p_i \), associated to each element \( x_i \), solving (2.48), and then \( p.d.f. \) and \( c.d.f. \) can be obtained.

Based on the Von Mises function for a discrete distribution above, an approximation method was developed to obtain \( p.d.f. \) and \( c.d.f. \) for an output \( r.v. \) [24]. In this method, the continuous and discrete distributions of input \( r.v.s \) are separately treated. The continuous distribution considered here is the normal distribution but it is possible to adopt other continuous distributions [24][25].

The method is based on the following principle: assume that the output \( r.v. \) \( \tilde{R} \) is the sum of two components:

\[
\tilde{R} = \tilde{R}_c + \tilde{R}_d 
\]

where \( \tilde{R}_c \) and \( \tilde{R}_d \) are the continuous and discrete parts, respectively.

Then, the \( p.d.f. \) of \( \tilde{R} \) is:
2.10. Time series analysis

\[ f_{\tilde{R}}(x) = \sum_{i=1}^{\nu} p_i f_{\tilde{R}_c}(x - x_i) \]  

(2.50)

where \( \nu, x_i, \) and \( p_i \) are the number of impulses considered, abscissas, and corresponding probabilities of \( \tilde{R}_d, \) and \( f_{\tilde{R}_c}(\cdot) \) is the p.d.f. of the normally distributed \( \tilde{R}_c: \)

\[ f_{\tilde{R}_c}(x) = \frac{1}{\sqrt{2\pi}\sigma_{\tilde{R}_c}} \exp\left(-\frac{(x - m_{\tilde{R}_c})^2}{2\sigma^2_{\tilde{R}_c}}\right) \]  

(2.51)

where \( m_{\tilde{R}_c} \) and \( \sigma_{\tilde{R}_c} \) are the mean and standard deviation of \( \tilde{R}_c, \) respectively.

The c.d.f. of \( \tilde{R} \) is formed as:

\[ F_{\tilde{R}}(x) = \sum_{i=1}^{\nu} p_i F_{\tilde{R}_c}\left(\frac{x - x_i - m_{\tilde{R}_c}}{\sigma_{\tilde{R}_c}}\right) \]  

(2.52)

where \( F_{\tilde{R}_c}(\cdot) \) is the c.d.f. of \( \tilde{R}_c: \)

\[ F_{\tilde{R}_c}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right)dt \]  

(2.53)

A suitable value of the order of cumulants to be adopted in the computations should be selected, as it determines the maximum number of impulses that can be considered in the resulting p.d.f. and c.d.f. It is also worth noticing that, in the approximation method based on Von Mises function, the accuracy will increase by using moments of higher order. However, increasing the number of impulses makes the method more cumbersome: accuracy and speed should be balanced. On the contrary, as discussed, the approximation methods based on the Gram-Charlier, Edgeworth, and Cornish-Fisher expansions do not ensure that increasing the chosen order of cumulants and moments the accuracy is increased: this is a significant drawback of the latter approaches.

2.10 Time series analysis

Time series is a sequence of observations ordered in time, usually at equally discrete intervals, for example, 10-minute wind speed measurement data for a period of one year, hourly load measurement data for a month and so on. It is a realization of a stochastic process — an ordered set of r.v. \( \{\tilde{X}_t\}, \ t \in \{1, 2, \ldots, N\}, \) in which the realization of each \( \tilde{X}_t \) is the value that is actually observed.

Time series can be classified in several ways: most commonly as univariate and multivariate or stationary and non-stationary.

Univariate time series is referred to as a time series that consists of a single observation recorded sequentially over, normally, equal time increments. On the contrary, multivariate time series is relevant to more than one observation recorded sequentially over time. In other words, a univariate time series is a vector of observed data associated with a stochastic process, while multivariate time series is a matrix of observed data associated with multiple stochastic processes. Therefore, multivariate time series
analysis is used to explore not only the information existing in each process like in univariate time series analysis but also the interaction or correlation between different processes.

Generally, a time series, associated with a stochastic process, is said to be stationary if its statistical properties do not change with time: its parameters such as the mean, variance, auto-covariance and auto-correlation structures, etc., do not change over time too. In other words, there is no trend, no periodic variations or seasonality, no change in mean and variance in a stationary time series. In fact, most processes in nature appear to be non-stationary. However, most techniques used for time series analysis are usually applicable to stationary processes. Regarding this, generally there are two ways to deal with a non-stationary time series in practice: Developing/choosing techniques which are possible to directly deal with non-stationary processes or pre-processing the time series to obtain an approximate stationary process prior to adopting techniques for stationary processes.

The examples of various univariate time series and bivariate time series are depicted in figures from 2.9 to 2.11 and Fig. 2.12 and Fig. 2.13 respectively. Figure 2.9 presents a stationary time series with mean 1 and variance 2, while the variance of time series in Fig. 2.10 increases over time, so this time series is non-stationary. Another type of non-stationarity with trend (increase) and seasonal pattern is plotted in Fig. 2.11. In Fig. 2.12, there is no correlation between two time series, while the patterns of two time series in Fig. 2.13 are very similar: there exists correlation among them.

Figure 2.9: Stationary time series

Based on time series data, time series models can be built. Time series models are used to gain the understanding of the underlying structure or process which produced the observed data or to make a forecast for time steps ahead. There are several techniques for modeling time series. Among them, Box-Jenkins model [26] is the most popularly used method in practice. In this regard, if a time series is stationary, it can be modeled by a typical linear Auto-regressive Moving Average (ARMA) model [26]. ARMA can be used to characterize a stationary process and for prediction as well. The ARMA model consists of two subsets: Auto-Regressive (AR) and Moving Average
2.10. Time series analysis

The model is usually referred to as the ARMA\((ar, ma)\) model where \(ar\) is the order of the AR part and \(ma\) is the order of the MA part. An ARMA\((ar, ma)\) model of a stochastic process can be mathematically represented as:

\[
w_s(t) = \sum_{j=1}^{ar} \alpha_j w_s(t-j) + \varepsilon_s(t) - \sum_{l=1}^{ma} \beta_l \varepsilon_s(t-l)
\] (2.54)

where, \(\alpha_1, \alpha_2, \ldots, \alpha_{ar}\) and \(\beta_1, \beta_2, \ldots, \beta_{ma}\) are the parameters of AR and MA, respectively. The stochastic process \(\{\varepsilon_s(t)\}\) is referred to as a white noise – a serially uncorrelated, zero-mean, constant, and finite variance process \([26]\), denoted by \(WN(0, \sigma^2)\) and plotted, e.g., for \(\sigma = 1\), as in Fig. 2.14.

If \(ma = 0\), then the ARMA\((ar, ma)\) model becomes an AR\((ar)\) model. On the other hand, when \(ar = 0\), the process becomes a MA\((ma)\) model. An AR model expresses a
time series as a linear combination of its past values. The order of $ar$ tells how many lagged past values are included in the model. The MA model includes lagged terms on the noise process.

According to Box-Jenkins methodology, there are three main steps to build a time series model: model identification, model estimation, and model validation. These steps are presented in detail in [26].

It should be noted that stationarity is a necessary condition in building an ARMA model. However, this condition may not always hold with real time series data. In such a case, Auto-Regressive Integrated Moving Average (ARIMA) model [26], in which an initial differencing step applied to remove the non-stationarity, may be applied. This model is not very difficult to build for a univariate time series; however, it is very complicated in case of multivariate non-stationary time series. An alternative way to effectively solve this problem is applying pre-processing techniques to the data.
2.11. Conclusions

The fundamental concepts in probability theory and statistics are presented in this chapter. Moreover, some discussions from the theoretical point of view are also given. These knowledge will be used as well as discussed in more detail in later chapters.

Figure 2.14: White noise WN(0,1)

to obtain an approximate stationary data before building an ARMA model.
3.1 Definitions

The main task of power system operators is to maintain the system in a secure state, meanwhile producing electricity at minimum cost to get the maximum benefit. However, there is a conflict between security and economics, so that obtaining a operating point which satisfies both secure and cost optimal requirements is a very difficult problem. In other words, the ultimate aim of security assessment is to balance security and economy for power system.

There are a number of definitions of power system security. In [27], security is defined as the art and science of ensuring the survival of power systems, while in [28–30], the author defines security in terms of satisfying a set of inequality constraints (i.e., system variables, such as currents and voltages, must not exceed the limitations of physical devices) over a subset of possible disturbances called the next contingency set. According to the definition of security from the North American Reliability Council (NERC) Planning Standards, security is the ability of the electric systems to withstand sudden disturbances such as electric short circuits or unanticipated loss of system elements [31]. Alternatively, security is interpreted here as the ability of the system to withstand contingencies (e.g., the outages of generators, transmission lines, transformers, etc.), uncertainty, variations of load and power generation, for example, from renewable sources such as wind and photovoltaic solar, without any undesirable disruption of customer service and without any violation of the system operational limits, like transmission line and transformer overloading, over-/under-voltage, over-/under-regulation limits and critical ramping of generators and so on. In particular, the classification of decision drivers of power system security is presented in Fig. 3.1 [32] in which it is divided into three categories: overload, voltage, and dynamic security. For overload security, power flows through transformers and transmission lines in the...
system are calculated and assessed whether their loading is over their capacities. The voltage security can be categorized into low voltage security and unstable voltage security. Generally, voltage security is the ability of the system to operate in a stable manner, so that voltages at all the buses in the system are within the limits, and maintain system voltage stable after any disturbance. In particular, the low voltage security assesses whether nodal voltages in the power system decrease below its lower limit, while the unstable voltage security evaluates the behavior of the voltage with system load levels and voltage collapse with system load levels. In terms of dynamic security assessment, the transient performance of the system following disturbance is evaluated. Dynamic security assessment can be classified as transient instability or oscillatory instability. It examines the ability of the system to remain customer service against dynamic problems.

![Decision drivers of power system security](image)

**Figure 3.1: Decision drivers of power system security**

In power system security assessment, depending on the time frame considered, the security-related decisions are different and briefly summarized in Table 3.1 [32]. Regarding on-line security assessment, the task is to collect all necessary information about the present system state, and then compute and assess whether the system maintains stable operation in case of any disturbance occurs. For a large-scale power system, the data collected are very huge, so the computation is very intensive and difficult. In terms of operational planning, necessary facilities are first determined and then a decision on whether the existing facilities satisfy fully the security criteria or need to be strengthened is made. On the other way, for operational time frame, system operators first identify the system operational limits in which the security criteria are met fully, after that make a decision whether it is necessary to take an action to adjust the operating conditions or not [33]. With a proper security assessment for either planning or operation, after a sudden disturbance or any abnormal phenomenon occurred, the power system will survive and keep working at an acceptable steady-state condition in which all components in the power system are operated within their established technical limits, taking also into account non-linearity such as protection system, and limits.
3.1. Definitions

A power system can be operated in different modes which are associated with the security states of the power system. Power system states are divided into five states: normal, alert, emergency, in extremis, and restorative \cite{[54]}. Figure \[5.2\] describes these states and the transition from one state to another. In normal state, the system operates in a secure manner: all system variables are within their operating ranges and no equipment is being overloaded. In this state, reserve margins of both transmission and generation systems are sufficient to provide an adequate level of security with respect to the limits. If either the security level decreases below a certain threshold of adequacy or the probability of disturbance increases, the system will switch to the alert state in which all system variables are still within the operating ranges and all constraints are still satisfied; however, the system becomes weaker with a lower security level where some disturbances may result in a violation of some inequality constraints, e.g., overloading of equipments. In this case, preventive actions such as generation shifting, increased reserves, tie-line rescheduling, etc., can be taken to restore the system to the normal state. Otherwise, the system enters the emergency state or may change to the in extremis state if the disturbance is very severe. In emergency state, inequality constraints are violated; however, the system is still intact and may be restored to the alert state by emergency control actions such as fault clearing, fast-valving, dynamic braking, exciter control, load curtailment, capacitor switching, etc. If these actions are neither taken in time nor effective and if either the initiating disturbance or a subsequent one is severe enough, the system is in extremis, possibly resulting cascading outages and shut-down of some parts of the system. In this state, both system equality (i.e., the balance between the summation of total load and losses and total generation in the system) and inequality constraints are violated. Emergency control actions, such as load shedding and controlled system separation, can be done to save the system as much as possible from total collapse. If successful, the system changes to restorative state. Further control actions need to be taken to pick up all lost load and reconnect the system: the system could transit either to the alert state or to the normal state depending on the system conditions.

Traditionally, deterministic methodologies are used for power system security assessment based on deterministic criteria, i.e., \((N-k)\)-type, normally \((N-1)\)-type. Under \((N-1)\) criterion, power systems are required to be able to withstand any outage of any single element in power system such as a generator, a transmission line, a transformer, etc., without any violation of the system operational limits. However, the traditional deterministic-based approach does not consider the stochastic nature of system be-

### Table 3.1: Security-related decisions in power system security assessment

<table>
<thead>
<tr>
<th>Time frame</th>
<th>Decision maker</th>
<th>Decision</th>
<th>Basis for decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-line assessment</td>
<td>Operator</td>
<td>How to constrain the economic operation to maintain the normal state?</td>
<td>Operating rules, online assessment, and cost</td>
</tr>
<tr>
<td>Operational planning</td>
<td>Analyst</td>
<td>What should be the operating rules?</td>
<td>Minimum operating criteria, reliability, and cost</td>
</tr>
<tr>
<td>Planning</td>
<td>Analyst</td>
<td>How to reinforce/maintain the transmission system?</td>
<td>Reliability criteria for system design and cost</td>
</tr>
<tr>
<td>(Minutes to hours)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Hours to months)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Months to years)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
haviour, the probabilities of contingencies and component failures as well as sources of uncertainty affecting the power system. It therefore cannot reflect actual risks in the system. On the contrary, a promising approach, i.e., probabilistic assessment \cite{33,35}, takes into account all the above factors, so providing a better view on system security compared to deterministic approach. Both deterministic and probabilistic assessment approaches are presented and discussed in this chapter.

### 3.2 Power system security assessment

#### 3.2.1 Deterministic security assessment

In deterministic security assessment, the decision is founded on the requirement that each event of the contingency set results in system performance that satisfies the selected performance evaluation criteria. The study involves the use of a limited number of operating parameters (called study parameters), e.g., load levels, generation levels, etc., to identify the limits of operating conditions. Currently, deterministic approach is widely used to assess the security of power systems. By using this approach, a set of contingencies is first predefined and then the system is constructed to operate safely if any event in the contingency set occurs.

The deterministic approach is based on two following criteria \cite{33}:

- **Credibility:** The network configuration, outage event, and operating conditions are credible, i.e., reasonably likely to occur.

- **Severity:** The extent to which failure criteria are violated. There should be no other credible combination of outage event, system configuration, and operating condition which results in more severe system performance.

In order to construct a deterministic security boundary, the procedure for a deterministic security assessment usually includes the following main steps \cite{33}:
3.2. Power system security assessment

1. Develop base cases for power flow computation based on time horizons and loading conditions, and then select network topology (normally all circuits in service) and perform a unit commitment for each case.

2. Select the set of contingencies which comprises of credible events likely to occur in the power system. Normally, the \((N-1)\) rule, considering loss of any single component in the system such as a generator, a transformer or a transmission line, etc., is normally used to identify what is credible.

3. Select the study parameters and determine the expected range (also called study range) of operating condition for each parameter during the considered time period.

4. Identify the event/events that first violate the performance evaluation criteria as operational stress is increased within the study range. The events are referred to as the limiting contingencies. If there are no such violations within the study range, the region is not security-constrained, and the study is complete.

5. Identify the set of operating conditions within the study range where a limiting contingency first violates the performance evaluation criteria. The set of operating conditions defines a security boundary which divides the operating space into acceptable and unacceptable regions.

6. Use visual representation such as a set of plots or tables to represent the security boundary so that the operators can easily understand and use.

3.2.2 Probabilistic security assessment

Deterministic security assessment methodology is the most widely used tool in practice to check the security of power systems. However, there exist some drawbacks when using a deterministic approach: relatively higher investment and operational costs because of determining the operating limit based on the most severe condition, treating all security problems with equal probability and impact, without taking into account the probability of operating conditions, without considering sources of uncertainty affecting the system, and so on. Therefore, the deterministic methods are no more sufficient for security assessment in power system, especially in case of significant penetration of highly variable resources, such as wind and photovoltaic solar; in these conditions, probabilistic approaches need to be used.

Probabilistic approaches are becoming more attractive because they can account for the probabilistic nature of operating conditions and contingencies to actually represent the system risk. Probabilistic indices for both reliability and security assessment in planning time frame, such as loss of load expectation (LOLE), loss of load probability (LOLP), loss of energy expectation (LOEE) and expected energy not supplied (EENS), are very well-known and widely used. However, for operation-based assessment, in order to assess the security level of a power system in decision-making problems for different future time periods, operational indices should be developed. For this purpose, a probabilistic risk index was introduced in \([33][36][38]\):

\[
\text{Risk} = \sum_i \sum_j P\{E_i\} P\{X^{(j,i)} | X^{(f,i)}\} \text{Sev}(X^{(j,i)}) \quad (3.1)
\]
where:

- \( X \) is a post-contingency performance measure such as a post-contingency line loading or bus voltage level to characterize a failure condition (e.g., line overloading, over-/under-voltage, etc.). This performance measure relies on both contingency and loading conditions. The loading condition is based on the last state estimation results and a forecast.
- \( X^{(j,i)} \) is the performance measure associated with the forecast loading condition and the \( i^{th} \) contingency \( E_i \).
- \( X^{(j,i)} \) is the performance measure associated with the \( j^{th} \) possible loading condition and the \( i^{th} \) possible contingency.
- \( P\{E_i\} \) is the probability of the \( i^{th} \) contingency \( E_i \). This term represents uncertainty in outage conditions.
- \( P\{X^{(j,i)} | X^{(f,i)}\} \) is the conditional probability of \( X^{(j,i)} \) given \( X^{(f,i)} \). It represents the uncertainty in operating conditions, e.g., forecast uncertainty.
- \( Sev(X^{(j,i)}) \) is the severity function used to quantify the severity of the system condition in terms of performance indicators.

The above risk index can represent the probability of possible loading condition and contingency event, uncertainty in the system, and the severity of disturbances.

Compared to a deterministic security assessment, the procedure for a probabilistic security assessment also consists of six main steps but steps 2, 4, and 5 are different [33]. The whole procedure is as follows:

1. Develop base cases for power flow computation based on time horizons and loading conditions, and then select network topology (normally all circuits in service) and perform a unit commitment for each case.

2. Select the set of contingencies which is usually created by state enumeration process. This process is terminated by some rules such as predefined level of minimum contingency probability.

3. Select the study parameters and determine the expected range (also called study range) of operating condition for each parameter during the considered time period.

4. Evaluate the probabilistic index that reflects the level of security throughout the study range. Make a decision on a particular threshold value of the index to categorize whether the operating condition is acceptable or not.

5. Identify the set of operating conditions within the study range that have the index values equal to the threshold level. The set of operating conditions constructs a security boundary which divides the operating space into acceptable and unacceptable regions.

6. Use visual representation such as a set of plots or tables to represent the security boundary so that the operators can easily understand and use.

