DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

SUBSTATION EXPANSION PLANNING IN POWER NETWORKS WITH RENEWABLE ENERGY GENERATION

by
Kaan YURTSEVEN

July, 2021
İZMİR

SUBSTATION EXPANSION PLANNING IN POWER NETWORKS WITH RENEWABLE ENERGY GENERATION

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science
in Electrical and Electronics Engineering Program

by
Kaan YURTSEVEN

July, 2021
İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled "**SUBSTATION EXPANSION PLANNING IN POWER NETWORKS WITH RENEWABLE ENERGY GENERATION**" completed by **KAAN YURTSEVEN** under supervision of **ASSOC. PROF. DR. ENGİN KARATEPE** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

....

Assoc. Prof. Dr. Engin KARATEPE

Supervisor

....

Asst. Prof. Dr. Emrah BIYIK | Asst. Prof. Dr. Abdül BALIKCI

(Jury Member) | (Jury Member)

Prof. Dr. Özgür ÖZÇELİK

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Engin KARATEPE for all his support, encouragement, invaluable guidance, the devotion of his valuable time, and all the facility he provided during the development of my thesis. His office door was always open whenever there was a question in my mind about my research. He encouraged me to do this research on my own but also kept me in the right direction by using his knowledge and experience during the research. I am grateful to him for his valuable advice that greatly contributed to my academic development.

I also would like to thank my family for their endless support and understanding throughout my study.

Finally, I wish to extend my appreciation to my fiancé for her continuous and unparalleled love, help, and support.

Kaan YURTSEVEN

SUBSTATION EXPANSION PLANNING IN POWER NETWORKS WITH RENEWABLE ENERGY GENERATION

ABSTRACT

A suitable probabilistic scenario set of load demand and natural characteristics of renewable energy is becoming a crucial issue in power system planning studies. The optimal planning of substation expansion is achievable through proper modeling of input parameters which describes the characteristics of the service areas. In that point, the disparity between the required resources for adequate and flexible power systems and the budget available to planners necessitates choosing between different expansion strategies. Under these circumstances, making investment plans with a convenient risk management strategy is crucial in order to maintain the balance between the cost and system adequacy. In this thesis, the SEP problem is handled from a risk management point of view by integrating different investment strategies. On the one hand, the co-existence of PV plants and the load in a service area under three different states such as daytime with clear-sky and no-fault, daytime with abnormal events, and nighttime are incorporated into the stochastic dynamic sub-transmission SEP, in which the decision-maker chooses the risk-neutral and risk-averse strategies, through systematic probabilistic scenario tree generation. The comparison between integrating inherent characteristics of PV plants with and without considering abnormal events into the optimization is performed to show the impact of a suitable probabilistic model on the dynamic nature of investment decisions. On the other hand, the distribution SEP problem in which the decision-maker chooses the risk-seeker strategy is merged with the different geographical configurations of the zones within a mixed-integer linear programming (MILP) framework by using approximate lumped models.

Keywords: Correlation, dynamic planning, mixed-integer linear programming, photovoltaic, stochastic programming, substation expansion planning

YENİLENEBİLİR ENERJİ ÜRETİMİ ENTEGRE EDİLMİŞ ELEKTRİK ŞEBEKELERİNDE TRAFO MERKEZLERİNİN PLANLANMASI

ÖZ

Yenilenebilir enerji üretiminin doğal davranış karakteristiğini ve sürekli değişen yük talebini betimleyen olasılıksal senaryoları içeren girdilerin oluşturulması güç sistemi planlama çalışmalarında önemli bir konu haline gelmiştir. Elektrik şebekelerinde trafo merkezlerinin tahsisi, hizmet alanlarının özelliklerini tanımlayan girdi parametrelerinin uygun şekilde modellenmesiyle elde edilebilir. Bu bağlamda, planlamacıların kullanabileceği bütçe miktarının ve güvenilir güç sistemleri için gerekli kaynakların birlikte gözetilmesi gereksinimi karar vericilerin farklı genişleme stratejileri arasında seçim yapmasını zorunlu kılmaktadır. Bu koşullar altında, yatırım planlarının maliyet ve sistem güvenilirliği arasındaki dengenin stokastik davranışlar gözetilerek sağlanması için uygun bir risk yönetimi stratejisi ile yapılması büyük önem taşımaktadır. Bu tezde, trafo merkez planlanması problemi karar verici tarafından alınan risk seviyesinden kaçınma miktarına göre oluşturulmuş farklı yatırım stratejileri göz önüne alınarak risk yönetimi çatısı altında ele alınmıştır. Bu bağlamda, çatı üstü fotovoltaik sistemlerinin ve yüklerin bir arada bulunduğu hizmet alanları, gündüz açık hava ve hatasız çalışma, gündüz kesikli koşullarda çalışma ve gece olmak üzere üç farklı durumda modellenmiştir ve bu hizmet alanları senaryo tabanlı bir yaklaşım kullanılarak karar vericinin risk-nötr ve riskten kaçınan yatırım stratejilerini seçtiği durumlardaki stokastik dinamik trafo merkezi tahsisi problemine entegre edilmiştir. Buna ek olarak, karar vericinin risk seven stratejisini seçtiği dağıtım trafo merkezlerinin planlanması problemi, bölgelerin farklı coğrafi konfigürasyonları ile karma tamsayılı doğrusal programlama çerçevesinde yaklaşık yığın modeller kullanılarak birleştirilmiştir.