Probabilistic-based approach can provide valuable information for security assessment. For example, in this research, it can provide a complete spectrum of all possible values of outputs such as nodal voltages, power flows, power output of generators and
3.3. Conclusions

so on. Based on this information, the issues associated with overloading, over-/under-voltage, over-/under-regulation limits, critical ramping of conventional generators, etc., can be assessed. For example, Fig. [3.3] illustrates an output r.v $\tilde{X}$, in terms of p.d.f. obtained by a probabilistic-based approach. The full spectrum of all possible values of $\tilde{X}$ is shown together with corresponding probabilities. $\bar{x}$, $x^{up}$, and $x^{low}$ are the value of $\tilde{X}$ at operating point, upper limit, and lower limit, respectively. The striped area describes the probability so that value of $\tilde{X}$ is higher than the upper limit $x^{up}$ (e.g., the thermal rating of real power flow): $\mathbb{P}\{\tilde{X} > x^{up}\}$. This information is very useful for security assessment.

![Figure 3.3: p.d.f. of r.v $\tilde{X}$](image)

3.2.3 Probabilistic vs. deterministic security assessment

An overview of the differences between probabilistic and deterministic approaches in power system security assessment is briefly presented in Table [3.2].

3.3 Conclusions

The increase of uncertainty in power system due to the increasing penetration of renewable energy sources, such as wind and photovoltaic solar, along with the conventional sources of uncertainty, the loads and random outages of power system components, e.g., transmission lines, transformers, generators, etc., make clear the limitations of the conventional deterministic-based approach for power system security assessment. The deterministic approach is no more sufficient for security assessment under uncertainty so probabilistic-based approach is critical need. This approach is capable of characterizing uncertainty in the system to give a "full" picture about security of the system, especially in power systems with deep penetration of highly variable generation and load resources.
## Table 3.2: Probabilistic vs. deterministic security assessment

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Deterministic approach</th>
<th>Probabilistic approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effects of uncertainty on operating conditions</td>
<td>Uncertainty in operating conditions is not accounted.</td>
<td>The stochastic nature of loads and power generation such as from renewable resources, e.g., wind and photovoltaic solar, and the forecast uncertainty from renewables are taken into account.</td>
</tr>
<tr>
<td>Uncertainty in outage conditions</td>
<td>It is not considered.</td>
<td>This is represented by $\mathbb{P}{E_i}$ – the probability of the $i^{th}$ contingency $E_i$ in (3.1).</td>
</tr>
<tr>
<td>Occurrence frequency of events [39]</td>
<td>The occurrence frequency of events is not measured.</td>
<td>Each event in contingency set occurs with a certain frequency, depending on its stochastic behaviour.</td>
</tr>
<tr>
<td>Reflecting the risk associated with the insecure region</td>
<td>There is no signal of severity of risk beyond the security boundary.</td>
<td>This approach can reflect the risk associated with the insecure region.</td>
</tr>
<tr>
<td>Economical evaluation</td>
<td>The investment and operational costs might be relatively high because of determining the operating limit based on the most severe condition.</td>
<td>The security of the system is integrated into economic decision making problems to obtain a operating point which satisfies both secure and cost optimal requirements.</td>
</tr>
<tr>
<td>Use in practice</td>
<td>It is a simple approach that is easy to be used in practice. Currently, deterministic security assessment methodology is the most widely used tool in practice.</td>
<td>There are some difficulties with probabilistic approach, for example, information about stochastic phenomena acquired in power systems might be not sufficient or not credible; lack of models for renewable generation which possible capture their stochastic features, etc., so that it is slow to be used in reality. However, nowadays, probability theory and statistics are widely applied to power system especially in power systems with the integration of renewable energy sources so some of the above difficulties are not problematic now while others might be overcome in the near future.</td>
</tr>
</tbody>
</table>
4.1 Introduction

Nowadays, in order to achieve environmental and economic benefits, RES, such as wind and photovoltaic solar, are widely used. The integration of renewable resources into power systems is one of the major challenges in planning and operations of modern power systems due to their salient features, e.g., high variability, temporal and spatial correlations, diurnal and seasonal patterns and so on. Among RES, wind power generation is one of the most important and the most challenging ones. To solve problems in planning and operation of power systems involving wind resources, good quality information about future wind power production is a critical need. Such wind power production related information can be obtained from forecast techniques: the improvement of wind forecasting techniques clearly plays an important role and brings several benefits to many operators in the electricity world. For instance, for TSOs: in the day-ahead operational planning, the improvement of wind forecasting techniques makes it possible to improve the accuracy of the estimation of the operation for the next day, to reduce the needs (and the related costs) for spinning and non-spinning reserve requirements, to compute more accurately the ramping requirements by conventional generators, to increase the quality of information given to market operators, etc.; in the planning environment, this helps them to evaluate the needs for new transmission investments more accurately; for market operators: this improvement helps them to make forecasts better, to improve their risk assessment and so on.

In this chapter, wind forecasting techniques and their use in power system studies are presented and discussed. Moreover, a model for wind generation, capturing all salient features of wind, is proposed. It can provide very useful information about wind power production needed for both planning and operations of power systems involving large amounts of wind resources from multiple wind sites. This model provides a good direc-
Chapter 4. Wind Power Models for Security Assessment

tion to deal with many problems in power system involving multi-site wind production. Among them, security assessment in the power system with deepening penetration of wind resources is the focus in this thesis. Furthermore, to transform wind speed into wind power for an entire wind farm or location, a procedure for constructing an aggregate power curve is also developed.

4.2 Wind power forecast techniques and use in power system studies

Wind power forecasting is a crucial tool for managing the inherent variability and uncertainty in wind power generation. So far, several wind power forecasting methods have been reported in the literature. The tools for wind forecasting have a solid background and have been applied to power systems in many different areas. An extensive overview of wind power forecasting techniques can be found in [40]. Forecasting of wind power generation can be implemented for different time horizons [41]:

- **Very short-term**: from few seconds to 30 minutes ahead; used for electricity market clearing, regulating actions, etc.;
- **Short-term**: from 30 minutes to 6 hours ahead; used for economic load dispatch planning, load increment/decrement decisions, etc.;
- **Medium-term**: from 6 hours to 1 day ahead; used for generator online/offline decisions, operational security, day-ahead electricity market, etc.;
- **Long-term**: from 1 day to 1 week ahead; used for unit commitment, reserve requirement decisions, maintenance scheduling to obtain optimal operating cost, etc.

Generally, wind power forecast techniques are classified as **point forecast** (also called deterministic forecast) or **uncertainty forecast**.

Traditionally, wind power forecasting results are provided in the form of point forecast – a single forecasted value (expectation) for each look-ahead time step. It is therefore easy to understand the result from point forecast that is expected to express everything about future wind power production by a single value. Nowadays, point forecast is still widely used in practice. However, this technique cannot provide any information about the distribution of observations around a predicted value, i.e., the forecasting uncertainty. This is very important and needs to be estimated, beside the predicted value. In addition, with point forecast wind power forecast error is usually assumed to have a Gaussian or near Gaussian distribution. However, this assumption is not always satisfied in reality and this drawback is indicated, for example, in [42], in which an indirect algorithm based on the Beta distribution is proposed to obtain a more appropriate probability density function of wind power forecast error. In any case, the time dependency information, that is very important for dealing with decision-making problems, is absent in point forecast techniques.

In contrast, uncertainty forecast can estimate the probability distribution for each time step which represents information about the uncertainty. Uncertainty forecast can provide the results in the forms of intervals, quantiles, scenarios, etc.

- a) For interval or quantile representation, there are some main features as follows:
4.2. Wind power forecast techniques and use in power system studies

• A quantile forecast is the value so that the observation has a predefined probability to be inferior or equal to this value, while a interval forecast provides a lower and upper bound for which the observation is expected to fall within a predefined probability [43]. These representations can provide the marginal distribution for each time step.

• In these models the set of either quantiles or intervals (also called probabilistic forecast) does not capture the temporal correlation between forecast errors (i.e., uncertainty) at different time steps. The result is a somewhat misleading form of presentation, because it gives the illusion of trajectories or time sequences. In other words, at each time step, a marginal distribution is provided, regardless of what happened in the previous time steps. This means that with interval/quantile representation, the time dependency is absent [44]. This is the main drawback of these techniques.

b) On the other way, a set of scenarios – the realizations of a stochastic process or the so-called sample paths over the predefined time horizon – is known as an alternative way of wind power forecasting and effectively representing its uncertainty. Scenario representation has the following attractive features:

• A discrete set of scenarios is able to characterize the uncertainty based on exploring information existing in the data and it can provide information on the development of the prediction errors through the set of look-ahead time steps.

• The most important advantage of a scenario set is that it can include the temporal dependency information of errors. At each time step, the probability distribution of any variable depends on probability distributions of previous time steps (see next sections). In other words, the temporal correlation of forecast errors are embedded in the scenarios values [44].

The main drawback of the scenario representation is its computational burden: in order to possibly cover all of outcomes likely to occur in power system operation, a very large number of scenarios needs to be generated. Hence, an appropriate scenario clustering/selection algorithm [45–49] is necessary to obtain a reduced number of representative scenarios.

Nowadays, wind power forecasting is widely used in many important applications in power systems such as in maintaining reliability, setting the operating reserve requirements, scheduling for day-ahead market, market clearing, unit commitment, real-time dispatch, deriving the optimal selling bids for electricity market, computing power flows, security assessment, assessing ramping events and so on. Generally, depending on the application, a suitable forecast technique should be chosen or developed. Among the applications of wind power forecasting, solving a decision-making problem under uncertainty is one of the most challenging goals and this needs a significant improvement in forecasting techniques. In this regard, scenario-based approach is a promising technique, because it is capable of capturing all characteristics of wind power production and providing information needed by decision makers.
4.3 A Multi-site model for wind speed and wind power production

4.3.1 Introduction

Planning and operations of modern power and energy systems are challenging due to the integration of many sources of uncertainty: among them, wind power generation is one of the most important ones. To deal with such problems, stochastic approaches have been used as efficient tools for power system planning and operations under uncertainty [50]. They have been indicated to provide more robust solutions compared to equivalent deterministic models [51, 52]. Uncertainties on the input data of stochastic methods can be modeled by continuous or discrete stochastic processes. Stochastic approaches can be very challenging to be solved in case of continuous stochastic processes; on the contrary, they can be carried out more easily with discrete stochastic processes. Consequently, some techniques are used to approximate continuous stochastic processes by discrete ones, e.g., the so-called scenarios, before adopting stochastic methods. Scenario-based stochastic approach, treating stochastic processes via a set of scenarios, is currently known as one of the most reliable and realistic methods and is gaining wider application in power systems with the deepening penetration of wind resources [19]. Scenarios of a stochastic process are constructed as a set of realizations – so-called sample paths – over the predefined time horizon.

Building the model for multi-site wind power production is one of the most challenging issues because of the salient characteristics of wind speed and wind power. Wind power is known as a time-varying, intermittent, uncertain, and non-dispatchable resource [53]. In general, the analytic characterization of the random and time-varying wind power output is not available, because it is considerably more complicated than that of wind speed due to the highly non-linear mapping of wind speed into wind power output. Moreover, the correlation among the wind speed and therefore the wind power output at the multi-site wind farm locations brings additional layer of complexity. In addition, when wind power data are not available due to, for example, commercial reasons or in case of new wind farms, the model for wind speed is firstly built and then wind power data are derived. Wind speed is a non-Gaussian and non-stationary stochastic process with distinct diurnal and seasonal patterns. Wind speed at a specific location is temporally correlated, while wind speeds at different locations are correlated both in space and in time. All of these features make the modeling for wind speed and wind power very challenging, especially in case of large number of wind farm locations requiring high-dimensional multivariate representation.

For generating wind speed and wind power scenarios to be used in scenario-based stochastic approaches for power system planning and operations, a number of methods are proposed in the literature [17]. In [18], the authors build the model for multi-site wind speed using a noise vector that drives a vector autoregressive process. This model accounts both spatial and temporal correlations of wind speed; it assumes joint Gaussian distribution and stationarity. However, these assumptions are not always satisfied in reality. This model is improved in [19, 20] by using some techniques to obtain approximate actual wind speed data to stationary and Gaussian data. In [18, 19], and [20], the problem is simplified by the assumption that the matrix of autoregressive coefficients is diagonal. This implies that spatial correlations of wind speed at different locations are modeled fully by the underlying noise vector so that the multivariate time
4.3. A Multi-site model for wind speed and wind power production

A Multi-site model can be decoupled into different univariate time series model for each wind site.

The goal of the research in this section is to build a set of scenarios for wind modeling by using Principal Component Analysis (PCA) \([10, 11, 54, 56]\) and time series analysis \([26]\). This combination is effective in obtaining the analytical characterization of the statistical features of the spatio-temporal model of the wind output at the multi-site wind farms. PCA transforms a set of observations of correlated variables into a set of uncorrelated variables, the so-called Principal Components (PCs). A particularly attractive feature of PCA is its ability to approximate the dimensions of a data set, which is very useful in cases of high-dimensional data, such as the wind data from a large number of wind farm locations. As PCA requires Gaussian distributions, in this work, some techniques to transform the wind data (typically non-stationary and non-Gaussian) and to make data sets suitable for PCA are proposed, thus removing any simplifying assumption. The proposed methodology is, therefore, realistic and applicable to any wind data set in reality. Depending on predefined time horizon, the spatio-temporal model proposed can be adopted for the operational planning like in stochastic power flow, stochastic optimal power flow, stochastic unit commitment, operating reserve requirement and so on as well as for planning studies such as transmission expansion with multi-site wind production, planning reserve requirement, etc.

### 4.3.2 Structural representation of wind data and Principal Component Analysis

PCA performs an orthogonal transformation on data in order to transform a correlated data set into an uncorrelated one. The underlying technique is the eigenanalysis applied to a symmetrical matrix, such as the correlation matrix or the covariance matrix.

PCA applies to a matrix \(W\) whose elements \(w_s(t)\) are the data of wind speed or power available at site \(s \in \{1, 2, \ldots, S\}\) and at time \(t \in \{1, 2, \ldots, N\}\):

\[
W = \begin{bmatrix}
w_1^{(1)} & w_1^{(2)} & \cdots & w_1^{(N)} \\
w_2^{(1)} & w_2^{(2)} & \cdots & w_2^{(N)} \\
\vdots & \vdots & \ddots & \vdots \\
w_S^{(1)} & w_S^{(2)} & \cdots & w_S^{(N)}
\end{bmatrix}
\] (4.1)

\(W\) can be interpreted as a realization of r.v.s \(\tilde{W}_s(t)\) over \(N\) time steps of the random process represented in Fig. 4.1 for \(T\) time steps.

We assume that all considered wind locations have the same number of observations \(N\) spanning – for example, one year – and that they are synchronized and equally spaced in time.

At first, data should be centered \([55]\) in matrix \(W_c\) by subtracting the mean of each time series at each site:

\[
\mu_s = \frac{\sum_{t=1}^{N} w_s(t)}{N}
\] (4.2)

\[
W_c = W - \mu
\] (4.3)
Chapter 4. Wind Power Models for Security Assessment

where $\mu = [\text{diag}\{\mu_1, \mu_2, \ldots, \mu_S\}] J$, where $J$ is a $(S \times N)$ matrix of ones.

Next, the correlation or covariance matrix of the centered data $W_c$ is calculated. PCA can use either correlation or covariance matrix; however, the correlation matrix must be adopted when the considered variables are not comparable [55]. For example, this is necessary when considering the real power outputs of wind parks of different capabilities. In this case, another option is also to normalize values and adopt the correlation matrix.

In the following, the covariance matrix is considered, for the sake of simplicity.

Covariance matrix $\Sigma$ is a symmetric $(S \times S)$ matrix:

$$\Sigma = \begin{bmatrix}
\Sigma_{1,1} & \Sigma_{1,2} & \cdots & \Sigma_{1,S} \\
\Sigma_{2,1} & \Sigma_{2,2} & \cdots & \Sigma_{2,S} \\
\vdots & \vdots & \ddots & \vdots \\
\Sigma_{S,1} & \Sigma_{S,2} & \cdots & \Sigma_{S,S}
\end{bmatrix} \quad (4.4)$$

where, $\Sigma_{i,j}$ is the covariance between the time series at site $i$ and the time series at site $j$; $\Sigma_{i,j} = \sigma^2_{i,j}$ and $\Sigma_{i,i} = \sigma^2_i$ is the variance ($\sigma$: standard deviation).

$$\Sigma_{i,j} = \frac{1}{N} \sum_{t=1}^{N} (w_i^{(t)} - \mu_i)(w_j^{(t)} - \mu_j) \quad (4.5)$$

The covariance matrix $\Sigma$ is a symmetric, positive, semi-definite matrix and all its eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_S$ are positive. Eigenvalues are ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_S$.

The $(S \times S)$ matrix $U$ is formed by the corresponding eigenvectors $u_1, u_2, \ldots, u_S$:

$$U = [u_1 | u_2 | \cdots | u_S] \quad (4.6)$$

Its elements are also known as PC coefficients. Finally, the PCs are derived [55] as

$$Z = U^T W_c \quad (4.7)$$

where $Z$ is a $(S \times N)$ matrix. The $i$-th row of matrix $Z$, $z_i$, is the $i$-th PC, that is a time series univariate and uncorrelated with other PCs. The techniques to characterize such a time series are reported in [26].
4.3. A Multi-site model for wind speed and wind power production

The reconstruction of wind data from PCs is implemented inversely:

\[
W = \mu + UZ
\]

\[
= \mu + [u_1 \ | \ u_2 \ | \ \cdots \ | \ u_S] [z_1^T \ | \ z_2^T \ | \ \cdots \ | \ z_S^T]^T
\]  

(4.8)

It is worth noting that PCA works properly under the assumption that the underlying distribution is multivariate Gaussian [55]. If the data set is not multivariate Gaussian, PCs will be still correlated, although their covariance matrix is diagonal. In fact, for multivariate non-Gaussian distribution, the first and second order statistics do not characterize totally the distribution. In this case, robust PCA is a possible approach [55]. In addition, outlier detection should be applied to the data prior to adopting PCA [11]. In the present work, an alternative approach is adopted: pre-processing techniques are used to obtain Gaussian data sets (see Section 4.3.3). This is a significant improvement in using PCA as confirmed by the simulation results in Section 4.3.4.

A very interesting feature of PCA, beside decorrelation, is that it provides an excellent tool to approximate large data set by reducing its dimension [54]. This function makes PCA a powerful tool for high-dimensional data analysis. As it derives from eigenanalysis, each PC, \(z_m\), can be seen as a mode, whose importance is weighted by the relevant eigenvalue \(\lambda_m\). Therefore, the importance of the \(m\)-th PC can be computed as:

\[
\gamma_m = \frac{\lambda_m}{\sum_{i=1}^{S} \lambda_i} \times 100\%
\]

(4.9)

and the cumulative contribution of the first \(m\) PCs is:

\[
\Gamma_m = \sum_{i=1}^{m} \gamma_i
\]

(4.10)

The first row vector \(z_1\) corresponding to the largest eigenvalue \(\lambda_1\) and eigenvector \(u_1\) is hence the most important component (dominant component) which contains most of the information in the data set, followed by the second component \(z_2\) and so on. If the first \(K\) (\(K < S\)) components are considered, \(W\) will be approximated as

\[
\hat{W} = \mu + [u_1 \ | \ u_2 \ | \ \cdots \ | \ u_K] [z_1^T \ | \ z_2^T \ | \ \cdots \ | \ z_K^T]^T
\]

(4.11)

This is also a very important use of PCA, which is able to describe most of the features of the data set by fewer variates.

4.3.3 Proposed methodology

In this section, the proposed approach to capture all characteristics of wind data from multiple sites is discussed in detail and the spatio-temporal model is built. The algorithm is described in the flow diagram in Fig. 4.2.

The input is observed wind speed or wind power data (in time series) for each location, in the form (4.1). The process is implemented step by step as follows:

**Step 1:** Various pre-processing techniques are applied to (4.1) for obtaining stationarity, i.e., to remove diurnal and seasonal effects (and also any other trend, if
Input
Observed wind speed at each location

Step 1
Pre-processing and Transformation

Step 2
Spatial decorrelation using PCA

Step 3
Characterizing PCs

Step 4
Generating time series of PCs

Step 5
Reconstruction of data from generated PCs

Step 6
Inverse of pre-processing and Retransformation

Output
Scenario set of wind speed

Figure 4.2: The flow diagram of the proposed approach

\[
\begin{align*}
    w_s^{(t)} &= \frac{w_s^{(t)} - \mu_s^{(t)}}{\sigma_s^{(t)}} \\
    w_s^{(t)} &= \Phi^{-1}[\hat{F}_s(w_s^{(t)})]
\end{align*}
\]

where \( \mu_s^{(t)} \) and \( \sigma_s^{(t)} \) are the mean and standard deviation at site \( s \) and time \( t \) for period \( p \) such as month, season, etc., which is selected based on the periodic features of the data. The resulting stationarity must be assessed by a statistic test on \( \{w_s^{(t)}\} \). In this research, Augmented Dickey-Fuller (ADF) test [59] and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test [60] are used. Should the \( \{w_s^{(t)}\} \) not pass the test, a different pre-processing must be carried out. As the obtained stationary data set could still be non-Gaussian, \( \{w_s^{(t)}\} \) is then transformed into Gaussian data set [18] by:

\[
    w_s^{(t)} = \Phi^{-1}[\hat{F}_s(w_s^{(t)})]
\]

where \( \hat{F}_s(\cdot) \) is the estimated c.d.f. of the stationary process associated with \( \{w_s^{(t)}\} \) and \( \Phi^{-1}(\cdot) \) is the inverse of the c.d.f. of the standard normal distribution (Section 2.8.2). Analogously, statistical tests, to assess if the resulting distribution is Gaussian, Lilliefors goodness-of-fit test [61] and Jarque-Bera hypothesis test [59] are carried out.

**Step 2:** PCA is adopted according to Section 4.3.2. Each PC is a univariate time
4.3. A Multi-site model for wind speed and wind power production

series, not cross-correlated with other PCs. The data, however, still contain temporally correlated information.

**Step 3:** Each PC time series is fitted by a time series model according to Section 2.10, as PC time series data herein satisfy all necessary conditions for building a time series model like ARMA.

**Step 4:** The obtained time series model for each PC is then used to generate an adequate number of time series for future time; for example, time frames of operations (e.g., 6 hours, 24 hours ahead, etc.) and/or planning (e.g., weeks, months, years ahead, etc.).

**Step 5:** The generated time series in terms of PCs are reconstructed by using (4.8). If dimensional approximation is desired, (4.11) can be used.

**Step 6:** The obtained data from Step 5 are post-processed and then the items removed in the pre-processing step are added back, to obtain scenarios obeying all the characteristics of the observed wind data for each location.

The outputs of the procedure are time series, i.e., scenario sets, of wind data over the predefined time horizon for each location. This means modeling a stochastic process as in Fig 4.1 where T changes according to the time frame to be considered. The proposed methodology can explicitly capture the salient features of the wind data:

- **The spatial correlation:** by PCA technique,
- **The temporal correlation:** by time series model (like ARMA),
- **Diurnal and seasonal non-stationarity and non-Gaussianity:** by pre- and post-processing and transformation techniques.

So far, a large number of time series of wind data have been generated. To solve this problem, it is therefore necessary to choose or develop a clustering/selection technique for obtaining a smaller number of representative time series, thus reducing computation time. Some methods such as backward reduction, forward selection, etc., using probability metrics, e.g., Kantorovich distance, presented in [45–47], can be used. This problem is, however, beyond the scope of the present research and will be considered in the future work.

4.3.4 Tests and results

In this section, the proposed approach is applied to a couple of case studies relevant to 10-minute wind speed in the Italian region of Basilicata and hourly wind speed from different sites in Sicily, Italy.

**Case 1: Wind speed in the Italian region of Basilicata**

1. Wind data

Basilicata is a region in the south of Italy which covers about 10000 km² and includes two provinces: Potenza and Matera. Figure 4.3 describes five wind locations in the region of Basilicata: Forenza, Pietragalla, Vaglio, Calvello, and Matera. For the sake of simplicity, the locations are denoted as F, P, V, C, and M, respectively. 10-minute wind speed measurement data from March 1, 2001 to February 28, 2002,
Table 4.1: Covariance matrix of observed wind speed data from five locations in Basilicata

<table>
<thead>
<tr>
<th>location</th>
<th>F</th>
<th>P</th>
<th>V</th>
<th>C</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>19.83</td>
<td>12.16</td>
<td>9.85</td>
<td>10.96</td>
<td>5.31</td>
</tr>
<tr>
<td>P</td>
<td>12.16</td>
<td>12.31</td>
<td>8.50</td>
<td>9.09</td>
<td>3.01</td>
</tr>
<tr>
<td>V</td>
<td>9.85</td>
<td>8.50</td>
<td>9.48</td>
<td>6.51</td>
<td>2.87</td>
</tr>
<tr>
<td>C</td>
<td>10.96</td>
<td>9.09</td>
<td>6.51</td>
<td>14.44</td>
<td>3.63</td>
</tr>
<tr>
<td>M</td>
<td>5.31</td>
<td>3.01</td>
<td>2.87</td>
<td>3.63</td>
<td>8.06</td>
</tr>
</tbody>
</table>

provided by RSE (Ricerca sul Sistema Energetico) S.p.A., Italy [62], resulting in 52560 samples, are used.

The observed wind speed from five locations is plotted in Fig. 4.4. The covariance matrix of observed wind speed is given in Table 4.1. It indicates that wind speed at five locations are strongly correlated. This makes the work of building a spatio-temporal model for such data very complicated. The scatter plots of wind speed between selected locations are also presented in figures from 4.5 to 4.7.

Figure 4.3: Wind locations in the region of Basilicata in Italy

2. Simulation results

As discussed, PCA and time series model like ARMA work properly if the data used are stationary and Gaussian. According to Step 1 in Section 4.3.3, the initial data set was pre-processed and passed the stationarity test (transformed stationary data as in Fig. 4.8); eventually, transformation (4.13) was applied and a stationary and Gaussian set was obtained and passed the relevant tests. The c.d.f.s of the stochastic processes associated with the data before and after using (4.13), for example, for location F are compared in Fig. 4.9.
4.3. A Multi-site model for wind speed and wind power production

It is worth noticing that, so far, techniques in statistics have been used to obtain the data set associated with a stationary Gaussian process for each location without using any assumption. This is a particularly attractive feature of the proposed methodology, i.e., the method can be used for any real wind data.

The refined data from all locations are then transformed into PCs by using (4.7). The plots of the time series relevant to each PC are shown in Fig. 4.10. The contribution of
each PC and the cumulative contribution are calculated by (4.9) and (4.10), respectively, and presented in Table 4.2. As can be seen from Fig. 4.10, PCs are quite different in terms of magnitudes. The first PC ($z_1$) contains the largest percentage of information in the data set (65.08%); the second PC ($z_2$) the second largest percentage (16.51%) and so on. In this case, the first two PCs account for 81.58% of the information and can be used as an approximation. This is also a very important aspect of the proposed model.
4.3. A Multi-site model for wind speed and wind power production

After applying PCA, the resulting covariance matrix of PC time series is $\Sigma = \text{diag}\{3.24; 0.82; 0.48; 0.28; 0.16\}$ showing that the PCs are uncorrelated; the same information can be deduced from Fig. 4.11 which depicts the scatter plot of $z_1$ and $z_2$. If the proper pre-processing and transformation had not been applied, the cross-

Figure 4.8: Transformed stationary data of five locations

Figure 4.9: c.d. f.s before and after using Gaussian transform for location $F$

so as to make it a very effective tool to reduce the size of very high-dimensional data sets.

After applying PCA, the resulting covariance matrix of PC time series is $\Sigma = \text{diag}\{3.24; 0.82; 0.48; 0.28; 0.16\}$ showing that the PCs are uncorrelated; the same information can be deduced from Fig. 4.11 which depicts the scatter plot of $z_1$ and $z_2$. If the proper pre-processing and transformation had not been applied, the cross-
Chapter 4. Wind Power Models for Security Assessment

Figure 4.10: The construction of five PCs

Table 4.2: The contribution of five PCs

<table>
<thead>
<tr>
<th>PC - m</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_m(%)$</td>
<td>65.08</td>
<td>16.51</td>
<td>9.67</td>
<td>5.55</td>
<td>3.20</td>
</tr>
<tr>
<td>$\Gamma_m(%)$</td>
<td>65.08</td>
<td>81.58</td>
<td>91.25</td>
<td>96.80</td>
<td>100.00</td>
</tr>
</tbody>
</table>

correlation would still have existed between PCs, as in Fig. 4.12 compared to Fig. 4.11.

Figure 4.11: Scatter plot of $z_1$ and $z_2$

The time series built on the PCs fulfill the assumptions for a successful use of the
4.3. A Multi-site model for wind speed and wind power production

Figure 4.12: Scatter plot of $z_1$ and $z_2$ in case of without using pre-processing and transformation techniques

ARMA models. The models adopted for PCs $z_1$, $z_2$, $z_3$, $z_4$, and $z_5$ are AR(4), AR(7), AR(10), AR(11), and AR(14), respectively. They were selected according to [26] and assessing the results by the residual (error) test. As an example, Fig. 4.13 illustrates the resulting residual test for $z_1$ model. As the residual is a white noise, the AR(4) model of $z_1$ is valid. The same process was carried out for the other PCs.

Figure 4.13: Residual test for time series model of $z_1$

The models of PCs are then used to generate a large number of time series for future time instants. In this study, 1000 time series with 10-minute resolution and different
time frames of both operations and planning of power system are implemented. After that, the generated data are reconstructed and the inverse process is used to obtain wind speed data.

In order to illustrate a possible use of the proposed methodology in the operational planning environment, the time horizon of six hours ahead is considered. One thousand scenarios are generated; each scenario represents the wind speed behaviour at each site according to the spatio-temporal correlations captured by the model. From the generated wind speed scenarios, first at each location (e.g., \(H\)) and each time step (e.g., \(\tau\)), the \(p.d.f. \ P_H(\tau)\) and \(c.d.f. \ C_H(\tau)\) could be obtained and could be correlated to \(p.d.f. \ P_L(t)\) and \(c.d.f. \ C_L(t)\) at other locations \(L\) and other times \(t\), thanks to the association in scenarios. This is the main contribution of this research, that can be directly exploited in probabilistic power flows (see Section 5.2 and Section 5.3 in Chapter 5) to assess the power system security for the next hours. Figure 4.14 presents, for example, the histogram and \(c.d.f.\) of wind speed at the time step of 30 minutes ahead for location \(F\) (observe that they are not Gaussian). These distributions are, of course, related to other distributions at both the same time step (at other sites) and different time steps (at both the same site and other sites) by the spatio-temporal structure proposed.

Also in the planning environment, probabilistic power flows can be adopted to study new network reinforcement taking into account spatio-temporal correlations, which allow reducing investments compared to other probabilistic traditional planning methods.

Moreover, the proposed model can provide a very appropriate input for solving decision-making problems \([50,63]\), especially under uncertainty, for both operating horizon (e.g., scheduling of electricity markets) and planning horizon (e.g., investment planning) in power systems involving multi-site wind production.

Another possible use of the proposed method is the improvement of wind power production forecast and its profile along the network, by means of the association of the current weather conditions provided by wind forecast to a subset of scenarios among those generated by the method. This tool, however, is currently under development and is over the purpose of the present research.

**Case 2: Wind speed from different locations in Sicily, Italy**

The proposed model is able to apply for different resolutions, e.g., 10 minutes, 15 minutes, one hour, etc., depending on the data available. In this case, hourly wind speed data are used.

**1. Wind data**

Hourly wind speed data from September 1, 2011 to August 31, 2012 measured at nine locations belonging to Sicily, provided by TERNA (the Italian TSO), are used. Analogously, for the sake of simplicity, the locations are denoted as \(L_1, L_2, \ldots, L_9\). The data are depicted in Fig. 4.15 while the covariance matrix of the wind speed is given in Table 4.3. The scatter plots of wind speed between selected locations are also shown in figures from 4.16 to 4.20 indicating that they are strongly correlated.
4.3. A Multi-site model for wind speed and wind power production

Figure 4.14: Histogram and c.d.f. of wind speed at the time step of 30 minutes ahead for location F

Table 4.3: Covariance matrix of observed wind speed data from nine locations in Italy

<table>
<thead>
<tr>
<th>location</th>
<th>$L_1$</th>
<th>$L_2$</th>
<th>$L_3$</th>
<th>$L_4$</th>
<th>$L_5$</th>
<th>$L_6$</th>
<th>$L_7$</th>
<th>$L_8$</th>
<th>$L_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_2$</td>
<td>5.62</td>
<td>10.09</td>
<td>5.20</td>
<td>4.08</td>
<td>6.68</td>
<td>7.59</td>
<td>4.93</td>
<td>6.33</td>
<td>6.10</td>
</tr>
<tr>
<td>$L_3$</td>
<td>8.19</td>
<td>5.20</td>
<td>9.99</td>
<td>5.90</td>
<td>4.79</td>
<td>5.18</td>
<td>6.47</td>
<td>3.79</td>
<td>7.04</td>
</tr>
<tr>
<td>$L_4$</td>
<td>9.14</td>
<td>4.08</td>
<td>5.90</td>
<td>11.33</td>
<td>4.21</td>
<td>4.74</td>
<td>9.87</td>
<td>2.36</td>
<td>7.16</td>
</tr>
<tr>
<td>$L_5$</td>
<td>5.49</td>
<td>6.68</td>
<td>4.79</td>
<td>4.21</td>
<td>13.27</td>
<td>8.21</td>
<td>46.0</td>
<td>7.19</td>
<td>5.37</td>
</tr>
<tr>
<td>$L_6$</td>
<td>6.76</td>
<td>7.59</td>
<td>5.18</td>
<td>4.74</td>
<td>8.21</td>
<td>13.98</td>
<td>6.06</td>
<td>8.54</td>
<td>6.79</td>
</tr>
<tr>
<td>$L_7$</td>
<td>9.62</td>
<td>4.93</td>
<td>6.47</td>
<td>9.87</td>
<td>4.60</td>
<td>6.06</td>
<td>13.36</td>
<td>3.40</td>
<td>7.58</td>
</tr>
<tr>
<td>$L_8$</td>
<td>4.13</td>
<td>6.33</td>
<td>3.79</td>
<td>2.36</td>
<td>7.19</td>
<td>8.54</td>
<td>3.40</td>
<td>10.92</td>
<td>5.03</td>
</tr>
<tr>
<td>$L_9$</td>
<td>8.44</td>
<td>6.10</td>
<td>7.04</td>
<td>7.16</td>
<td>5.37</td>
<td>6.79</td>
<td>7.58</td>
<td>5.03</td>
<td>12.11</td>
</tr>
</tbody>
</table>

2. Simulation results

First, data set of each location is pre-processed as presented in Step 1 in Section 4.3.3 to obtain transformed stationary data set. However, all resulting data sets are still non-Gaussian, shown in Fig. 4.21 for all locations. These data sets are then transformed into Gaussian ones using (4.13); Fig. 4.22 is depicted, for example, for location $L_1$.

After using pre-processing and transformation techniques, PCA is applied to spatially de-correlate the resulting data, obtaining PC time series as in Fig. 4.23. As can be seen from the figure, the magnitudes of PCs are quite different showing that the contribution of PCs into capturing information existing in the data are different, as shown in Table 4.4.

Each PC time series is uncorrelated to others and the relevant covariance matrix is:

$$\Sigma = diag\{0.41; 1.54; 0.71; 0.53; 0.46; 0.42; 0.37; 0.32; 0.20\}$$

Adopting the proposed method step by step, finally, the wind speed scenarios for a desired time frame (of both operating and planning) for each location can be obtained.
As discussed, the proposed method can work well with high-dimensional data sets. Instead of using all components, dimensional approximation is considered here: the first $K$ components are used as in (4.11). In this test, 500 scenarios are generated for day-ahead horizon. Figure 4.24 shows the residuals of dimensional approximation, for example, for location $L_5$, in which $R^K_{L_5}$ denotes the residual when the first $K$ components used. Residual $R^K_{L_H}$ for location $L_H$ is calculated: $R^K_{L_H} = W_{L_H} - \hat{W}^K_{L_H}$, where $W_{L_H}$ and $\hat{W}^K_{L_H}$ presented in (4.8) and (4.11), respectively. As can be seen from Fig. 4.24, if the first six components, for example, are used, the corresponding residual will be very small: this means that the information lost in this case is negligible. In
4.3. A Multi-site model for wind speed and wind power production

![Figure 4.16: Scatter plot of observed wind speed for locations L1 and L3](image)

![Figure 4.17: Scatter plot of observed wind speed for locations L2 and L6](image)

other words, one dimension of data set, i.e., space, (in total of two dimensions, i.e.,
time and space, in this case) can be reduced from nine to six. This is a very attractive
feature of the proposed method to deal with high-dimensional data.
Figure 4.18: Scatter plot of observed wind speed for locations $L_4$ and $L_9$

Figure 4.19: Scatter plot of observed wind speed for locations $L_5$ and $L_6$
4.3. A Multi-site model for wind speed and wind power production

Figure 4.20: Scatter plot of observed wind speed for locations $L_2$ and $L_8$
Figure 4.21: c.d.f.s of transformed stationary data at nine locations
4.3. A Multi-site model for wind speed and wind power production

Figure 4.22: c.d.f.s before and after using (4.13) for location $L_1$
Figure 4.23: PC time series
4.3. A Multi-site model for wind speed and wind power production

Figure 4.24: Residuals of dimensional approximation for location L₅.
4.4 Wind power curve

The power output from a single wind turbine can be determined from wind speed by the so-called wind turbine power curve – the relationship between wind power output and wind speed across the turbine blades. With this curve, power output can be obtained without detailed technical knowledge of the wind turbine and its components. Figure 4.25 describes a typical wind turbine power curve in which:

- $v_i$ is the cut-in wind speed – the minimum wind speed at which the turbine blades overcome friction and begin to rotate for generating wind power;
- $v_o$ is the cut-out wind speed – the speed at which the turbine blades are brought to rest for preventing damage to the turbine when the wind is too fast and beyond that speed;
- $v_r$ is the rated wind speed;
- $P_r$ is the rated power output, corresponding to $v_r$.

**Figure 4.25: Typical wind turbine power curve**

From Fig. 4.25, it can be seen that there are four parts in the wind turbine power curve:

- $v < v_i$: standby, no power production generated;
- $v_i \leq v \leq v_r$: working with non-linear wind power-wind speed relationship;
- $v_r < v \leq v_o$: rated power region, power output kept at constant level of $P_r$;
- $v > v_o$: cut-out part, wind generator disconnected for protection.
4.4. Wind power curve

For a single wind turbine, power curve is given by its manufacturer [64,65]. However, this curve can be used only when the operating conditions are ideal [1]: air density must be standard; wind speed should be uniform horizontally across the face of the turbine; vertical wind speed profile should be the same as that experienced during the calibration of the turbine; and the wind speed cannot be influenced by obstacles. These conditions are, of course, very unlikely to be satisfied in reality so the "real" power curve should be estimated.