Anahtar Kelimeler: Korelasyon, dinamik planlama, karma tamsayılı doğrusal programlama, fotovoltaik, stokastik programlama, trafo merkezi planlanması

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER ONE
INTRODUCTION

1.1 Background and Motivation

The evolving nature of the power system due to the growth in demand and renewable energy production necessitates planners to select expansion strategies in order to maintain a reliable power system (Abedi, Hosseini, & Jalilvand, 2020). In that manner, substation expansion planning (SEP) holds a critical part of the electric power systems as a result of inevitable uncertainties that are brought about by the increasing renewable energy production (Behal, 2019). Moreover, SEP determines the way of enhancing the capacity of a power system in the interest of adequately supplying the future demand along with paving the way for enabling renewable energy production as much as possible. In order to maintain a reliable network in an environment with an increase in demand and renewable energy production, starting with substations, the power grid must be upgraded properly.

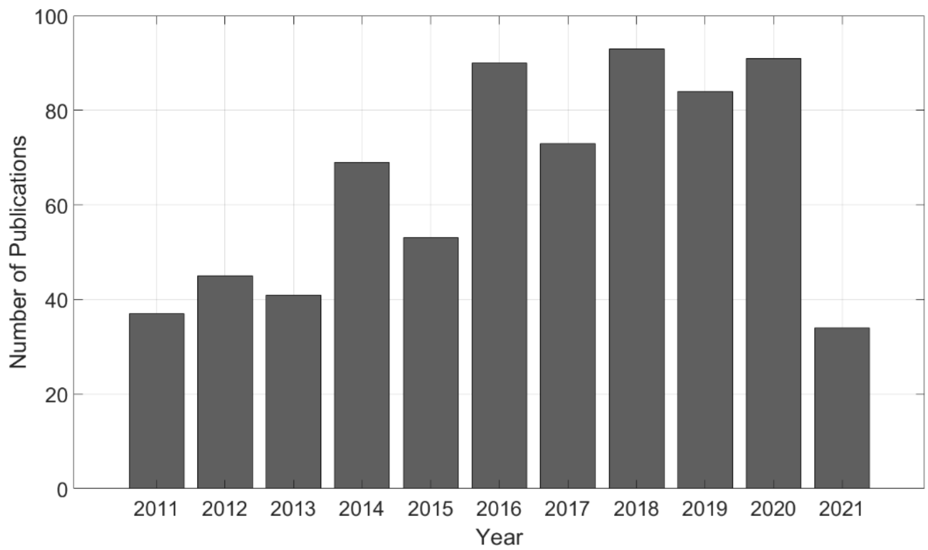

Figure 1.1 The number of publications whose title, abstract, and keywords consist of the words “substation” and “planning” for the last ten years in the Scopus database

The aim of SEP is to determine the necessary expansions in capacities of the existing substations as well as the locations and sizes of new substations while maintaining the network adequacy to supply all the service areas within certain technical limitations (Franco, Rider, & Romero, 2016). Due to the fact that substations form the intermediate system which links the power supplies and service areas, it is evident that substation expansion planning is an essential subject in power system planning studies. Figure 1.1 illustrates the number of publications whose title, abstract, and keywords consist of the words “substation” and “planning” for the last ten years in the Scopus database. It can clearly be seen that substation expansion planning is still an ongoing research area and one of the hot topics of power system planning studies. The studies on the SEP problem in the literature are presented below in terms of solution methods, approaches and the handling of renewable energy sources.

In the SEP problem, power system planners initially should determine the capacity required by the geographically distributed service areas (Li et al., 2018) and then allocate their respective substations (Seifi & Sepasian, 2011). The capacity requirements of service areas simply correspond to the substation capacity and the feeder capacity. Substation capacity requirement refers to the transformer sizes needed in order to maintain a reliable network and adequately supply all the service areas, whereas feeder capacity requirement refers to the need for selecting lines with appropriate technical limits. In that point, it is obvious that a highly complex problem structure is handled in the substation expansion planning. Due to the complex structure of the problem (Temraz & Salama, 2002), SEP is dealt with in the literature from two different perspectives such as substation allocation (Sepasian, Seifi, Foroud, Hosseini, & Kabir, 2006) and optimal feeder routing (Avilés, Mayo-Maldonado, & Micheloud, 2020), which can be handled by long-term and short-term planning, respectively (Haghifam & Shahabi, 2002). In other words, long-term planning deals with the allocation of the substations while approximating the feeder requirements, whereas short-term planning is more about the details in feeder routing on the network (Melgar-Dominguez, Pourakbari-Kasmaei, & Mantovani, 2019). Substation allocation studies are heavily distinguished from optimal feeder routing studies in terms of problem formulation, network existence, planning outputs, etc. The aim of the optimal feeder routing is to extract an optimum network structure to supply the service areas considering thoroughly the limits associated with lines in the network. On the other hand, substation allocation studies provide the planners with an insight into the future of the power system in the long-term by determining the locations and sizes of new and existing substations and approximately connecting them to specific service areas. By doing so, the power system planning in the long-term is handled by considering the players of the power system from the bulk point-of-view without dealing with the details that concern the operational planners.

On the basis of solution methodology, there are two different optimization techniques when dealing with the complex SEP problem such as mathematical (Arias, Tabares, Franco, Lavorato, & Romero, 2018; Seifi & Sepasian, 2011; Zare, Chung, Zhan, & Faried, 2018), and meta-heuristic (Kiani Rad & Moravej, 2017; Hincapie, Gallego, & Mantovani, 2019; Abedi, Hosseini, & Jalilvand, 2019; Esmaeeli, Kazemi, Shayanfar, Haghifam, & Siano, 2016; Salyani, Salehi, & Samadi Gazijahani, 2018; Karimi & Haghifam, 2017; Fan, Dong, & Meng, 2020) optimization. In the study of Kiani Rad & Moravej (2017), a meta-heuristic approach for simultaneous planning of transmission, sub-transmission, and distribution network expansion planning is presented without the effect of the uncertainty characteristics of the power system. A bilevel meta-heuristic approach is presented in the study of Hincapie et al. (2019) for finding a joint global simultaneous solution for primary and secondary networks considering DGs by using a deterministic approach. In the study of Seyed Mahdi Mazhari & Monsef (2013), a method is proposed to solve the dynamic SEP problem by using an algorithm based on learning automata without considering the uncertainties in the power system. A genetic algorithm optimization method is proposed in the study of Sepasian et al. (2006) to solve the SEP problem again from a deterministic point of view. However, recent studies indicate that uncertainty modeling is an integral part of power system planning due to the load and stochastic generation introduced by renewable energy resources. That being said, the planners must decide the level of accuracy of the optimization problem considering the increased model complexity. At that point, the level of the network, on which the planner is focused, is an important factor in deciding whether to model the inputs in a probabilistic way or a deterministic way. Taking uncertainties into account with probability distribution functions in the chance constraint based approaches or requiring sub-programs while modeling the problem forces planners to use meta-heuristic optimizations that are searching a solution space to find the best feasible solution identified (Ugranlı, 2020). In that manner, a new sub-transmission SEP approach is presented in the study of Abedi et al. (2019) that considers both the uncertainties in locations and sizes of the service areas by using meta-heuristic optimization in which mathematical clustering is performed at each iteration. Another way of considering uncertainties in planning problems is to use risk-based optimization techniques by using risk measures. A risk-based meta-heuristic distribution substation planning problem is formulated in the study of Esmaeeli et al. (2016) with a mixed-integer nonlinear programming (MINLP) model that deals with the uncertainties of load growth, failure rate and duration of substations, and energy price growth without considering renewable energy resources. On the other hand, a promising meta-heuristic joint chance constrained programming method is utilized in the study of Salyani et al. (2018) with an MINLP formulation that addresses distribution network expansion planning and allocation of various DG units with uncertainty considering the risk strategies. Moreover, a meta-heuristic dynamic programming method is proposed in the study of Karimi & Haghifam (2017) for optimal timing, siting, and sizing of feeders, substations, DGs, and capacitors simultaneously by considering the uncertainties of wind, load, and market price.