The power curve for wind turbine can be estimated by a quadratic [64,66] or a cubic model [67–69]. Also, it can be estimated by using measurement data of wind power-wind speed pairs which are widely used in practice. The procedure for construction of a power curve for a wind turbine is described as a standard, i.e., IEC 61400-12, by International Electrotechnical Commission [70]. In [1] and [70], "the method of bins" was investigated and used. In this approach, wind speed is divided into bins of 0.5 m/s for wind speed between 0 m/s and the cut-out speed. For each bin, the average wind speed and average wind power are computed and used as an estimated point. In addition, some techniques are applied to rejected erroneous data to improve the estimation of power curve. Erroneous data usually exist in any measurement data set. Figure 4.26 shows an example of plotting 10-minute wind power versus wind speed measurement for a real wind turbine [1]. In Fig. 4.26 wrong data can be divided into four zones: zone 1 is caused by shade effect (due to the presence of obstacles); zone 2 corresponds to abnormal working states of the turbine, e.g., due to pitch misoperation; data in zones 3 and 4 are probably bad but reasons are not clear because of many effects. In fact, the quality of measurement data is affected by many factors [1] such as accuracy of sensors, electro-magnetic interferences, information processing errors, storage or communication fault, alarms, faults in control system, and so on.

![Figure 4.26: Measured wind power against measured wind speed for a real wind turbine](image)

67
Chapter 4. Wind Power Models for Security Assessment

For a wind farm or location with many wind turbines (from tens to hundreds of wind turbines), the problem of wind turbine power curve estimation becomes much more complicated due to additional effects between wind turbines. Moreover, due to the fact that for an entire wind farm or location, aggregate wind power production needs to be estimated for power system planning and operations, an aggregate power curve is necessary. In other words, instead of estimating for a single wind turbine as discussed above, estimating for an aggregate power curve is required here. In reality, the relationship between wind speed measured for a wind farm and its aggregate power output is more complicated than a simple transformation based on power curve of each turbine. In fact, even in a farm with all identical turbines the output powers from its turbines would not be necessarily equal.

In order to build an equivalent power curve for a wind farm or location based on wind speed-wind power measurement data, so far various methods have been developed and used in practice [20, 70–74]. Among them, curve fitting [20] and method of bins [70] are widely used.

The method of bins is considered and developed in this section. The method needs the following steps:

**Step 1:** Due to erroneous values existing in the measurement data of wind power-wind speed pairs, the “first” filtering process should be carried out. Some criteria are proposed to eliminate spurious data points; data points falling into the following cases (do not represent the normal working states or are caused by wrong measurement and other effects as discussed above) need to be filtered:

- data points that do not coincide with the number of turbines available for generation (e.g., power output measured is greater than \( \sum_{l=1}^{N_t} P_{r,l} \), where \( P_{r,l} \) is the rated power output of turbine \( l \) and \( N_t \) is the total number of turbines available);

- data points at very high wind speed with the corresponding power outputs not equal to zero (e.g., if all turbines are identical and when wind speeds are greater than the cut-out speed of these turbines, the power outputs should be equal to zero because all these turbines are not in operation in this condition; if cut-out speed of turbines are different, the greatest one will be used to specify the very high wind speed region);

- data points corresponding to very low wind speed (similarly to the above, cut-in wind speed is used to specify this region) and positive power output (but in standby region, no power generated);

- data points corresponding to power output of zero and wind speed within normal operation region;

- data points with constant wind speed over a long period...

**Step 2:** Wind speed is divided into bins, for example, of 0.5, 1.0, 1.5, etc., m/s. For each bin, the average wind speed and average wind power are computed as follows [70]...
4.4. Wind power curve

\[
\bar{v}_k = \frac{\sum_{l=1}^{N_k} v_{k,l}}{N_k} \quad (4.14)
\]

\[
\bar{p}_k = \frac{\sum_{l=1}^{N_k} p_{k,l}}{N_k} \quad (4.15)
\]

where:

- \(\bar{v}_k\) is average wind speed of bin \(k\);
- \(\bar{p}_k\) is average wind power of bin \(k\);
- \(v_{k,l}\) is the value of the \(l\)-th point of wind speed in bin \(k\);
- \(p_{k,l}\) is the value of the \(l\)-th point of wind power in bin \(k\);
- \(N_k\) is the total number of points in bin \(k\).

**Step 3:** The "second" filtering process is then implemented. The standard deviation of the power in each bin is calculated:

\[
\sigma_{p,k} = \sqrt{\frac{\sum_{l=1}^{N_k} (p_{k,l} - \bar{p}_k)^2}{N_k}} \quad \forall k \quad (4.16)
\]

A filtering criterion is then applied [70]: data points deviating more than \(\xi \sigma_{p,k} \) (\(\xi \in \mathbb{R}, \xi \) chosen) can be rejected as outliers and by repeating Step 2 and Step 3 (number of times selected) the accuracy can be increased.

**Step 4:** The aggregate power curve is finally obtained based on fitting a curve to \((\bar{v}_k, \bar{p}_k)\) pairs.

Taken as an example to illustrate the above developed approach, measurement data of wind speed-wind power pairs for locations presented in Case 2, Section 4.3.4 are used (hourly wind power measurement data also provided by Terna). Figures from 4.27 to 4.31, for example, show the measurement data for locations \(L_1\), \(L_3\), \(L_5\), \(L_7\), and \(L_8\), respectively. Generally, it is clear that the aggregate power curve for each location may have a general shape that is similar to the power curve of a single turbine (see Fig. 4.25). Also, there are many erroneous data existing in the measurement data and need to be rejected.

In this test, bins of 1m/s are adopted. After using the "first" filtering process, Step 2 and Step 3 are repeated 100 times. \(3\sigma_k\) is chosen in this case so that values of wind power in each bin \(k\) outside the range of \([\bar{p}_k - 3\sigma_k, \bar{p}_k + 3\sigma_k]\) are rejected. Following the above process step by step, the aggregate power curves for locations \(L_3\) and \(L_5\), for example, are finally obtained as shown in figures 4.32 and 4.33, respectively, in which the range \([\bar{p}_k - 3\sigma_k, \bar{p}_k + 3\sigma_k]\) corresponds to the lower and upper bounds.

In practice, the procedure for constructing a power curve requires data analysis, i.e., checking, detecting, filtering the data and so on, to reduce the effects of erroneous measurement data. The amount of measurement data for wind speed and wind power is
very huge but this work is usually done manually by using the range checks, statistical confidence checks, etc., so very time and cost consuming. However, by using the above approach all the work for constructing an aggregate power curve can be done automatically. This is an interesting feature of the developed approach.
4.4. Wind power curve

Figure 4.29: Wind power versus wind speed for location L5

Figure 4.30: Wind power versus wind speed for location L7
Chapter 4. Wind Power Models for Security Assessment

Figure 4.31: Wind power versus wind speed for location $L_8$

Figure 4.32: Approximate power curve for location $L_3$
Figure 4.33: Approximate power curve for location L5
4.5 Conclusions

In this chapter, wind power forecast techniques as well as their use in power system studies are discussed. An effective approach to characterize all statistical features of wind data and a spatio-temporal model of wind generation for both planning and operation studies are proposed. In the model, PCA is adopted, combined with time series analysis that enables us to capture both spatial and temporal correlations for wind data from multiple wind sites. In building the model for high-dimensional and correlated data like wind data at multiple sites, the problem becomes much more complicated when the dimensions of the data set are very high. The proposed model is a tool for both decorrelating data and reducing computational burden when the problem becomes intractable. As wind power is a time-varying, intermittent, uncertain, and non-dispatchable resource while wind speed is a non-Gaussian and non-stationary stochastic process with distinct diurnal and seasonal patterns, the proposed model adopts techniques in statistics, i.e., pre-processing and transformation techniques, that remove the need for any further assumptions: this is why it can be used for any real wind data. The proposed method was tested for 10-minute wind speed data from five wind locations in the Italian region of Basilicata and hourly wind speed from nine different locations in Sicily, Italy. The representative simulation results were shown and discussed in detail.

The proposed spatio-temporal model provides a very good direction for dealing with many problems in power systems involving wind power production from multiple locations. In terms of applications, the proposed model results can be directly exploited in probabilistic power flows (see Chapter 5) to assess the power system security for the next hours or in the planning environment for studying new network reinforcement taking into account spatio-temporal correlations, which allow reducing investments compared to other probabilistic traditional planning methods. Also, such kind of results are very appropriate for solving decision-making problems, especially under uncertainty, for both operating and planning horizons. Another possible use of the proposed method is to improve wind power production forecast and its profile along the network, by means of the association of the current weather conditions provided by wind forecast to a subset of scenarios among those generated by the method.

Moreover, for the purpose of mapping wind speed to wind power for an entire wind farm or location to be used in planning and operations, an approach to obtain aggregate power curve is also developed. The proposed procedure for constructing an aggregate power curve can be done automatically, so reducing cost and time consumption.
5.1 Probabilistic models for security assessment of power systems under uncertainty

5.1.1 Introduction

In order to carry out computations for power system security assessment by a probabilistic approach, probabilistic models of all probabilistic factors in the system are required. This section presents the probabilistic models of load, wind power production, and contingencies due to branch and generating unit outages to be used in the next sections.

5.1.2 Probabilistic model of load

In deterministic security assessment, load at each bus is considered as a deterministic term with a specified value. In fact, during the operation, load varies randomly due to its stochastic nature and the system operation should adapt to this behaviour at any time [75] to maintain the system in a secure state. Probabilistic security assessment should be adopted in this case to manage the variability and uncertainty of the load and also other probabilistic factors in the system. Therefore, building a probabilistic model for load is very important. In general, models of load can be classified as time-instance model and time-period model.

In terms of time-instance model, load at bus $i$ in the system and at time $t$ is modeled by a probability distribution. Such a probability distribution can be obtained from a forecast technique. Traditionally, load forecasting results are provided in the form of point forecast (Section 4.2) in which the forecasted value and forecast error can be understood as the mean and standard deviation of the normal distribution $N(m_i^{(t)}, \sigma_i^{(t)^2})$.
Chapter 5. Probabilistic Security Assessment

(Section 2.8.2). If provided by an uncertainty forecast technique, the probability distribution is already presented in the result. The probability distribution can also be obtained by adopting the model proposed in Section 4.3 if load data are available.

For long-term probabilistic security assessment, the probability distribution of a load can be estimated based on its historical time series data. In case of security assessment for a certain time period \( T \), the probability distribution of the load can be obtained by the load duration curve for this time period \( [75] \). Figure 5.1, for example, shows a load duration curve: suppose that point \((p, t)\) of the curve represents the time duration \( t \) for which the load is greater than or equal to \( p \); the relationship is:

\[
t = F(p)
\]  

(5.1)

Dividing both sides of (5.1) by \( T \), gets:

\[
f(p) = F(p)/T
\]  

(5.2)

where \( f(p) \) represents the probability that the load is larger than or equal to \( p \), thus representing the distribution of the load.

![Figure 5.1: Load duration curve](image)

In some special cases, load connected to a bus in the system varies discretely so that it can be modeled by a discrete distribution with a number of discrete values and corresponding probabilities, e.g., as shown in Fig. 5.2.
5.1. Probabilistic models for security assessment of power systems under uncertainty

5.1.3 Probabilistic model of wind power production

In general, building a probabilistic model of wind power production is much more challenging than for load, due to the stochastic features of wind speed and wind power as discussed in detail in the previous chapter.

There are many approaches for building a probabilistic model of wind power production depending on what the model will be used for. Generally, the probabilistic model can be based on wind speed measurements or wind power measurements [76] as presented in Fig. 5.3.

The model based on measurements of wind speed can be used in almost all cases, while that based on measurements of wind power may face with some difficulties in reality. In fact, wind power measurement data are sometimes difficult to have due to,
Chapter 5. Probabilistic Security Assessment

for example, commercial reasons, or might even be not available, e.g., in case of new wind farms. In contrast, wind speed measurement data are usually available and even free to access, e.g., wind speed data from European wind Atlas. When wind speed measurement data approaches are used, as shown in Fig. 5.3, the process needs two steps, i.e., building wind speed model and building wind turbine/farm model which can be based on physical factors of wind turbine/farm or based on estimation of wind speed-wind power curve for wind turbine/farm (presented in Section 4.4).

For security assessment by using a probabilistic approach for a time-instant or a specified horizon, the probabilistic information can be provided by a wind forecast technique. Analogously to the load forecast, the form of information obtained by a wind forecast technique, i.e., point forecast or uncertainty forecast, depends on the technique used (discussed in Section 4.2). By an alternative and effective way, the model proposed in Section 4.3 is very useful to provide necessary information for probabilistic security assessment.

In terms of long-term probabilistic security assessment, the distribution can be estimated by using observed wind data. Regarding this, Weibull distribution (Section 2.8.4) is widely used to estimate the probability distribution for wind speed [77], then using wind turbine/farm model to obtain probability distribution for wind power production.

5.1.4 Probabilistic models of branch outage and generating unit outage

Contingencies in power system are probabilistic events mostly due to component outages such as branch and generating unit outages. From a computational point of view, adding contingencies makes the computation more challenging. However, with the probabilistic models presented in the following, the probabilistic methods developed in the next sections can take into account these contingencies straightforwardly.

In power system planning, a basic parameter used for modeling uncertainty is the Forced Outage Rate (FOR), estimating the probability of a component being forced out of service in the future. FOR for a two-state component is given by [4]:

\[
\text{FOR} = \frac{\lambda}{\lambda + \xi} = \frac{\text{FOH}}{\text{FOH} + \text{SH}}
\]  

(5.3)

where:

- \(\lambda\): Failure rate (failures/hour),
- \(\xi\): Repair rate (repairs/hour),
- \(\text{FOH}\): Forced outage hours,
- \(\text{SH}\): Service hours.

However, \(\text{FOR}\) is the "steady-state" probability of being in the down state and it is not capable of estimating the probability of failure for a component in a short or very short time frame. In those cases, another parameter to represent the time-dependent unavailability of a component should be adopted: The Outage Replacement Rate (ORR) is used analogously to the \(\text{FOR}\); however, while \(\text{FOR}\) is a fixed quantity associated with a component, \(\text{ORR}\) is a time-dependent quantity, depending on the lead time.
considered. \( ORR \) [4] is a function of the lead time \( t \) into the future. For a two-state component, it is given by:

\[
ORR(t) = \frac{\lambda}{\lambda + \xi} + \frac{e^{-(\lambda+\xi)t}}{\lambda + \xi} \left[ \xi U(t_0) - \lambda A(t_0) \right]
\]

\[= FOR + \frac{e^{-(\lambda+\xi)t}}{\lambda + \xi} \left[ \xi U(t_0) - \lambda A(t_0) \right] \quad (5.4)
\]

where:

- \( A(t_0) = 1 \) if the component is available at \( t_0 \); 0 otherwise,
- \( U(t_0) = 1 \) if the component is unavailable at \( t_0 \); 0 otherwise.

From (5.4), the steady-state value of \( ORR(t) \) is exactly equal to \( FOR \) of the component (Fig. 5.4):

\[
\lim_{t \to \infty} ORR(t) = FOR \quad (5.5)
\]

![Figure 5.4: ORR vs. FOR](image)

If at \( t_0 \), the component is in service, (5.4) becomes:

\[
ORR(t) = \frac{\lambda}{\lambda + \xi} (1 - e^{-(\lambda+\xi)t}) \quad (5.6)
\]

In a short or very short time frame (i.e., for \( t \to 0 \)), if a component fails, it will not be repaired during the lead time \( t \); so, \( \xi = 0 \) and:

\[
ORR(t) = 1 - e^{-\lambda t} \quad (5.7)
\]
Applying Taylor series expansion to (5.7), we obtain:

\[ ORR(t) = \lambda t + \frac{(\lambda t)^2 e^{-\lambda t}}{2!} + \frac{(\lambda t)^3 e^{-\lambda t}}{3!} + \cdots \]  

(5.8)

For a short or very short lead time \( t \), \( \lambda t \ll 1 \) so:

\[ ORR(t) \approx \lambda t \]  

(5.9)

From either the value of \( FOR \) for planning or \( ORR \) for operations, the contingencies, due to the random branch outage and generating unit outage, can be modeled:

1) For a power plant or a group with multiple units, if all of them are identical, a binomial distribution (Section 2.8.3), using either \( FOR \) or \( ORR \) as the probability of failure, will be used. Otherwise, a combination of 0-1 distributions ("0" denotes outage state; "1" denotes operating state) of units should be adopted. Figure 5.5 presents the probabilistic modeling for, for example, the power injection from a power plant including 10 identical units with the rated power of 5MW each (assume all units operating at the rated power) and \( FOR \) of 0.09. In this case, the power injection from the power plant is modeled by a binomial r.v. \( \tilde{P} \) with the p.m.f. as in Fig. 5.5

![Diagram](image_url)

**Figure 5.5**: An example of probabilistic modeling for generating unit outage

2) For a branch, the outage can be modeled as shown in Fig. 5.6 in which \( P_{ik} \) and \( Q_{ik} \) are the real and reactive power flows from bus \( i \) to bus \( k \), and \( P_{ki} \) and \( Q_{ki} \) are the real and reactive power flows from bus \( k \) to bus \( i \) before the outage. If the outage occurs, this phenomenon will be modeled by two fictitious power injections, i.e., \( P_{iout} \), \( Q_{iout} \) and \( P_{kout} \), \( Q_{kout} \), at both ends of the branch. With these fictitious injections,
the state of the branch is kept the same as before outage. This is very useful because in this case sensitivity matrices, such as the sensitivity matrix of power flows with respect to nodal power injections, are kept constant for convenience in computation, especially with analytical approach.

Two power injections are calculated as follows:

\[
\begin{bmatrix}
P_{i_{\text{out}}} \\
Q_{i_{\text{out}}} \\
P_{k_{\text{out}}} \\
Q_{k_{\text{out}}}
\end{bmatrix} = ([I] - [L])
\begin{bmatrix}
P_{ik} \\
Q_{ik} \\
P_{ki} \\
Q_{ki}
\end{bmatrix}
\]  

(5.10)

where \([I]\) is a \(4 \times 4\) identity matrix and \([L]\) is the sensitivity matrix of power flows through branch \(ik\) with respect to nodal power injections at the two ends of the branch, computed as follows:

\[
[L] = \begin{bmatrix}
\frac{\partial P_{ik}}{\partial P_i} & \frac{\partial P_{ik}}{\partial Q_i} & \frac{\partial P_{ik}}{\partial P_k} & \frac{\partial P_{ik}}{\partial Q_k} \\
\frac{\partial Q_{ik}}{\partial P_i} & \frac{\partial Q_{ik}}{\partial Q_i} & \frac{\partial Q_{ik}}{\partial P_k} & \frac{\partial Q_{ik}}{\partial Q_k} \\
\frac{\partial P_{ki}}{\partial P_i} & \frac{\partial P_{ki}}{\partial Q_i} & \frac{\partial P_{ki}}{\partial P_k} & \frac{\partial P_{ki}}{\partial Q_k} \\
\frac{\partial Q_{ki}}{\partial P_i} & \frac{\partial Q_{ki}}{\partial Q_i} & \frac{\partial Q_{ki}}{\partial P_k} & \frac{\partial Q_{ki}}{\partial Q_k}
\end{bmatrix}
\]  

(5.11)

When (5.10) is solved, the random branch outage can be modeled by two random power injections, i.e., \(P_{i_{\text{out}}}\), \(Q_{i_{\text{out}}}\), and \(P_{k_{\text{out}}}\), \(Q_{k_{\text{out}}}\), with 0-1 distributions using \(FOR\) or \(81\)
Chapter 5. Probabilistic Security Assessment

ORR, at two ends of the branch \([78]\). In Fig. 5.6, the 0-1 distribution of \(\tilde{P}_{out}\) plotted, for example, corresponds to \(P_{out} = 24.8\) MW and \(FOR = 0.01\). Multiple branch outages can also be modeled in a similar way.

5.1.5 Conclusions

The probabilistic modeling for load, wind production, branch outage, and generating unit outage are discussed in this section. Among them, building the model for wind production is one of the most challenging one due to its stochastic features and it is considered more difficult than that for load. The models presented above allow the computations in the next sections to take into account these probabilistic factors.

5.2 Probabilistic power flow

5.2.1 Introduction

The conventional deterministic power flow (PF) is the most widely used tool in power system analysis, operations, planning, and control. Deterministic PF uses the specified values of power generation and load, and the parameters of the network topology to compute system steady-state operating conditions without taking into account any sources of uncertainty affecting the power system. The deepening penetration of renewable energy sources, such as wind and photovoltaic solar, has introduced additional uncertainty into power system operations and control. This added uncertainty, along with the conventional sources of uncertainty such as the loads and the availability of resources and transmission assets, makes clear the limitations of the conventional deterministic PF in power system analysis and security assessment applications. In order to deal with the above problems, a probabilistic approach needs to be used, where sources of uncertainty are explicitly represented by r.v.s. This approach is referred to as Probabilistic Power Flow (PPF).

Unlike deterministic PF, PPF is capable of managing the wide spectrum of all possible values of the input and state variables so as to provide a complete spectrum of all possible values of outputs of interest such as nodal voltages, line power flows, etc., in terms of p.d.f.s and/or c.d.f.s. Such kind of result can be used for power system analysis and security assessment by probabilistic approaches [35].

In this section, first, various PPF methods which have been introduced so far are summarized and discussed, then different cumulant-based PPF methods are presented in detail. To illustrate the applications of the developed methods, extensive tests on 14-bus and 300-bus IEEE test systems are carried out. The results obtained from different approaches are compared and discussed.

5.2.2 Overview of probabilistic power flow methodologies

The early contribution in the PPF area was the scheme published in 1974 [80]. Since then, it has been developed and applied to various areas in power systems where the uncertainty needs to be managed. PPF procedures can generally be divided into two categories: numerical methods and analytical methods.

Monte Carlo Simulation (MCS) is a typical numerical method. It is a systematic methodology for the emulation of events under uncertainty [81]. The MCS for power
5.2. Probabilistic power flow

flow studies under uncertainty uses multiple deterministic PF solutions for the sampled values of the realizations of the r.v.s that are used to represent the various sources of uncertainty.

The basic idea of the analytical approach is to apply a defined algorithm, i.e., point estimate methods \[82,83\], cumulant techniques \[22,24,78,84,85\], or convolution techniques \[86,87\] with p.d.f.s and/or c.d.f.s of r.v.s of inputs so that p.d.f.s and c.d.f.s of r.v.s of system states and line flows can be obtained. Point estimate methods are based on approximations: input variables are decomposed into a series of values and corresponding weights, and then the moments of the output variables of interest are computed as a function of the inputs. The cumulant methods adopt the properties of moments and cumulants, based on the probability distributions of input r.v.s and linearized power flow equations, while the convolution methods convolve all r.v.s.

Both the numerical and analytical methods have their own advantages and disadvantages. In MCS, the non-linear PF equations can be directly used. The accuracy of the results is one of the advantages of MCS and mostly depends on the number of samples \[88,90\]. However, MCS is usually very computationally intensive. On the contrary, the analytical approaches are much faster in computation speed. Nevertheless, analytical approaches usually involve complex mathematical computation and their result accuracy are less than that of MCS result mostly due to using linearized PF equations and/or approximation techniques. Generally, among analytical approaches, cumulant-based approach is the fastest, followed by point estimate method and convolution method. By using cumulants, it is possible to obtain the p.d.f. and c.d.f. of a linear combination of several r.v.s by a quite simple arithmetic \[84\] and less computationally intensive process while maintaining an appropriate level of accuracy, so suitable for large power systems. That is why cumulant method is chosen to develop and apply to probabilistic security assessment in this chapter.

For the literature on PPF, reference \[91\] provides a very good summary and bibliography on PPF published before 1988. Also, PPF techniques are summarized in \[92\] with an extensive bibliography attached. PPF is recently developed and applied to many areas including security assessment, stability analysis, power system planning and operation, power system control, economic analysis and so on, especially in the systems with significant integration of renewable resources such as wind and photovoltaic solar, etc., which can be found in \[93-103\].

5.2.3 Formulation of cumulant-based probabilistic power flow methods

The deterministic PF requires specific values for loads, generations and network conditions. Therefore, the uncertainties of loads, power outputs of generators, etc., and component parameter cannot be taken into account by deterministic PF. However, the probabilistic approach makes it possible to solve these problems. Each uncertainty is modeled by a probabilistic distribution represented by a r.v. Depending on its stochastic behaviour, a suitable probability distribution can be used for modeling: for example, normal distribution for load, binomial distribution for a contingency, etc. PPF calculations can provide p.d.f. of line power flows, nodal voltage probability distributions, and they can estimate the probability of line overloading, over-voltage, and other useful security-related information. In this section, different cumulant-based PPF calculation methods are presented and compared.
Chapter 5. Probabilistic Security Assessment

1. Series expansions and sensitivities (M1)

Method M1 is based on series expansions (see Section 2.9.1, Chapter 2) and sensitivity matrices [75]. Continuous and discrete distributions are dealt with together during the computations.

The form of the PF equations generally is:

- **Real and reactive injected powers:**
  \[ P_i = V_i \sum_{k=0}^{n-1} V_k (G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik}) \]  
  \[ Q_i = V_i \sum_{k=0}^{n-1} V_k (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik}) \]  

- **Real and reactive line power flows:**
  \[ P_{ik} = V_i V_k (G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik}) - t_{ik} G_{ik} V_i^2 \]  
  \[ Q_{ik} = V_i V_k (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik}) + t_{ik} B_{ik} V_i^2 - B'_{ik} V_i^2 \]

where:

- \( P_i \) and \( Q_i \) are real and reactive power injections at bus \( i \),
- \( P_{ik} \) and \( Q_{ik} \) are real and reactive power power flows on the branch \( ik \),
- \( V_i \) and \( V_k \) are voltage magnitudes at buses \( i \) and \( k \),
- \( \theta_{ik} \) is voltage angle difference between buses \( i \) and \( k \),
- \( G_{ik} \) and \( B_{ik} \) are the real and imaginary parts of the admittance matrix of branch \( ik \),
- \( B'_{ik} \) is the haft of susceptance of branch \( ik \),
- \( t_{ik} \) is the transformation ratio of branch \( ik \),
- \( n \) is the total number of buses in the system.

The above equations can be expressed by matrix form as

\[ w = g(x) \]  
\[ z = h(x) \]

where:

- \( w \) is the vector of nodal injected powers,
- \( x \) is the vector of state variables,
- \( z \) is the vector of line power flows,
• \( g(x) \) are the power flow equations,
• \( h(x) \) are the functions to compute line power flows.

The solution \( \bar{x} \) and \( \bar{z} \) of the above equations is obtained by solving a deterministic PF.

Using Taylor series expansion to linearize the above equations around the solution point gives:

\[
\Delta x = [S]\|_{\bar{x}} \Delta w
\]
\[
\Delta z = [T]\|_{\bar{x}} \Delta w
\]

where \([S]\|_{\bar{x}}\) is the inverse of the well-known Jacobian matrix and \([T]\|_{\bar{x}}\) is the sensitivity matrix of power flows with respect to nodal power injections, both computed at the solution point \( \bar{x} \).

To solve the PPF, each element of \( w \) is considered as the realization of the r.v. associated with each nodal power injection \( \tilde{W}_j \), \( j \in \{1, \ldots, n-1+\ell\} \) (assume bus 0 is the slack bus in single slack bus model; \( \ell \) is the total number of load buses; elements from 1 to \( n-1 \) are relevant to real power equations; the others to reactive power equations). From relationships in (5.18) and (5.19), elements of \( x \) and \( z \) are, therefore, the realizations of corresponding r.v.s \( \tilde{X}_\chi \), \( \chi \in \{1, 2, \ldots, n-1+\ell\} \) (similarly, elements from 1 to \( n-1 \) are related to angle terms, while the rest are voltage magnitude terms) and \( \tilde{Z}_\zeta \), \( \zeta \in \{1, 2, \ldots, 2b\} \) \((b \) is the total number of branches), respectively. Consequently, we have the relationships of r.v.s as follows:

\[
\begin{bmatrix}
\tilde{X}_1 \\
\tilde{X}_2 \\
\vdots \\
\tilde{X}_{n-1+\ell}
\end{bmatrix} = [S]\|_{\bar{x}} 
\begin{bmatrix}
\tilde{W}_1 \\
\tilde{W}_2 \\
\vdots \\
\tilde{W}_{n-1+\ell}
\end{bmatrix}
\]
\[
\begin{bmatrix}
\tilde{Z}_1 \\
\tilde{Z}_2 \\
\vdots \\
\tilde{Z}_{2b}
\end{bmatrix} = [T]\|_{\bar{x}} 
\begin{bmatrix}
\tilde{W}_1 \\
\tilde{W}_2 \\
\vdots \\
\tilde{W}_{n-1+\ell}
\end{bmatrix}
\]

Finally, based on (5.20), (5.21) and (2.23) (see Section 2.7, Chapter 2), the calculation procedure for this method can be described as follows [104]:

(1) Run the deterministic PF calculation to obtain the expected values \( \bar{x} \) and \( \bar{z} \) and the sensitivity matrices \([S]\|_{\bar{x}}\) and \([T]\|_{\bar{x}}\);
(2) Calculate the self (see Section 2.5) and joint (see Section 2.6) cumulants of nodal power injections;
(3) Compute cumulants of state variables and line power flows using (5.20), (5.21), and (2.23) (if r.v.s are dependent; otherwise, (2.22) used);
Chapter 5. Probabilistic Security Assessment

(4) Obtain the p.d.f.s and/or c.d.f.s of the outputs of interest using approximation methods based on series expansions presented in Section 2.9.1. The expected values \( \bar{x} \) and \( \bar{z} \) and the sensitivity matrices \( [S]_{\bar{x}} \) and \( [T]_{\bar{x}} \) can be obtained from either the decoupled deterministic PF (in the following, M1a) or the non-decoupled deterministic PF (in the following, M1b).

2. Von Mises function and sensitivities (M2)

Method M2 also uses sensitivity matrices; however, continuous and discrete distributions of input parameters are separately treated and the Von Mises function, presented in Section 2.9.2, is employed instead of series expansions.

The output-input relationships are separated as follows

- For continuous part:
  \[
  \begin{bmatrix}
  \tilde{X}_{c,1} \\
  \tilde{X}_{c,2} \\
  \vdots \\
  \tilde{X}_{c,n-1+\ell}
  \end{bmatrix}
  = [S]_{\bar{x}}
  \begin{bmatrix}
  \tilde{W}_{c,1} \\
  \tilde{W}_{c,2} \\
  \vdots \\
  \tilde{W}_{c,n-1+\ell}
  \end{bmatrix}
  \]  
  \( (5.22) \)

- For discrete part:
  \[
  \begin{bmatrix}
  \tilde{X}_{d,1} \\
  \tilde{X}_{d,2} \\
  \vdots \\
  \tilde{X}_{d,n-1+\ell}
  \end{bmatrix}
  = [S]_{\bar{x}}
  \begin{bmatrix}
  \tilde{W}_{d,1} \\
  \tilde{W}_{d,2} \\
  \vdots \\
  \tilde{W}_{d,n-1+\ell}
  \end{bmatrix}
  \]  
  \( (5.24) \)

- For continuous part:
  \[
  \begin{bmatrix}
  \tilde{Z}_{c,1} \\
  \tilde{Z}_{c,2} \\
  \vdots \\
  \tilde{Z}_{c,2b}
  \end{bmatrix}
  = [T]_{\bar{x}}
  \begin{bmatrix}
  \tilde{W}_{c,1} \\
  \tilde{W}_{c,2} \\
  \vdots \\
  \tilde{W}_{c,n-1+\ell}
  \end{bmatrix}
  \]  
  \( (5.23) \)

- For discrete part:
  \[
  \begin{bmatrix}
  \tilde{Z}_{d,1} \\
  \tilde{Z}_{d,2} \\
  \vdots \\
  \tilde{Z}_{d,2b}
  \end{bmatrix}
  = [T]_{\bar{x}}
  \begin{bmatrix}
  \tilde{W}_{d,1} \\
  \tilde{W}_{d,2} \\
  \vdots \\
  \tilde{W}_{d,n-1+\ell}
  \end{bmatrix}
  \]  
  \( (5.25) \)

The calculation process for this method is [104]:

(1) Run the deterministic PF calculation to obtain the expected values \( \bar{x} \) and \( \bar{z} \) and the sensitivity matrices \( [S]_{\bar{x}} \) and \( [T]_{\bar{x}} \);

(2) Calculate the self and joint cumulants of nodal power injections for both continuous and discrete parts;
5.2. Probabilistic power flow

(3) The continuous part is carried out like in M1 using (5.22) and (5.23), and the corresponding cumulants of state variables and line power flows are computed; from these values, the p.d.f. and c.d.f. needed in (2.51) and (2.53) (Section 2.9.2, Chapter 2) are determined;

(4) Compute cumulants of state variables and line power flows for discrete part using (5.24) and (5.25); next, convert these cumulants to moments according to (2.15), and then use the process described in Section 2.9.2 to calculate abscissas $x_i$ and corresponding probabilities $p_i$;

(5) Obtain the p.d.f.s and/or c.d.f.s of the outputs of interest using (2.50) and (2.52).

Similarly to M1, the expected values $\bar{x}$ and $\bar{z}$ and the sensitivity matrices $[S]|_{\bar{x}}$ and $[T]|_{\bar{x}}$ can be obtained from either decoupled deterministic PF (denoted by M2a) or non-decoupled one (denoted by M2b). However, the above equations can be used in different ways.

In [78], each discrete distribution with $\xi$ impulses is split into $\xi$ single-impulse distributions and equations (5.24) and (5.25) are applied to all possible combinations of the latter single-impulse distributions. However, the method is based on the linearization of power flow equations: its accuracy should be assessed, especially in the case of high uncertainty level of inputs, as some impulses of the discrete part of nodal power injections could cause great changes, thus affecting the accuracy of the linearization. Therefore, for all the above combinations, the nodal voltage angle variations are checked and if they are too large, a new deterministic PF is carried out and equations (5.24) and (5.25) are applied with the new sensitivity matrices. In the presence of many discrete distributions, the method can become very time consuming, although it obtains accurate results, especially in the presence of contingencies.

In the present work, according to the technique adopted in [24] in another framework, equations (5.24) and (5.25) are applied only once, where the cumulants of all discrete distributions are taken into account in one shot. The results show that generally the obtained cumulants of order higher than one are accurate enough but the mean values are slightly different compared to the MCS. The idea of shifting technique in [86] can be adopted to overcome this difficulty: the mean values are shifted to coincide with the expected values obtained by deterministic PF calculation. The shapes of p.d.f. and c.d.f. curves remain the same.

3. Von Mises function and direct cumulant method (M3)

Method M3 deals with both continuous and discrete distributions too; as in M2, they are accounted for separately by exploiting the Von Mises function and the cumulant properties; the difference is in the technique adopted for the linearization of the PF equations [87].

Taken as an example to illustrate the technique for linearization, two $r.v.s$ $\tilde{X}_1$ and $\tilde{X}_2$ are considered. Assume that the two $r.v.s$ have corresponding expected values $\bar{X}_1$ and $\bar{X}_2$ and their values may be randomly changed around their expected values by $\Delta X_1$ and $\Delta X_2$, respectively. The $r.v.s$ can therefore be expressed as:

$$\tilde{X}_1 = \bar{X}_1 + \Delta X_1$$

(5.26)
Also assume that the product (also a r.v.) \( \tilde{Y} = \tilde{X}_1 \tilde{X}_2 \) needs to be approximated to obtain the more simple representation which is suitable for analytical approach. We have:

\[
\tilde{Y} = \tilde{X}_1 \tilde{X}_2 = (\bar{x}_1 + \tilde{X}_1^\Delta)(\bar{x}_2 + \tilde{X}_2^\Delta) = \bar{x}_1 \bar{x}_2 + \bar{x}_1 \tilde{X}_2^\Delta + \bar{x}_2 \tilde{X}_1^\Delta + \tilde{X}_1^\Delta \tilde{X}_2^\Delta
\]

(5.28)

If the variations \( \tilde{X}_1^\Delta \) and \( \tilde{X}_2^\Delta \) are small, the quantity \( \tilde{X}_1^\Delta \tilde{X}_2^\Delta \) will be very small, and can be neglected. Also, take \( \tilde{X}_1^\Delta \) and \( \tilde{X}_2^\Delta \) from (5.26) and (5.27) then substitute into (5.28), finally, \( \tilde{Y} \) can be approximated as:

\[
\tilde{Y} \approx \tilde{x}_2 \tilde{x}_1 + \bar{x}_1 \bar{x}_2 - \bar{x}_1 \bar{x}_2
\]

(5.29)

In PF equations, such product terms exist and make the work of linearization more difficult. The above technique can be adopted for the following terms in (5.12) to (5.15) because the variations of voltage magnitudes around the operating point in a power system are generally small:

\[
\tilde{V}_i \tilde{V}_k = \bar{v}_i \bar{V}_k + \tilde{v}_i \tilde{V}_k - \bar{v}_i \bar{V}_k
\]

(5.30)

\[
\tilde{V}_i^2 = 2\bar{v}_i \tilde{V}_i - \bar{v}_i^2
\]

(5.31)

where \( \tilde{V}_i \) and \( \tilde{V}_k \) are r.v.s of voltages at buses \( i \) and \( k \), \( \tilde{v}_i \) and \( \tilde{v}_k \) are the corresponding expected values (values at the operating point).

Also, due to the small changes around the operating point of the angles in power system, the Maclaurin’s series is applied to the following terms:

\[
\sin \tilde{\Theta}_{ik} \approx \tilde{\Theta}_{ik} - \frac{\tilde{\Theta}_{ik}^3}{6}
\]

(5.32)

\[
\cos \tilde{\Theta}_{ik} \approx 1 - \frac{\tilde{\Theta}_{ik}^2}{2}
\]

(5.33)

Again, (5.29) is applied to the high-order terms at the right side of (5.32) and (5.33):

\[
\sin \tilde{\Theta}_{ik} \approx \frac{\tilde{\Theta}_{ik}^3}{3} + (1 - \frac{\tilde{\Theta}_{ik}^2}{2})\tilde{\Theta}_{ik}
\]

(5.34)

\[
\cos \tilde{\Theta}_{ik} \approx (1 + \frac{\tilde{\Theta}_{ik}^2}{2}) - \tilde{\Theta}_{ik} \tilde{\Theta}_{ik}
\]

(5.35)

where \( \tilde{\Theta}_{ik} \) and \( \tilde{\Theta}_{ik} \) denote the r.v and corresponding expected value of angle difference between buses \( i \) and \( k \).