1.2 Challenges with Renewable Energy Generation in SEP

It is important to note that rooftop PV plants are expected to be spread more and more in the service areas due to recent developments in solar energy technologies. Furthermore, not only the load of residential customers but also their rooftop PV plants have started to become an important role in the uncertainty of the power system. For these reasons, the uncertainty in SEP problems should be handled from a prosumer-centered point of view by taking into account the service areas where generation and consumption co-exist simultaneously. It is evident that the optimal planning of the substation expansions is achievable through proper modeling of the input parameters which describes the characteristics of the service areas. The main challenge here is to model a suitable set of scenarios that describes the interrelated behaviors of the load and PV generation to be fed into the planning process. However, despite modeling the stochastic behavior of the PV generation and the load is the basis of the modern power system planning studies, the progress in the development of the suitable probabilistic model fell behind the spread of PV generation in the world. A method is introduced in the study of Fan et al. (2020) that takes into account the DG and load uncertainties by generating a scenario tree and reducing it with Kantorovich techniques in scenario-based stochastic optimization. It is known that there are a finite number of possible results in each stage of the scenario tree when the probability distribution function is discrete. In that manner, a scenario that represents one possible realization can be created by fixing each random variable to one of their possible results. Quantifying all combinations of possible results allows the planners to describe the stack of scenarios as a tree in which each scenario represents a path from the root to one of the leaves. In that point, there are some scenario tree reduction techniques available to deal with the computational effort of considering a large number of possible outcomes (Esmaeeli et al., 2016; Fan et al., 2020; Karimi & Haghifam, 2017; Salyani et al., 2018). The main differences between the aforementioned studies and the uncertainty handling method developed in this thesis are that the service areas are considered from a prosumer-centered point of view and the corresponding scenario tree of the problem consists of a suitable set of scenarios and it is branched from three different bases simultaneously instead of only one as in the conventional approaches. Furthermore, this thesis is a major step towards addressing the gap of modeling the stochastic behavior of the PV generation and the load. This thesis introduces a hybrid model which integrates detailed inherent characteristics including abnormal conditions of PV plants in the service areas into the stochastic dynamic sub-transmission SEP through systematic probabilistic scenario tree generation. In addition, in the presence of different input scenarios, the decision-maker is expected to choose an investment strategy to establish a trade-off between the budget and the risk level taken. Moreover, the risk level here is quantified with the conditional value-at-risk (CVaR) to maintain the linearity of the problem (Rockafellar & Uryasev, 2002). In that point, this study combines the risk-constrained stochastic dynamic sub-transmission SEP problem and MILP framework under one roof. This allows the SEP problem to be solved by conventional branch and cut solvers.

With regard to modeling the inputs, the conventional way of stochastic power system planning is mostly done by considering the inputs independently of each other. On the other hand, some researchers also model the inputs as they are dependent random variables by using Copula functions. However, both approaches lack the impact of abnormal conditions, such as faults, cloudy sky, soling, partial shading, etc., of PV plants on the planning problem. An example of the abnormal events in PV plants is the cloudy-sky condition and due to its obvious effects on the power output of the PV plants, researchers try to create indexes to describe these conditions (Engerer & Mills, 2014) as well as investigating their occurrence rates ("Climatic Atlas of Clouds Over Land and Ocean," n.d.; Reno & Hansen, 2016). It is observed that PV plants encounter a large duration of abnormal conditions resulting in a significantly lower output for PV plants than clear-sky with no-fault conditions (Yurtseven, Karatepe, & Deniz, 2021). At this point, it is apparent that reflecting the abnormal events along with inherent characteristics of PV plants to the planning studies is unavoidable. Since using the Copula theory for modeling correlations between PV generation and load characteristics only acceptable for normal operation of the PV plants, there is still a hole in systematically generating a suitable probabilistic set of scenarios that also considers abnormal events of PV plants to be used in the SEP problem.

In the first part, a novel model is presented that integrates the detailed inherent characteristics of PV plants through systematic probabilistic scenario tree generation into stochastic dynamic sub-transmission SEP under the MILP framework. It is worthy to note that the percentages of abnormal events and daytime duration change when the studied area changes. In this context, the proposed model allows planners to easily modify the stochastic dynamic model to the characteristics of any area by separately assigning the probabilities of both abnormal events and daytime characteristics. This kind of approach in power system planning has great importance to the long-term organizational decision-makers due to its flexibility of adjusting the model to any region systematically with little effort.

In the second part of thesis, distribution SEP problem, in which the decision-maker chooses the risk-seeker strategy, is merged with the different geographical configurations of the zones within a MILP framework by utilizing analytical approximate models. The impact of the configurations on SEP result is investigated.

Based on the proposed SEP model, the contributions of this thesis are summarized as follows:

- Risk management strategy is incorporated into the SEP problem by using CVaR.

- Distribution SEP problem is merged with the different configurations of the zones within a MILP framework by using approximate lumped models and risk-seeker investment strategy.

- A suitable probabilistic set of scenarios that introduces inherent characteristics including abnormal conditions of PV plants in the service areas is integrated into the SEP problem.

- Scenarios are characterized considering abnormal events along with clear-sky and no-fault operating conditions through systematic probabilistic scenario tree generation.

- A stochastic-dynamic SEP problem is formulated using the MILP model.

1.3 Organization of Thesis

The studies within the scope of this thesis are presented in five chapters as follows;

In Chapter 1, the motivation behind this thesis is presented. Afterward, the types of substation expansion planning are discussed. Then, the importance of suitable modeling of the inputs is presented by explaining the relation between the type of the concerned substations and the nature of input parameters. Finally, the role of renewable energy generation in SEP is briefly given.

Chapter 2 introduces stochastic dynamic modeling of the service areas as well as the suitable probabilistic scenario set generation to be fed into the SEP optimization problem. The reason for creating different modeling states for PV generation and load is presented.

In Chapter 3, the problem formulation and the case studies of stochastic dynamic substation expansion planning are presented. The CVaR risk measure and the contribution of risk management to the SEP problem are propitiously discussed. In order to conduct a risk management study, the decision maker is assumed to choose two different kinds of investment strategies which are risk-neutral and risk-averse.

In Chapter 4, the distribution substation expansion planning problem is handled when the decision-maker chooses risk-seeker strategy under the MILP framework. The effect of voltage constraints and the cost of feeder loss on the SEP problem is investigated and the comparison between the SEP problem with rectangular zones and with triangular zones is performed to show the impact of different geographical configurations on the investment decisions. Finally, the results for different cases are presented.

In Chapter 5, the discussion of this thesis study is presented.

In Chapter 6, the general overview of this thesis study and some concluding remarks are presented.

CHAPTER TWO
STOCHASTIC SERVICE AREA MODELING AND SYSTEMATIC SCENARIO TREE GENERATION

In case of uncertainties in the input variables of the power system, it is desirable to assess the system output variables for many load and generation conditions. The problem of representing the uncertainties in the suitable form for quantitative models is often faced by planners in decision-making problems. In that manner, when the uncertainties are described by probability distribution functions the planners have to choose between two possibilities such as either creating a decision model with internal sampling or scenario-based input modeling (Høyland & Wallace, 2001). Most of the studies consider the load and PV generation as if they are independent random variables while some studies consider them dependent random variables and model their dependency with Copula functions (Hou et al., 2019; Othman, Abdelaziz, Hegazi, & El-Khattam, 2015; Park & Baldick, 2020). Considering the load and PV generation as independent random variables leads to introducing unrealistic combinations of the load and PV generation into the scenario-based problem. Even though the load and PV generation own uncertain characteristics, they follow a certain pattern given some environmental limitations. For instance, when the load is at its minimum daily value, the irradiance is almost zero. In the same way, both the irradiance and the load start to increase after sunrise. In that manner, it is obvious that PV generation correlates with the load (Othman et al., 2015). However, this point of view is valid only when the PV plant is assumed to be operating under clear-sky and no-fault conditions. In order to accurately integrate the inherent characteristics including abnormal conditions of PV plants in the service areas, a hybrid model is introduced to the SEP problem.