Applying all the above linearizations to equations from (5.12) to (5.15), the following equations are obtained and used to directly calculate cumulants for the continuous
and discrete parts of the outputs [87] [104], instead of the sensitivities used by M1 and M2:

\[
\tilde{\Theta}_i = \sum_{j=1}^{n-1} \xi_{ij} \tilde{P}_j - \sum_{j=1}^{n-1} \xi_{ij} r_j \quad (5.36)
\]

\[
\tilde{V}_{i(l)} = \sum_{j=1}^\ell \eta_{ij} \tilde{Q}_{j(l)} + \sum_{j=1}^{\ell} \eta_{ij} h_j \quad (5.37)
\]

\[
\tilde{P}_{ik} = g_{ik} \tilde{\Theta}_i - g_{ik} \tilde{\Theta}_k + h_{ik} \quad (5.38)
\]

\[
\tilde{Q}_{ik} = \alpha_{ik} \tilde{V}_i + \beta_{ik} \tilde{V}_k + \gamma_{ik} \quad (5.39)
\]

where:

- \( \tilde{P}_j \) is the r.v. of real power injection at bus \( j \),
- \( \tilde{Q}_{j(l)} \) is the r.v. of reactive power injection at load bus \( j \) (\( \ell \) is the total number of load buses as denoted),
- \( \tilde{\Theta}_i \) is the r.v. of angle at bus \( i \),
- \( \tilde{V}_{j(l)} \) is the r.v. of voltage magnitude at load bus \( j \),
- \( \tilde{P}_{ik} \) and \( \tilde{Q}_{ik} \) are the r.v.s of real and reactive power power flows on branch \( ik \).

The others factors in equations from (5.36) to (5.39) are constant terms that can be calculated from network data [87].

The method was proposed in [24] with the decoupled approach (M3a) using equations from (5.36) to (5.39).

In the present research, the above method is further developed by a non-decoupled approach (M3b) to gain a greater accuracy, similarly to [105] obtaining equations:

\[
\begin{bmatrix}
\tilde{\Theta}_1 \\
\vdots \\
\tilde{\Theta}_{n-1} \\
- \tilde{V}_{1(l)} \\
\vdots \\
- \tilde{V}_{\ell(l)}
\end{bmatrix} = [C] \begin{bmatrix}
\tilde{P}_1 \\
\vdots \\
\tilde{P}_{n-1} \\
- \tilde{Q}_{1(l)} \\
\vdots \\
- \tilde{Q}_{\ell(l)}
\end{bmatrix} + [D] \quad (5.40)
\]

\[
\begin{bmatrix}
\tilde{P}_{ik,1} \\
\vdots \\
\tilde{P}_{ik,h}
\end{bmatrix} = [C_P] \begin{bmatrix}
\tilde{\Theta}_0 \\
\vdots \\
\tilde{\Theta}_{n-1} \\
- \tilde{V}_1 \\
\vdots \\
- \tilde{V}_{n-1}
\end{bmatrix} + [D_P] \quad (5.41)
\]
where $[C], [D], [C_{P}], [D_{P}], [C_{Q}], \text{ and } [D_{Q}]$ are constant matrices, computed from network data \cite{87,105}.

The differences between M3 and the above methods is that the first order cumulant (mean) of the desired outputs can be calculated directly using explicit formulas while $\bar{x}$ and $\bar{z}$ are considered as the means of outputs in M1 and M2.

5.2.4 Tests and numerical results

In order to illustrate and compare the above methods, two case studies relevant to IEEE 14-bus and IEEE 300-bus test systems are presented and discussed in the following. The aim of these case studies is to check and compare the performances of the cumulant-based PPF methods developed. In these tests, input r.v.s are assumed to be independent. However, this assumption will be discussed carefully in the next section.

**Case 1: IEEE 14-bus test system**

The single line diagram of the IEEE 14-bus test system is shown in Fig. [5.7], while the branch, bus, and generator data are presented in Appendix [A]. Loads at buses 2-6 and 10-14 are modeled by normal distributions characterized by mean and standard deviation, while load at bus 9 is modeled by a discrete distribution characterized by 5 impulses and by the relevant probabilities. The generators at bus 1 and 2 are modeled by binomial distributions with relevant FORs. All probabilistic information for IEEE 14-bus test system is also presented in Appendix [A].

In order to assess the accuracy and efficiency of all the considered PPF methods presented in Section 5.2.3, a MCS with 50,000 samples was carried out and taken as reference.

In M1, approximation method based on Gram-Charlier series expansion is used, while the shifting technique is applied to M2. The results point out that the shifting technique not only can provide satisfactory outcomes but it is also simpler.

In terms of the mean values, all the above methods give good results compared to MCS. However, looking at the standard deviations (see figures from 5.8 to 5.11), the results indicate clearly that the active-reactive coupling problem is to be considered and affects much more voltage magnitudes and reactive power flows than angles and real power flows. The standard deviations of all non-decoupled methods (M1b, M2b, and M3b) are near to MCS results in all these figures; on the contrary, in Fig. 5.9 and 5.11, the large difference among the methods are easy to be seen, owing to decoupling. In particular, in Fig. 5.10 and 5.11, the standard deviations of both $\tilde{P}_{1-2}$ and $\tilde{Q}_{1-2}$ are
5.2. Probabilistic power flow

Figure 5.7: Single line diagram of the IEEE 14-bus test system [2]

much higher than in other cases maybe because line 1-2 is connected to the slack bus (bus 1) and it is the most important line in the system with the heaviest loading and in the single slack bus model the slack bus injection is responsible for reflecting the combination of uncertainty of all injections, resulting in the large value of standard deviation. This will be further discussed in Section 5.3.

However, the mean and standard deviation are not enough to describe non-Gaussian distributions. The c.d.f.s and/or p.d.f.s, for example, of a selected voltage at bus 12 ($\tilde{V}_{12}$), real and reactive power flows through branch 3-4 ($\tilde{P}_{3-4}$ and $\tilde{Q}_{3-4}$, respectively) are presented in figures from 5.12 to 5.15 respectively. It is interesting that the curves obtained by non-decoupled model are very close to the MCS curve. Moreover, the results of M2b and M3b, which treat separately continuous and discrete distributions and employ Von Mises functions, are more accurate than M1b. In this test, the level of uncertainties of the inputs is not very high because, so far, branch outages are not considered and for example possible large variability of wind power production is not
Figure 5.8: Standard deviation of selected nodal voltage angles

Figure 5.9: Standard deviation of nodal voltage magnitudes

simulated. However, their differences will increase as the uncertainty increases.
5.2. Probabilistic power flow

Figure 5.10: Standard deviation of selected real power flows

Figure 5.11: Standard deviation of selected reactive power flows
Chapter 5. Probabilistic Security Assessment

Figure 5.12: p.d.f.s of $\tilde{V}_{12}$

Figure 5.13: c.d.f.s of $\tilde{Q}_{3-4}$
5.2. Probabilistic power flow

Figure 5.14: p.d.f.s of $\tilde{Q}_{3-4}$

Figure 5.15: c.d.f.s of $\tilde{P}_{3-4}$
Chapter 5. Probabilistic Security Assessment

Table 5.1: ARMS for 3 selected output r.v.s

<table>
<thead>
<tr>
<th>ARMS(%)</th>
<th>M1a</th>
<th>M1b</th>
<th>M2a</th>
<th>M2b</th>
<th>M3a</th>
<th>M3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{V}_{12} )</td>
<td>0.4176</td>
<td>0.0201</td>
<td>0.3685</td>
<td>0.0109</td>
<td>0.3579</td>
<td>0.0106</td>
</tr>
<tr>
<td>( \tilde{P}_{3-4} )</td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td>( \tilde{Q}_{3-4} )</td>
<td>2.3118</td>
<td>0.0056</td>
<td>2.2815</td>
<td>0.0023</td>
<td>2.2423</td>
<td>0.0021</td>
</tr>
</tbody>
</table>

Table 5.2: ARMS of \( \tilde{P}_{3-4} \) with random outage line 2-4

<table>
<thead>
<tr>
<th>ARMS(%)</th>
<th>M1a</th>
<th>M1b</th>
<th>M2a</th>
<th>M2b</th>
<th>M3a</th>
<th>M3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{P}_{3-4} )</td>
<td>0.0609</td>
<td>0.0568</td>
<td>0.0156</td>
<td>0.0079</td>
<td>0.0153</td>
<td>0.0078</td>
</tr>
</tbody>
</table>

In order to perform an overall evaluation of the accuracy of all considered methods, the Average Root Mean Square (ARMS) error is also computed using the MCS results as in [22]. ARMS is defined as:

\[
ARMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (MCS_i - C MP PF_i)^2}
\]

(5.43)

where \( MCS_i \) and \( C MP PF_i \) are the \( i^{th} \) value on c.d.f. curves obtained by MCS and cumulant-based PPF method, respectively, and \( N \) is the number of samples considered in the range of the c.d.f. (in this test, an interval of \( 10^{-4} \) p.u. was adopted).

The ARMS of \( \tilde{V}_{12} \), \( \tilde{P}_{3-4} \), and \( \tilde{Q}_{3-4} \) are presented in Table 5.1. These numerical results indicate again both the different effects of the coupling and the accuracy of the considered cumulant methods: non-decoupled methods (M1b, M2b, M3b) appear to be more accurate.

Some tests were also carried out to assess the accuracy of the considered PPF methods in the presence of contingencies. A branch outage can be modeled as presented in Section 5.1.4. Both c.d.f. curves in Fig. 5.16 and ARMS values in Table 5.2 of \( \tilde{P}_{3-4} \) with random outage line 2-4 (\( FOR = 0.03 \)) show that M2 and M3, which adopt Von Mises function, can provide better results than M1.
Case 2: IEEE 300-bus test system

In order to evaluate the features of the proposed methods in terms of both accuracy and computation times, a larger power system, namely the IEEE 300-bus test system, has been considered.

The network diagram, branch, bus, and generator data of the IEEE 300-bus test system can be obtained from [2]. For probabilistic data, all loads, except loads at bus 9, 162, and 246, are assumed to be normally distributed with standard deviations equal to 8% of the nominal power. Generating unit forced outages are not taken into account. Discrete loads at bus 9, 162, and 246 with 5 impulses each are shown in Appendix B.

All the non-decoupled methods provide good results compared to MCS with 10,000 samples. The p.d.f. curves of real and reactive power flows, for example, through branch 126-132 ($\tilde{P}_{126-132}$ and $\tilde{Q}_{126-132}$, respectively) are presented in Fig. 5.17 and 5.18 respectively. The obtained results indicate again that coupling problem impacts particularly on voltages and reactive power flows (Fig. 5.18).

Table 5.3 compares the computation times referred to the different methods with MCS. Method M2a and M2b times are determined using the shifting technique described Section 5.2.3. All the methods based on cumulants can quickly obtain the desired results even for a large network, while the MCS is significantly time consuming. M1 is the fastest while M2 and M3 need a slightly higher computation time. Table 5.4 shows the computation times of M2a (called M2as in the following) and M2b (called M2bs in the following) with different thresholds of corresponding voltage angles when each impulse of the discrete part is handled one by one. The relevant times are higher than M2a and M2b, but the accuracy is improved as depicted by Fig. 5.19 in which the
c.d.f. curves of $\tilde{Q}_{126-132}$ of the above methods are presented. The curves obtained by M2as and M2bs are relevant to the threshold of 0.1 radian. Figure 5.19 shows that the
5.2. Probabilistic power flow

Figure 5.19: c.d.f.s of $\tilde{Q}_{126-132}$

Table 5.3: Computation time comparison for IEEE 300-bus test system

<table>
<thead>
<tr>
<th>Method</th>
<th>M1a</th>
<th>M1b</th>
<th>M2a</th>
<th>M2b</th>
<th>M3a</th>
<th>M3b</th>
<th>MCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation time (second)</td>
<td>1.61</td>
<td>1.98</td>
<td>2.21</td>
<td>2.47</td>
<td>2.25</td>
<td>2.52</td>
<td>596</td>
</tr>
</tbody>
</table>

Table 5.4: Computation time of method M2 with different thresholds

<table>
<thead>
<tr>
<th>Threshold (radian)</th>
<th>Computation time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M2as</td>
</tr>
<tr>
<td>0.05</td>
<td>8.68</td>
</tr>
<tr>
<td>0.1</td>
<td>8.43</td>
</tr>
<tr>
<td>0.2</td>
<td>6.23</td>
</tr>
<tr>
<td>0.3</td>
<td>2.88</td>
</tr>
<tr>
<td>0.4</td>
<td>2.43</td>
</tr>
</tbody>
</table>

curve relevant to M2bs is very close to MCS.

To evaluate the overall accuracy, the ARMS errors of all desired outputs are computed. Grouped by type of variable, the average value ($m$) and standard deviation ($\sigma$) of their means errors are presented in Table 5.5. It is clear that the errors are very small in terms of angles and real power flows with all considered methods. Nevertheless, for voltages and reactive power flows, the errors are still small with the non-decoupled methods but much larger with the decoupled approaches.
Table 5.5: ARMS (%) of IEEE 300-bus test system (large errors in bold)

<table>
<thead>
<tr>
<th>Method</th>
<th>M1a</th>
<th>M1b</th>
<th>M2a</th>
<th>M2b</th>
<th>M3a</th>
<th>M3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>0.0174</td>
<td>0.0170</td>
<td>0.0173</td>
<td>0.0169</td>
<td>0.0173</td>
<td>0.0169</td>
</tr>
<tr>
<td></td>
<td>0.0043</td>
<td>0.0041</td>
<td>0.0043</td>
<td>0.0041</td>
<td>0.0043</td>
<td>0.0041</td>
</tr>
<tr>
<td>Voltage</td>
<td>1.1691</td>
<td>0.2303</td>
<td><strong>1.1596</strong></td>
<td>0.2243</td>
<td><strong>1.1594</strong></td>
<td>0.2242</td>
</tr>
<tr>
<td></td>
<td>1.1863</td>
<td>0.2815</td>
<td><strong>1.1861</strong></td>
<td>0.2742</td>
<td><strong>1.1860</strong></td>
<td>0.2742</td>
</tr>
<tr>
<td></td>
<td>0.0059</td>
<td>0.0045</td>
<td>0.0053</td>
<td>0.0039</td>
<td>0.0052</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>0.0111</td>
<td>0.0095</td>
<td>0.0102</td>
<td>0.0090</td>
<td>0.0101</td>
<td>0.0088</td>
</tr>
<tr>
<td>Real power</td>
<td>2.3263</td>
<td>0.1468</td>
<td><strong>2.3102</strong></td>
<td>0.1343</td>
<td><strong>2.3101</strong></td>
<td>0.1342</td>
</tr>
<tr>
<td></td>
<td>5.0681</td>
<td>0.2359</td>
<td><strong>5.0727</strong></td>
<td>0.2107</td>
<td><strong>5.0727</strong></td>
<td>0.2105</td>
</tr>
</tbody>
</table>

5.2.5 Final comments on the application of the cumulant-based PPF methods

From the above analysis and comparison, each method has its own advantages and disadvantages, which should be considered carefully before adopting them for real applications.

M1 is the fastest one; however, its main disadvantage is the use of series expansions as discussed in Section 2.9.1. For Gaussian distribution, series expansions provide very good results. On the contrary, they have a limited domain of applicability for non-Gaussian distributions, because of their rather poor convergence properties. Consequently, the method has some restrictions in application when taking into account different distributions and high uncertainties such as branch outages, large variability of wind power production, etc.

Both M2 and M3 can overcome the drawbacks of M1 by treating separately continuous and discrete distributions and adopting the Von Mises function. Nevertheless, these two methods are slightly more time consuming than M1. M2 with shifting technique and M3 are suggested in the presence of several randomly distributed inputs such as discrete loads and unplanned outages of generating units and without taking into account very high level of uncertainties. M2 method, which treats individually each disturbance due to discrete parts, as in [78], should be applied to deal with branch outages problems and the thresholds are used in these cases to balance between accuracy and efficiency.

For the comparison between decoupled and non-decoupled methods, the decoupled methods provide accurate results for angles and real power flows, but are less accurate in case voltage magnitudes and reactive power. Of course, the increased complexity results in increased computation time. This should also be considered depending on the application. Indications for the application of methods are briefly presented in Table 5.6.
5.2. Probabilistic power flow

Table 5.6: Indications for the application of methods

<table>
<thead>
<tr>
<th>Input probability distributions</th>
<th>M1</th>
<th>M2s</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian distributions</td>
<td>⋆</td>
<td>⋆</td>
<td>⋆</td>
</tr>
<tr>
<td>Gaussian and non-Gaussian distributions without very high level of uncertainty (e.g., discrete loads, generating unit outages)</td>
<td>⌂ ∧</td>
<td>⋆</td>
<td>⋆</td>
</tr>
<tr>
<td>Gaussian and non-Gaussian distributions with very high level of uncertainty (e.g., branch outages)</td>
<td>⋆</td>
<td>⋆</td>
<td>*</td>
</tr>
</tbody>
</table>

Legend: ⋆ very good, ⋆ not applied, * good, ∧ not guarantee the convergence, ⌂ fair

5.2.6 Conclusions

This section described the cumulant-based methods in PPF studies: an in-depth analysis and comparison are presented.

A first conclusion is that the desired results can be quickly obtained by cumulant-based methods, even for a large power system.

Secondly, the effect of active-reactive coupling in PPF has been evaluated: the decoupled techniques cause small effects on angles and real power flows but can affect significantly nodal voltage magnitudes and reactive power flows accuracy.

Thirdly, in order to improve accuracy, an enhanced method (M3b) has been developed in this section by using a non-decoupled model. In addition, another method (M2) was also improved by applying the shifting technique to obtain both satisfactorily accurate results and simpler computational procedures.

Fourthly, cumulant-based PPF methods can explicitly take into account the correlation between input nodal power injections (e.g., among load and among wind) and the contingencies due to branch and generating unit outages. However, these factors will be considered carefully in the next section.

Finally, the PPF variants illustrated in this research are found to be suitable for different applications depending on the type of distributions considered. In particular, the results coming from the application of the PPF methods to a small and a large power system allowed to compare the different methods under different aspects (computational burden, accuracy, etc.) and to assess the most suitable one depending on the application requirements.
5.3 Distributed slack bus probabilistic power flow

5.3.1 Background and motivation

Traditionally, the slack bus is known in PF computation as the bus which is responsible for balancing power mismatch in power system and losses. The conventional deterministic PF is based on the concept of a single slack bus due to the lack of prior knowledge of the losses in power system. In addition, this works well as far as the injections are known with good confidence. The slack bus can be regarded as a mathematical model to take into account a priori uncertainty on losses; in the actual operation, however, the loading of generators is also determined by the frequency control model. For example, in the very short-term, any imbalance is compensated by the droops of primary frequency regulators; in a longer term, the same occurs for the participation factors of secondary frequency control (or AGC) even or tertiary control. In any case, any power imbalance can be allocated to some controlling generators, according to some known participation factors. The deterministic PF model can be enhanced to take this into account, by the Distributed Slack Bus (DSB) [106]. With the DSB, any power mismatch can be assigned to a set of generating units participating in a real power allocation process, thus reflecting the actual power system operation.