The smallest components after modules are solar cells in a PV plant. Temperature and irradiance received from a PV site are not always distributed to each PV module evenly, which is an essential issue related to the PV plant's performance. It is known that even a defect in the smallest component results in the decrease in the output power of the system (Xue Lin, Yanzhi Wang, Pedram, Jaemin Kim, & Naehyuck Chang, 2014). It is important to note that the term inherent characteristics are used to describe the natural and environmental causes of changes in the output power of PV plants. In that point, the abnormal events refer to the conditions that the PV plants experience other than clear-sky and no-fault conditions. While abnormal conditions are a broad term describing the performance reductions caused by faults, cloudy-sky conditions, soiling, partial shading, etc., every one of them results in very close outputs to each other in terms of reduced power generation (Yurtseven et al., 2021). Thus, abnormal conditions should be systematically included as a performance reduction in scenario-based long-term planning. Moreover, due to the fact that the output power of PV plants is highly dependent on weather and environmental conditions and also PV plants consist of many interconnected units, it is inevitable to encounter a large duration of abnormal conditions in the long term (European Commission. Joint Research Centre., 2020). At that point, the ever-changing characteristics of abnormal events break the correlated nature between the load and PV generation. Consequently, the relations between the load and PV generation should be simultaneously described with different models that distinguish clear-sky with no-fault conditions and abnormal events (Silvestre, Kichou, Chouder, Nofuentes, & Karatepe, 2015) as well as the daytime and nighttime characteristics (Anukoolthamchote, Assané, & Konan, 2020; Palanichamy & Babu, 2002).

The behavior of the load consumption has been modeled by many researchers such that the long-term load demand has a Gaussian distribution (Dadkhah & Venkatesh, 2012; Dukpa, Venkatesh, & Chang, 2011; Evangelopoulos & Georgilakis, 2014; Hemmati, Hooshmand, & Taheri, 2015; Karimi & Haghifam, 2017; Liu, Wen, & Ledwich, 2011; Salyani et al., 2018; Soroudi & Afrasiab, 2012) with the mean value of a pre-defined load value and a standard deviation of a desired amount of stochasticity. In that point, the trade-off between computational complexity and the model accuracy is established by dividing the Gaussian distribution into a number of scenarios and the probability of each scenario is specified by their corresponding area under the probability distribution function (Dukpa et al., 2011; Evangelopoulos & Georgilakis, 2014; Karimi & Haghifam, 2017; Soroudi & Afrasiab, 2012).

Table 2.1 Summary of the Modeling States

	Modeling States	**PV and Load Dependency**	**PV Generation**	**Load**	**Contribution to the hybrid Model**
1	Daytime with Clear-sky and no-fault	Dependent	Correlated with the load	Correlated with PV generation	Correlated behavior of PV and Load
2	Daytime with Abnormal Events	Independent	Reduced Performance	Gaussian Distribution	Abnormal Conditions
3	Nighttime	Independent	None	Gaussian Distribution	Nighttime characteristics

In this chapter, the characteristics of uncertain input parameters are modeled separately for daytime and nighttime characteristics for long-term planning. In that point, it is known that the mean value of the load during daytime is different from the mean value of the load during nighttime (Anukoolthamchote et al., 2020; Chapaloglou et al., 2019). This work is structured around the idea of systematically generating realistic scenarios of the load and the PV generation to construct the scenario tree of the problem under the MILP framework. In that manner, there are two different Gaussian distributions for the load during daytime and nighttime, and a number of scenarios are generated from these distribution functions with the same approach in (Dukpa et al., 2011; Evangelopoulos & Georgilakis, 2014; Karimi & Haghifam, 2017; Soroudi & Afrasiab, 2012). In this context, the summary of the modeling states, which are incorporated in the optimization problem in the proposed method, are given in Table 2.1. Three different modeling states, each have 1000 samples, are generated for service areas.

2.1 Modeling of The Load and PV for Daytime With Clear-Sky

There are several methods that can quantify the correlations between random variables. The Copula theory is an effective tool to analyze the correlation between random variables (Cai, Shi, & Chen, 2014). In this thesis, the Copula theory is utilized to model the correlation between the load and PV generation during daytime in clear-sky with no-fault conditions. Copula Theory is first proposed in the study of Sklar (1959) as the function that couples one-dimensional cumulative distribution functions (CDFs) to their multivariate joint CDF. Let the marginal CDFs of random variables x_1 and x_2 be $F_1(x_1)$ and $F_2(x_2)$, respectively, and the joint CDF be $F_{12}(x_1, x_2)$. Copula theory indicates that there exists a Copula function C that satisfies (2.1).

$$F_{12}(x_1, x_2) = C\big(F_1(x_1), F_2(x_2)\big) \tag{2.1}$$

Herein, the correlation between the load and PV generation under clear-sky and no-fault is modeled using the most commonly used Copula function which is the Gaussian Copula (Papaefthymiou & Kurowicka, 2009). The marginal PDFs of the load and PV generation are Gaussian and Beta distributions, respectively (Youcef Ettoumi, Mefti, Adane, & Bouroubi, 2002). The illustration of the correlation between the load and PV generation can be seen in Figure 2.1.

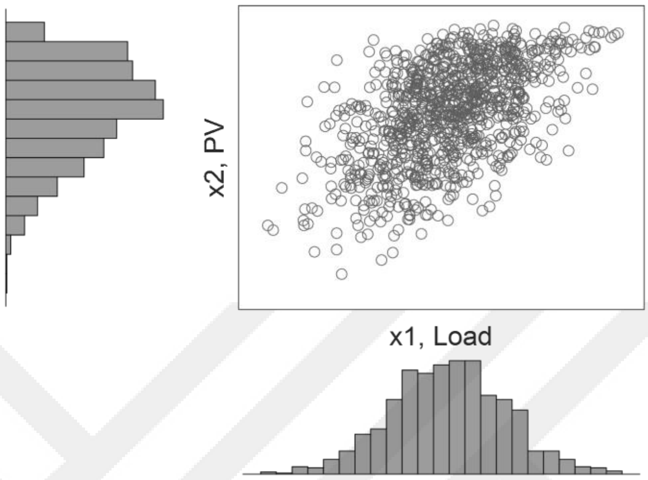

Figure 2.1 Correlation between load and PV generation with Gaussian copula

2.2 Modeling of The Load and PV for Daytime With Abnormal Events

Neglecting the abnormal events (e.g., cloudy-sky, environmental impacts, partial shading, inverter failures, faults in the dc installation, soiling, MPPT failures, global shading, line-to-line faults, AC disconnect, string and module-based failures, recloser tripping) (Klise, Lavrova, & Gooding, 2018; Silvestre, Chouder, & Karatepe, 2013; Yurtseven et al., 2021) while modeling PV generation behavior is inadequate since it is clear that PV generation is highly dependent on various events besides normal conditions. Even though PV generation correlates with the load in clear-sky and no-fault conditions, abnormal events change this dependent relation due to reducing the PV output power drastically. During abnormal events, it is inevitable that PV generation and the load should be considered as independent random variables.

The abnormal conditions consist of a number of events whose occurrence rates are investigated in the literature and it is observed that the cloudy-sky conditions have the maximum occurrence rate among other abnormal events such as inverter and PV module faults and other defects (Geoffrey et al., 2014; Laukamp, 2002; Reno & Hansen, 2016). In that regard, there are a number of studies already available that quantifies the frequency of the cloudy-sky conditions in different regions in the world (Reno & Hansen, 2016; “Climatic Atlas of Clouds Over Land and Ocean,” n.d.). The study in (“Climatic Atlas of Clouds Over Land and Ocean,” n.d.) presents atlases of cloud climatological data obtained from visual observations at the Earth’s surface. In that point, the data in (“Climatic Atlas of Clouds Over Land and Ocean,” n.d.) indicates that the annual average cloud amount is 56% globally during the daytime, resulting in a significantly lower output for PV plants than clear-sky with no-fault conditions. Moreover, the algorithm developed by Long and Ackerman (Long & Ackerman, 2000) and another algorithm developed by Reno and Hansen (Reno & Hansen, 2016) agree on the results that 51% of the data of 2011 experiences clear-sky conditions in California. At this point, it is apparent that reflecting the abnormal conditions along with the other inherent characteristics of PV plants to the planning studies is unavoidable. Nevertheless, integrating abnormal events into planning studies is a complicated challenge since they are region-specific in nature, as shown in Figure 2.2. In other words, the probabilities that represent the percentage of abnormal events’ occurrences differ from region to region (“Climatic Atlas of Clouds Over Land and Ocean,” n.d.) which indicates that the models without the flexibility to be customized according to specific regions are inadequate. Therefore, it is of great importance to develop an appropriate model that is flexible enough to be adjusted to various regions effortlessly and systematically. In that manner, the occurrence rate of the cloudy-sky condition is here conservatively considered as the probability of abnormal conditions due to its dominance over other events. In addition, since the output of each abnormal event is a reduction in the output of the PV plants, it is more than convenient to generate performance reduction scenarios with conservatively assigned probabilities in order to model the abnormal events. It is worth noting that the proposed model can simply be adapted to any area and PV plant site with respect to the desired risk level on the consideration of the included abnormal events by adjusting the probabilities without vitiating the applicability of the model.