In modern power systems, the uncertainty of bus power injections is significant, for example, due to the stochastic nature of loads and wind power resources; in this framework, the slack bus injection reflects the combination of uncertainty of all injections too. The power output of all generators under frequency control will be affected by that uncertainty, as well. Therefore, for such generators, it becomes necessary to assess both reserve margins and ramping constraint under uncertainty by using a probabilistic tool. Such a probabilistic tool is proposed in this section to include the steady-state behaviour of the frequency regulation of conventional generation and provide more valuable information for power system analysis and security assessment (e.g., probability of over-/under-regulation limits, probability of not meeting ramping requirements, probability of line overloading, probability of over-/under-voltage, etc.).

5.3.2 Distributed slack bus in power flow calculation

Let us assume that there are \( n \) buses in the power system; in the DSB there is no more a bus defined as a 'slack bus', so 0 to \((m - 1)\) are generator buses that take care of the overall real power imbalance in the power system, while buses from \( m \) to \((n - 1)\) (including load buses, buses connected to renewable resources, and remaining generator buses) are not involved in the imbalance sharing. In DSB, modification in the formulation is only relevant to the real power part of the Jacobian matrix as follows [106].
where $P_B$ is the total real power imbalance (real power mismatch) and $\Delta P_B$ represents the change in power mismatch.

Each element of the first column of the modified Jacobian matrix is the participation factor of each generator in the real power allocation process and zero otherwise [106]:

$$\frac{\partial P_i}{\partial P_B} = \alpha_i \sum_{i=0}^{m-1} \alpha_i = 1 \quad (5.45)$$

An iterative process is then adopted to solve the real PF equations (5.44) and, after each iteration, the voltage magnitudes, voltage angles as well as the power imbalance are updated (by solving the reactive subproblem, according to the conventional model).

### 5.3.3 Distributed slack bus probabilistic power flow

As discussed, the uncertainty of nodal power injections in modern power systems is significant due to the deepening penetration of renewable resources such as wind and photovoltaic solar together with the probabilistic behaviour of loads and the random outages. These sources of uncertainty will affect the total power imbalance in the system and make it not deterministic as considered in conventional PF computation. Moreover, under the frequency control the power imbalance will be charged to generators participating in the regulation process; therefore, the power outputs of these generators should obey certain probabilistic phenomena (i.e., certain probability distributions).

Figure 5.20 describes what differences inside PPF computations for single slack bus (presented in Section 5.3.2, denoted by SSBPPF) and DSB (denoted by DSBPPF). In the figure, $g$ and $g'$ are the functions mapping between state variables and power injections in SSB and DSB. Let us assume that the system has $n$ injected powers, in which $m$ injections from dispatchable generators, $r$ injections from renewable variable resources, and the rest with $\ell$ injections from loads. In the part of SSB, bus 0 is the slack bus. Injections from renewable resources and loads are probabilistic, so the slack bus injection is probabilistic too because, it is in charge of all power mismatch, losses, and uncertainty in the system, while power outputs of other dispatchable generators (from 1 to $m-1$) are deterministic. On the contrary, in the DSBPPF computation, all power mismatch, losses, and combination of uncertainty of all injections are characterized by total power imbalance $\tilde{P}_B$. In any case, the power imbalance in the system is shared
by \( m \) dispatchable generators according to their participation factors \( \alpha_i \), resulting the outputs of their generators as follows:

\[
\tilde{P}_{gi} = P_{gi}^0 + \alpha_i \tilde{P}_B \quad (5.46)
\]

where \( P_{gi}^0 \) is the expected value of \( \tilde{P}_{gi} \) obtained by solving PF with DSB.

The procedure for DSBPPF is as follows [99]:

- Run PF with DSB to obtain:
  - \( \bar{x} \): the expected value of the random state variables,
  - \( \bar{P}_B \): the expected value of the power imbalance,
  - the sensitivity matrix \( [S']|_{\bar{x},\bar{P}_B} \) (the inverse of the modified Jacobian matrix in DSB) and the sensitivity matrix of power flows with respect to nodal power injections \( [T']|_{\bar{x},\bar{P}_B} \) computed at the expected value \( \bar{x} \) and \( \bar{P}_B \);
- Calculate the self and joint cumulants of nodal power injections;
- Compute cumulants of state variables, line power flows, and total imbalance;
- Calculate cumulants of power output of each generator injection based on (5.46);
- Obtain p.d.f.s and/or c.d.f.s of the outputs of interest using the methods presented in Section 2.9 depending on the type of distributions considered.

From p.d.f.s and c.d.f.s obtained by DSBPPF, probability of line overloading, probability of over-/under-voltage, probability of not meeting ramping requirements and
probability of violation of over-/under-regulation limit of generators can be assessed,
e.g., for the next hours, based on the forecast, the electricity market output, etc.

5.3.4 Tests and numerical results

To illustrate the application of the DSBPPF proposed, a couple of case studies on
the modified IEEE-14 bus test system and on the power system of Sicily in Italy are
carried out.

Case 1: Modified IEEE 14-bus test system

The branch and generator data of the modified IEEE 14-bus test system are pre-
sented in Appendix C. Five wind farms were added to the system at buses 9, 10, 12, 13,
and 14 with the nominal powers provided in Appendix C and depicted as in Fig. 5.21.
The information about wind power production, load, and their uncertainties for a con-
sidered look-ahead horizon are assumed to be known by forecast techniques. The infor-
mation about correlation among load/wind production are assumed as well and could
also be obtained by time series techniques. Nevertheless, they are beyond the scope
of this research. Here, we assume that a forecast technique is available and provides a
15-minute (or other time steps) sequence of load and wind production forecasts, their
expected forecast errors, and correlations during the considered horizon. For the sake
of simplicity, the forecast errors are assumed normally distributed but it is possible to
adopt any other distributions. In addition, the scheduled power output of generators are
assumed to be known. In the system, generators at buses 1 ($G_1$) and 2 ($G_2$) play the
role of the distributed slack so that, in any case, the power imbalance is shared among
them with corresponding participation factors.

Let us assume that at time $t_0$ a forecast for both wind and loads is available for time
$t_k$ and $t_{k+1} = t_k + \Delta t$. Depending on the time frame considered, the value for $t_k$
may range from one hour to a dozen of hours, as well as the $\Delta t$ values may range from 5
minutes to an hour. According to the time frame considered, the values of uncertainty
are likely to be very different. However, the model can take into account all these
cases; in the present work, just an application is shown. Assume that for time $t_k$ the
information about wind power and load forecasts are provided as in Table 5.7 and 5.8,
respectively. The forecast errors are assumed normally distributed represented in terms
of standard deviation (in %) from the mean. In this test, the correlation coefficients
among wind resources are assumed to be 0.5 and those among loads are shown in
Table 5.9.

<table>
<thead>
<tr>
<th>bus</th>
<th>9</th>
<th>10</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>wind power (MW)</td>
<td>24</td>
<td>12</td>
<td>9</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>error (%)</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

For the next time step $t_{k+1}$, load of power system is assumed to be increased 8%
and the forecasting error is kept equal while wind power forecasts are provided as in
Table 5.10. The real power schedules of dispatchable generators at buses 1 and 2 are
assumed as in Table 5.11.
The first test aims at understanding DSBPPF in detail and at exploring the impacts of explicit representation of correlations on the output: the correlations among loads and among wind are taken into account, while contingencies are not considered at this point.

The c.d.f.s and/or p.d.f.s of generator power outputs, power flows, and voltages can be obtained by the proposed DSBPPF. The power imbalance in the system is shared by $G_1$ and $G_2$ according to their participation factors. For instance, at time step $t_{k+1}$, the
5.3. Distributed slack bus probabilistic power flow

Table 5.9: Correlation coefficients among loads

<table>
<thead>
<tr>
<th>bus</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.05</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
<td>0.20</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.00</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.10: Wind power forecasts at time step $t_{k+1}$

<table>
<thead>
<tr>
<th>bus</th>
<th>9</th>
<th>10</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>wind power (MW)</td>
<td>25</td>
<td>14</td>
<td>10</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>error (%)</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 5.11: Real power schedules (MW) at the considered time steps

<table>
<thead>
<tr>
<th>gen</th>
<th>$G_1$</th>
<th>$G_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_k$</td>
<td>238</td>
<td>34</td>
</tr>
<tr>
<td>$t_{k+1}$</td>
<td>249</td>
<td>43</td>
</tr>
</tbody>
</table>

power imbalance computed according to a deterministic PF with DSB is 1.21 MW. This imbalance, depending on losses, is shared among $G_1$ and $G_2$ with $\alpha_1 = 0.55$ and $\alpha_2 = 0.45$, resulting in the following expected values: $P^0_{g_1} = 249.67$ MW and $P^0_{g_2} = 43.54$ MW. At this point, the probabilistic part of the DSBPPF is carried out, resulting in p.d.f.s shown in Fig. 5.22 and 5.23 for $\tilde{P}_{g_2}$ at time steps $t_k$ and $t_{k+1}$, respectively. In this test, the base power of 100 MVA is used.

The proposed method is presented here with reference to the evaluation of security at times $t_k$ and $t_{k+1}$, as well as for ramping constraint evaluation between times $t_k$ and $t_{k+1}$. At the same time, in order to assess the accuracy of the proposed method, a MCS with 30,000 samples has been carried out and taken as reference. The ARMS error is computed similarly to (5.43) to compare the DSBPPF and the MCS results.

The curves obtained by the proposed method for $\tilde{P}_{g_2}$ at the two time steps are very close to the results obtained by MCS: the ARMS values are $1.75 \times 10^{-2} \%$ and $1.34 \times 10^{-2} \%$, respectively. The expected values of $\tilde{P}_{g_2}$ at $t_k$ and $t_{k+1}$ are 34.30 and 43.54 MW, respectively; they take into account both the share of the losses (determinis-
tic amount) and the uncertainties on both wind generation and loads (with correlation), that provide the shape of the p.d.f.s.

Figure 5.22: p.d.f.s of $\tilde{P}_{g_2}$ at time step $t_k$

Based on the output of the DSBPPFs, we can first assess the probability of violation of over-/under-regulation limits of generators. Assume that the upper limit of the regulation band of generator $G_2$ is 0.6 p.u. (the vertical line in Fig. 5.22 and 5.23): we can
5.3. Distributed slack bus probabilistic power flow

calculate the probability that $\tilde{P}_{g_2}$ is higher than its upper limit as

$$\begin{align*}
  t_k &: \mathbb{P}\{\tilde{P}_{g_2} > 0.6\} \simeq 0\% \\
  t_{k+1} &: \mathbb{P}\{\tilde{P}_{g_2} > 0.6\} = 4\%
\end{align*}$$

The above increase mostly results from the increase of load as well as forecast errors of wind: the forecast errors impact on security of the system.

The DSBPPF allows us to evaluate the probability of rampability violation for generator $G_2$ from time step $t_k$ to time step $t_{k+1}$ depending on $\Delta t = t_{k+1} - t_k$. For example, if $\Delta t = 15$ minutes, from the distributions of power output at the two time steps, the p.d.f. of ramping can be obtained as in Fig. 5.24, thanks to cumulant properties applied to a linear combination of two r.v.s $\tilde{P}_{g_2}^{(t_1)}$ and $\tilde{P}_{g_2}^{(t_2)}$. Analogously, assume that the maximum ramp rate for increasing the output of $G_2$ is 0.4 p.u./15 min (the vertical line in Fig. 5.24): the probability of rampability violation is: $\mathbb{P}\{\tilde{R}_{g_2} > 0.4\} = 0.7\%$ (the filled area in Fig. 5.24).

![Figure 5.24: p.d.f. of ramping $\tilde{R}_{g_2}$ of generator $G_2$.](image)

For illustration of power flow and voltage security assessments, the c.d.f.s and/or p.d.f.s of a selected voltage at bus 9 ($\tilde{V}_9$) and real power flow through line 2-3 ($\tilde{P}_{2-3}$) at $t_{k+1}$ are presented in Fig. 5.25 to 5.27 respectively. If the upper bound of the real power flow (e.g., owing to thermal rating) of line 2-3, for instance, is 0.93 p.u. (the vertical line in Fig. 5.26 and 5.27), then the probability being greater than its upper bound is zero at $t_k$, but at $t_{k+1}$ is 2.3%, mainly due to the increase of load. In this test, as the operating range of voltage at bus 9 is [0.94, 1.07] p.u., voltage at bus 9 is within the range. The ARMS calculated for $\tilde{V}_9$ and $\tilde{P}_{2-3}$ above are $6.31 \times 10^{-2}\%$ and $1.59 \times 10^{-2}\%$, respectively, showing the very good accuracy of the proposed method.

The impacts of explicit representation of correlations among nodal power injections from wind power generation and among loads on the outputs, for example, $\tilde{P}_{g_2}$, $\tilde{P}_{2-3}$,
Figure 5.25: c.d.f. s of $\tilde{V}_9$ at time step $t_{k+1}$

Figure 5.26: p.d.f. s of $\tilde{P}_{2-3}$ at time step $t_{k+1}$

$\tilde{Q}_{2-3}$ (reactive power flow through line 2-3), and $\tilde{V}_9$ at $t_{k+1}$, are evaluated and compared in Fig. 5.28 to 5.31. It is clearly indicated that the correlation, in this test, makes the range of variation of these outputs larger.

The second test aims at considering intensively the effects of contingencies: the contingencies due to the random branch outages and generating unit outages are taken into account. The power outputs of $G_1$ and $G_2$ and random branch outages are modeled as presented in Section 5.1.4. In this case, the time frame considered is short, so ORR
is used instead of FOR given in [87]. It is assumed that the values of ORR at $t_k$ and $t_{k+1}$ are as in Table 5.12.

It is worth noticing that, in the first test (without contingencies), the Gram-Charlier series expansion is chosen to obtain the p.d.f.s and c.d.f.s of the outputs of interest. However, with the presence of contingencies (discrete distributions, i.e., the binomial distributions of generating units and the 0-1 distributions of branch outages), the approximation method based on Von Mises function [24, 78] should be adopted, as discussed in detail in Section 5.2.5. A threshold is used to check nodal voltage angle
Chapter 5. Probabilistic Security Assessment

**Figure 5.29:** Impacts of explicit representation of correlations on $\tilde{P}_{2-3}$ at $t_{k+1}$

**Figure 5.30:** Impacts of explicit representation of correlations on $\tilde{Q}_{2-3}$ at $t_{k+1}$

**Table 5.12:** Outage replacement rate

<table>
<thead>
<tr>
<th>ORR(%)</th>
<th>$t_k$</th>
<th>$t_{k+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>lines</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>$G_1$</td>
<td>1.23</td>
<td>1.28</td>
</tr>
<tr>
<td>$G_2$</td>
<td>1.37</td>
<td>1.43</td>
</tr>
</tbody>
</table>
5.3. Distributed slack bus probabilistic power flow

Figure 5.31: Impacts of explicit representation of correlations on $\tilde{V}_9$ at $t_{k+1}$

variations when accounting for discrete distributions caused by random outages (similarly method M2b presented in Section 5.2.3), i.e., nodal voltage angle variations are compared with the predefined threshold; if they are larger, a new deterministic PF with DSB is carried out and the resulting sensitivity matrices are adopted. In this test, the threshold of 0.05 radian is used.

The p.d.f. of $\tilde{P}_{g2}$ at $t_{k+1}$ is shown in Fig. 5.32 with a comparison to the case without contingencies (Fig. 5.23). The filled area describes the difference in probabilities that $\tilde{P}_{g2}$ is higher than its upper limit (i.e., 0.6 p.u.): 4% (without contingencies) and 5% (with contingencies). Moreover, looking at Fig. 5.32, the curve with contingencies is deviated to the right from the curve without contingencies: the corresponding expected values calculated for two cases are 44.47 MW and 43.54 MW. This is mainly due to the generating unit outages which make the expected values of other generators decreased, so the expected value of $\tilde{P}_{g2}$ increases.

Analogously, the probability of not meeting ramping requirements of generator $G_2$ from time step $t_k$ to time step $t_{k+1}$ is also evaluated. The p.d.f.s of ramping $\tilde{R}_{g2}$ for the two cases are compared in Fig. 5.33. From the figure, it can be seen that the range of variation of ramping is larger (p.d.f. with long tails on both sides) in the presence of contingencies, than that in the case without contingencies. The probability that the ramp is higher than its upper bound (i.e., 0.4 p.u./15 min) increases to 2.3%, compared to the value of 0.7% in the case without contingencies. On the other side, if the lower bound is assumed to be -0.4 p.u./15 min, the probability that the ramp is lower than the lower bound will be 1.5%, while the value is 0% in the normal case.

The c.d.f.s and/or p.d.f.s of $\tilde{P}_{2-3}$ and $\tilde{Q}_{2-3}$ in the presence of contingencies are presented in Fig. 5.34 to 5.36. Taken as examples, the comparisons between the two cases (with and without contingencies) for $\tilde{V}_9$ and $\tilde{P}_{2-3}$ at $t_{k+1}$ are depicted in Fig. 5.37 and 5.38 respectively. Moreover, for assessing the effects of random branch outages
Figure 5.32: Impacts of contingencies on $\tilde{P}_{g_2}$ at $t_{k+1}$

Figure 5.33: Impacts of contingencies on ramping $\tilde{R}_{g_2}$ of generator $G_2$

on the system, the calculations with different values of $ORR$ of branches at $t_{k+1}$, i.e., 0.14% and 0.5%, are also implemented and compared. From the figures, it can be seen that, in the presence of contingencies, $\tilde{V}_9$ probably tends to be lower ($p.d.f.$ with a long left-hand tail), while $\tilde{P}_{2-3}$ tends to increase ($p.d.f.$ with a long right-hand tail): the probability being greater than the upper bound of $\tilde{P}_{2-3}$ increases from 2.3% (without contingencies) to 3.0% ($ORR = 0.14\%$) and 4.9% ($ORR = 0.5\%$). For evaluating the accuracy, the ARMS calculated for $\tilde{P}_{2-3}$, $\tilde{Q}_{2-3}$, and $\tilde{V}_9$ at $t_{k+1}$ ($ORR = 0.14\%$), are
0.10%, 0.13%, and 0.18%, respectively, indicating that the proposed method can give a very accurate result even in the case of contingencies.

Figure 5.34: c.d.f. curves of $\tilde{P}_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies

Figure 5.35: p.d.f. curves of $\tilde{Q}_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies
Figure 5.36: c.d.f. curves of $\tilde{Q}_{2-3}$ at time step $t_{k+1}$ in the presence of contingencies

Figure 5.37: Impacts of contingencies on $\tilde{V}_9$ at $t_{k+1}$
Figure 5.38: Impacts of contingencies on $\tilde{P}_{2-3}$ at $t_{k+1}$
Case 2: Sicilian power system

Sicily is an island in the south of Italy. The proposed method has been tested on the Sicilian network for the year 2015 with data provided by TERNA. In 2015, the Sicilian power system is interconnected to the continental Italian power system (so the continental European power system) via three AC links (380 kV). The network includes: MV, HV, and EHV levels, 539 buses (261 buses connected to generators) and 664 branches. The total load is 4269 MW while the total installed capacity of power plants is 9687 MW, in which, about 5275 MW of fossil-fuelled units (mainly steam turbines, combined cycle gas turbines, and combined heat and power units), 787 MW of hydro units (about 664 MW from pumping units), 1915 MW from wind power resources, 1430 MW from photovoltaic solar resources, and 280 MW from biomass power plants.

For probabilistic data, all loads and wind power injections are assumed to be normally distributed with standard deviations equal to 8% and 12% of the nominal power, respectively. Correlation among wind resources is taken into account with correlation coefficients calculated from historical wind data: they range from 0.34 to 0.88, indicating that injections by wind power generation from different farms are correlated and some of them are strong. Random outages of 170 lines are considered and modeled as presented in Section 5.1.4 with the probability of failure equal to 0.1%, while generating unit outages are not taken into account.

In order to evaluate both the accuracy and the computation efficiency of the proposed method, a MCS with 30,000 samples has been carried out and taken as benchmark. As discussed in Section 5.2.3, because some branch outages may cause great changes for the system, thus affecting the accuracy of the linearization of power flow equations, the nodal voltage angle variations are checked by comparing to a predefined threshold: if they are greater than the threshold, a new deterministic PF is carried out to obtain new sensitivity matrices. As can be seen from Table 5.13, the lower threshold value is used, the more intensive computation is. However, the computation time of the proposed method is much less than that of MCS: this indicates that the proposed method is applicable to large power systems.

Table 5.13: Computation time comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>DSBPPF</th>
<th>MCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (rad)</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Computation time (s)</td>
<td>10.5</td>
<td>14.8</td>
</tr>
</tbody>
</table>

The results obtained show that all bus voltages are within their limits, while few lines are overloaded with low probability. Figures from 5.39 to 5.42 present the p.d.f.s and/or c.d.f.s of a selected voltage at bus 112 ($\tilde{V}_{112}$) and real and reactive power flows between buses 110 and 66 ($\tilde{P}_{110-66}$ and $\tilde{Q}_{110-66}$) using threshold 0.05 radian. The results point out that DSBPPF can provide a good result for a large-scale power system: the curves obtained by DSBPPF are very close to MCS curves; ARMS errors are small for voltage magnitudes (e.g., 0.16% for $\tilde{V}_{112}$) and reactive power flows (e.g., 0.11% for $\tilde{Q}_{110-66}$), and very small in case of real power flows (e.g., 0.08% for $\tilde{P}_{110-66}$). In
particular, Fig. 5.40 shows that $\tilde{P}_{110-66}$ tends to increase with a right-hand tail caused by random branch outages: the probability of overloading is $P\{\tilde{P}_{110-66} > P_{110-66}^{\max}\} = 0.28\%$ ($P_{110-66}^{\max} = 44.5$ MW). This value is low; however, if the power flow through the line increases due to the increase of loads and/or the variations of nodal power injections increases, the probability of overloading will rise and if it is significant, some preventive actions (mentioned in Section 3.1) will be necessary to avoid the violation.

Figure 5.39: p.d.f.s of $\tilde{V}_{112}$ (voltage level: 150kV)

In this test, conventional generators connected to 18 buses are distributed slack, so that any mismatch, and/or any uncertainty, in the system can be shared by their generators with corresponding participation factors. For example, generator at bus 468 is assigned in real power allocation process with the participation factor 0.154, resulting in: its expected power output equal to 100.03 MW (calculated by deterministic PF with DSB) and its p.d.f. shown in Fig. 5.43. Assume that the regulation band of the generator is $\pm 6\% P_{\text{rated}}^{g_{468}}$ (where $P_{\text{rated}}^{g_{468}}$ is the rated power and it is equal to 250 MW), the probabilities of violation of over-/under-regulation limits of the generator are:

$$
\begin{align*}
P\{\tilde{P}_{g_{468}} > P_{\text{up}}^{g_{468}}\} &= 1.06\% \\
P\{\tilde{P}_{g_{468}} < P_{\text{low}}^{g_{468}}\} &= 1.05\%
\end{align*}
$$

Given the scheduling of generators, the participation factors of controlling genera-
tors can be determined and the probabilities of violation of over-/under-regulation limits at each instant of time as well as the probability of not meeting ramping requirements (based on distributions between different time steps as in the previous test case) of their generators can be assessed. Regarding the regulation band of controlling generators, they are determined depending on the time frame considered as well as the feature and ability of generators.
Figure 5.41: c.d.f.s of $\bar{P}_{110-66}$
Chapter 5. Probabilistic Security Assessment

Figure 5.42: p.d.f.s of $\tilde{Q}_{110-66}$
Figure 5.43: p.d.f.s of $\tilde{P}_{g_{166}}$
5.3.5 Conclusions

In this section, the distributed slack bus probabilistic power flow is proposed and tested on the modified IEEE-14 bus test system and on the Sicilian power system. The results are then discussed in detail to illustrate the application of the proposed approach.

Similarly to the cumulant-based PPF methods presented in the previous section, the distributed slack bus probabilistic power flow provides an effective representation of correlation effects (e.g., for load and wind) and is able to take into account the contingencies due to branch and generating unit outages.

The proposed approach can explicitly represent the steady-state behaviour of the frequency control. The probabilistic power flow with distributed slack bus is developed so that any power imbalance in power system can be charged to a set of dispatchable generators, for example, as a response to a signal from the secondary frequency regulation, so reflecting the actual power system operation.

Distributed slack bus probabilistic power flow application results are useful in power system analysis and the assessment of power system security: specific metrics, such as probability of line overloading, probability of over-/under-voltage and, in particular, the probability of not meeting ramping requirements and probability of violation of over-/under-regulation limits of conventional generators in systems with deepening penetration of highly variable generation and load resources.
Conclusions and Future Work

6.1 Conclusions

The deepening penetration of renewable energy sources, such as wind and photovoltaic solar, has introduced additional uncertainty into various study areas of power system. This added uncertainty, together with the conventional sources of uncertainty due to the stochastic nature of both the load and the availability of generation resources and transmission assets, makes clear the limitations of the conventional deterministic power flow in power system analysis and security assessment applications because the deterministic approach does not consider any stochastic factor and source of uncertainty affecting power systems. In order to address these issues, probabilistic power flow methodologies can provide a valuable contribution. Various approaches are investigated in this thesis; in particular, an in-depth analysis and comparison of the cumulant-based methods in probabilistic power flow studies are presented. Extensive testing indicates good performance of probabilistic power flow techniques. First, an accurate solution can be quickly obtained by cumulant-based methods, even for a large power system. Second, the effect of active-reactive coupling in probabilistic power flow has been evaluated: the decoupled techniques cause small effects on angles and real power flows but can significantly affect and make worse nodal voltage magnitudes and reactive power flows accuracy. Moreover, further researches on cumulant methods have been developed to increase the accuracy and the computational efficiency. Cumulant-based probabilistic power flow methods can take into account the correlation between input nodal power injections (e.g., among load and among wind) so the impacts of explicit representation of correlation on the output can be explored. In addition, the contingencies due to branch and generating unit outages can also be straightforwardly accounted for: this gives added value of probabilistic power flow in security analysis. Finally, the probabilistic power flow variants illustrated in this research are found to be suitable for
different applications depending on the type of distributions considered. In particular, the results coming from the application of the probabilistic power flow methods to a small and a large power system allowed the comparison of the different methods under different aspects, i.e., computational burden, accuracy, etc., and the assessment of the most suitable one depending on the application requirements.

Moreover, probabilistic power flow with distributed slack bus is proposed in the thesis so as to explicitly represent the steady-state behaviour of the frequency control. The approach is developed so that any power imbalance in power system can be shared among a set of dispatchable generators, for example, as a response to a signal from the secondary frequency regulation, so reflecting the actual power system operation. In addition, distributed slack bus probabilistic power flow application results are useful in power system analysis and in the assessment of power system security: specific metrics, such as probability of line overloading, probability of over-/under-voltage as well as the probability of not meeting ramping requirements and probability of violation of over-/under-regulation limits of conventional generators in systems with deep penetration of highly variable generation and load resources. The proposed methodology has been extensively tested on both a small system – IEEE-14 bus test system – and a real large-scale power system – the Sicilian power system, giving good results.

To be taken into account in probabilistic power flow, the probabilistic modeling for load, wind production, branch outage, and generating unit outage is required and considered in the research. Among them, building the model for wind power production is one of the most challenging issues, due to its stochastic features. In this thesis, an effective approach to characterize all statistical features of wind data and a spatio-temporal model of wind generation for both planning and operation studies is proposed. In the model, principal component analysis is adopted combined with time series analysis that enables us to capture both spatial and temporal correlations for wind data from multiple wind sites. In building the model for high-dimensional and correlated data like wind data at multiple sites, the problem becomes much more complicated when the dimensions of the data set are very high. The proposed model is useful for both decorrelating data and reducing computational burden. As wind power is a time-varying, intermittent, uncertain, and non-dispatchable resource while wind speed is a non-Gaussian and non-stationary stochastic process with distinct diurnal and seasonal patterns, the proposed model adopts techniques in statistics, i.e., pre-processing and transformation techniques, without any further assumptions: this is why it can be used for any real wind data. The proposed spatio-temporal model provides a very good direction for dealing with many problems in power systems involving wind power production from multiple locations. In terms of applications, the proposed model results can be directly exploited in probabilistic power flows to assess the power system security for the next hours or in the planning environment for studying new network reinforcement taking into account spatio-temporal correlations. Also, such kind of result is very appropriate for solving decision-making problems under uncertainty for both operating and planning horizons.

For the purpose of mapping wind speed to wind power for an entire wind farm or location to be used in planning and operations, an approach to build an aggregate power curve is also developed. The proposed procedure can be done automatically, so reducing time consumption.
6.2 Future work

A number of interesting and promising research areas arising from the study is as follows:

1) **Integrating advanced photovoltaic solar power and load forecast models into probabilistic power flow analyses**: The model of wind power production proposed can be applied to photovoltaic solar power as well as load. In fact, building a model for load is not so challenging as for wind power.

2) **Clustering/selection algorithm**: The proposed model of wind power production can be used to generate a large number of scenarios. It is therefore necessary to choose or develop a clustering/selection technique for obtaining a smaller number of representative scenarios, thus reducing computation time. There are some methods such as backward reduction, forward selection [45,47], etc., using probability metrics, e.g., Kantorovich distance, to solve this problem. Nevertheless, such methods are suitable for a single stochastic process and unsuitable for multiple stochastic processes associated with wind production at multiple locations.

3) **Stochastic/robust model for scheduling**: This work takes advantage of the model of wind power production proposed to define as accurately as possible a scheduling for the next day or for the next time period considered. This is especially important for TSOs and for electricity market operators which are facing, in operating the grid and power markets, high levels of uncertainty. For TSOs, this results in the need to enlarge security margins, which in turn results in higher operating costs (for example in the ancillary service market); for electricity market operators, uncertainty in RES production estimation can result in inaccurate scheduling which in turn can cause economic penalties. That is why stochastic scheduling is at any rate a critical need.

4) **Improvement of wind power production forecast**: Another possible use of the proposed method is the improvement of wind power production forecast and its profile along the network, by means of the association of the current weather conditions provided by wind forecast to a subset of scenarios among those generated by the method.

5) **Reserve requirements for wind power integration**: The deepening penetration of wind resources in power systems introduces one of the major challenges to power system in which the determination of reserve requirements is a current area of interest. This work becomes very promising with the model of wind power production proposed in the thesis.

6) **Integration of different types of probability distribution into probabilistic power flow tool**: In real power systems, there exist many stochastic factors which follow various types of probability distribution. Understanding their stochastic behaviors and building probabilistic models for them to be taken into account in probabilistic power flow computation are critical needs.
IEEE 14-bus test system

The single line diagram of the IEEE 14-bus test system is shown in Fig. 5.7, while the branch data are presented in Table A.1 [2].

Loads at buses 2-6 and 10-14 are assumed to have Gaussian distributions, characterized by corresponding expected value (mean) and standard deviation (percentage of the mean value) as in Table A.2 [87]. On the contrary, load at bus 9 is modeled by a discrete distribution characterized by 5 impulses and by the relevant probabilities as in Table A.3. Figures A.1 and A.2 show the probability mass functions (p.m.f.s) of $\tilde{P}_l$ and $\tilde{Q}_l$, respectively.

Generators at buses 1 and 2 are modeled by binomial distributions with relevant FORs as in Table A.4. Figures A.3 and A.4 show the p.m.f.s of $\tilde{P}_g$ and $\tilde{Q}_g$, respectively.
### Appendix A. IEEE 14-bus test system

#### Table A.1: Branch data for IEEE 14-bus test system

<table>
<thead>
<tr>
<th>from bus</th>
<th>to bus</th>
<th>resistance (p.u.)</th>
<th>reactance (p.u.)</th>
<th>susceptance (p.u.)</th>
<th>transformer tap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0.01938</td>
<td>0.05917</td>
<td>0.0264</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.05403</td>
<td>0.22304</td>
<td>0.0264</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.04699</td>
<td>0.19797</td>
<td>0.0219</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.05811</td>
<td>0.17632</td>
<td>0.0187</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.05695</td>
<td>0.17388</td>
<td>0.0170</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.06701</td>
<td>0.17103</td>
<td>0.0173</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.01335</td>
<td>0.04211</td>
<td>0.0064</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0.20912</td>
<td>0</td>
<td>0.978</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>0</td>
<td>0.55618</td>
<td>0</td>
<td>0.969</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0</td>
<td>0.25202</td>
<td>0</td>
<td>0.932</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>0.09498</td>
<td>0.1989</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>0.12291</td>
<td>0.25581</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>0.06615</td>
<td>0.13027</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>0</td>
<td>0.17615</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0</td>
<td>0.11001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>0.03181</td>
<td>0.0845</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>0.12711</td>
<td>0.27038</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>0.08205</td>
<td>0.19207</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>0.22092</td>
<td>0.19988</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>0.17093</td>
<td>0.34802</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

#### Table A.2: Normally distributed loads for IEEE 14-bus test system

<table>
<thead>
<tr>
<th>bus</th>
<th>type</th>
<th>voltage (p.u.)</th>
<th>real power</th>
<th>reactive power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>real power</td>
<td>reactive power</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m (MW)</td>
<td>σ (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m (MVAr)</td>
<td>σ (%)</td>
</tr>
<tr>
<td>2</td>
<td>PV</td>
<td>1.045</td>
<td>-21.74</td>
<td>9.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-12.70</td>
</tr>
<tr>
<td>3</td>
<td>PV</td>
<td>1.010</td>
<td>-94.20</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-19.00</td>
</tr>
<tr>
<td>4</td>
<td>PQ</td>
<td>-</td>
<td>-47.80</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.90</td>
</tr>
<tr>
<td>5</td>
<td>PQ</td>
<td>-</td>
<td>-7.60</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.60</td>
</tr>
<tr>
<td>6</td>
<td>PV</td>
<td>1.070</td>
<td>-11.20</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-7.50</td>
</tr>
<tr>
<td>7</td>
<td>PQ</td>
<td>-</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>PV</td>
<td>1.090</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>PQ</td>
<td>-</td>
<td>-9.00</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-5.80</td>
</tr>
<tr>
<td>11</td>
<td>PQ</td>
<td>-</td>
<td>-3.50</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.80</td>
</tr>
<tr>
<td>12</td>
<td>PQ</td>
<td>-</td>
<td>-6.10</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.60</td>
</tr>
<tr>
<td>13</td>
<td>PQ</td>
<td>-</td>
<td>-13.50</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-5.80</td>
</tr>
<tr>
<td>14</td>
<td>PQ</td>
<td>-</td>
<td>-14.90</td>
<td>8.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-5.00</td>
</tr>
</tbody>
</table>
Table A.3: Discretely distributed load at bus 9 for IEEE 14-bus test system

<table>
<thead>
<tr>
<th>bus</th>
<th>type</th>
<th>voltage (p.u.)</th>
<th>real power</th>
<th>reactive power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>value (MW)</td>
<td>probability</td>
</tr>
<tr>
<td>9</td>
<td>PQ</td>
<td>-</td>
<td>-13.4</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-19.6</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-30.2</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-34.8</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-37.3</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Table A.4: Binomial distributions for IEEE 14-bus test system

<table>
<thead>
<tr>
<th>bus</th>
<th>type</th>
<th>voltage (p.u.)</th>
<th>unit rating (MW)</th>
<th>FOR</th>
<th>number of units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Slack</td>
<td>1.060</td>
<td>25</td>
<td>0.08</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>PV</td>
<td>1.045</td>
<td>22</td>
<td>0.09</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure A.1: p.m.f. of $\tilde{P}_9$
Figure A.2: p.m.f. of $\hat{Q}_b$
Figure A.3: p.m.f. of $\tilde{P}_{g_1}$.
Appendix A. IEEE 14-bus test system

Figure A.4: p.m.f. of $\hat{P}_{g2}$
The system includes 300 buses and 411 branches. All network data and single diagram are provided in [2]. For probabilistic data, all loads, except loads at bus 9, 162, and 246, are modeled by normal distributions with standard deviations equal to 8% of the nominal power. Discrete loads at bus 9, 162, and 246 with 5 impulses each are shown in Table B.1.
Table B.1: Discrete loads for IEEE 300-bus test system

<table>
<thead>
<tr>
<th>bus number</th>
<th>type</th>
<th>real power</th>
<th></th>
<th>reactive power</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>value (p.u.)</td>
<td>probability</td>
<td>value (p.u.)</td>
<td>probability</td>
</tr>
<tr>
<td>9</td>
<td>PQ</td>
<td>-0.4267</td>
<td>0.10</td>
<td>-0.1911</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.6400</td>
<td>0.15</td>
<td>-0.2867</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.2800</td>
<td>0.30</td>
<td>-0.5733</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.0667</td>
<td>0.25</td>
<td>-0.4778</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.8533</td>
<td>0.20</td>
<td>-0.3822</td>
<td>0.20</td>
</tr>
<tr>
<td>162</td>
<td>PQ</td>
<td>-0.3778</td>
<td>0.10</td>
<td>-0.1911</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.5667</td>
<td>0.15</td>
<td>-0.2867</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.1333</td>
<td>0.30</td>
<td>-0.5733</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.9444</td>
<td>0.25</td>
<td>-0.4778</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.7556</td>
<td>0.20</td>
<td>-0.3822</td>
<td>0.20</td>
</tr>
<tr>
<td>246</td>
<td>PQ</td>
<td>-0.3422</td>
<td>0.10</td>
<td>-0.1911</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.5133</td>
<td>0.15</td>
<td>-0.2867</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.0267</td>
<td>0.30</td>
<td>-0.5733</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.8556</td>
<td>0.25</td>
<td>-0.4778</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.6844</td>
<td>0.20</td>
<td>-0.3822</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Modified IEEE 14-bus test system

The single line diagram of the modified IEEE 14-bus test system is shown in Fig. 5.21 while the branch data are the same as in Table A.1. The network has been used for testing the probabilistic power flow with distributed slack bus in Section 5.3 in which there is no slack bus: buses 1 and 2 are distributed slack. The rating of generating units at buses 1 and 2 are as in Table A.4. Five wind farms were added to the system at buses 9, 10, 12, 13, and 14 with the nominal power provided in Table C.1.

Table C.1: Nominal power of wind farms

<table>
<thead>
<tr>
<th>bus</th>
<th>9</th>
<th>10</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal power (MW)</td>
<td>27</td>
<td>15</td>
<td>12</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

Information about load, wind power, real power schedules, and ORRs of lines and generating units are presented in Section 5.3.4.


Bibliography


Bibliography