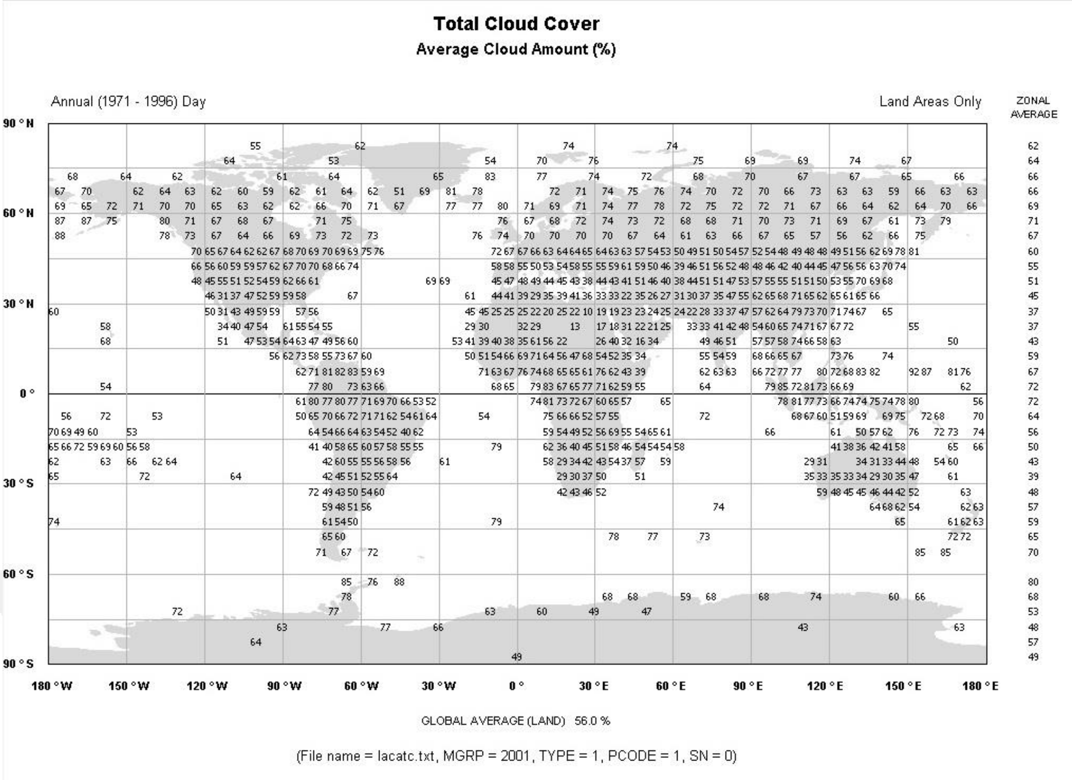

Figure 2.2 Land map of the average cloud amount (%) (Climatic Atlas of Clouds Over Land and Ocean, n.d.)

Due to this intermittent phenomenon which becomes a considerable period in the long term when added up, the output power can be modeled by assuming its input has a Gaussian pdf with a mean of a relatively low irradiance value. The load is also modeled with a Gaussian distribution independently. It is worth noting that the proposed model can simply be adapted to any area and PV plant site with respect to the desired risk level on the consideration of the included abnormal events by adjusting the probabilities without vitiating the applicability of the model.

2.3 Modeling of The Load and PV for Nighttime

In the interest of properly modeling the load and PV generation, it is obvious that nighttime and daytime should be considered separately in the optimization process. During the nighttime, PV generation is modeled with a constant term zero while the load is modeled with a Gaussian distribution with a mean of slightly lower than the mean of the load during daytime (Anukoolthamchote et al., 2020; Chapaloglou et al., 2019).

2.4 Scenario Generation for Stochastic Dynamic SEP

After modeling the necessary load and PV generation characteristics, the service area is modeled by its corresponding injected power. The rooftop PV plants are assumed to be on every house in a service area. Additionally, a typical service area contains fifty distribution transformers such that each distribution transformer feeds approximately one hundred residential customers which each typically occupies between 5 and 7 kW capacity in a sub-transmission substation (European Commission. Joint Research Centre., 2019; Warne, 2005). It is worthy to note that in rooftop PV plants the peak power is limited by the available roof area (Goodrich, James, & Woodhouse, 2012; Jaeger-Waldau, 2019). The injected powers are calculated for every three modeling states individually. Consequently, three different sets which contain a number of injected power scenarios are created, as illustrated in Figure 2.3. The load and PV generations in each distribution transformer are assumed to have the same characteristics. This simplistic approximation is employed here in order to maintain the stochastic dependence (Papaefthymiou & Kurowicka, 2009). Afterward, each set is simply clustered into the desired number of scenarios along with their probabilities in order to ease the dynamic approach in stochastic optimization, as depicted in Figure 2.4.

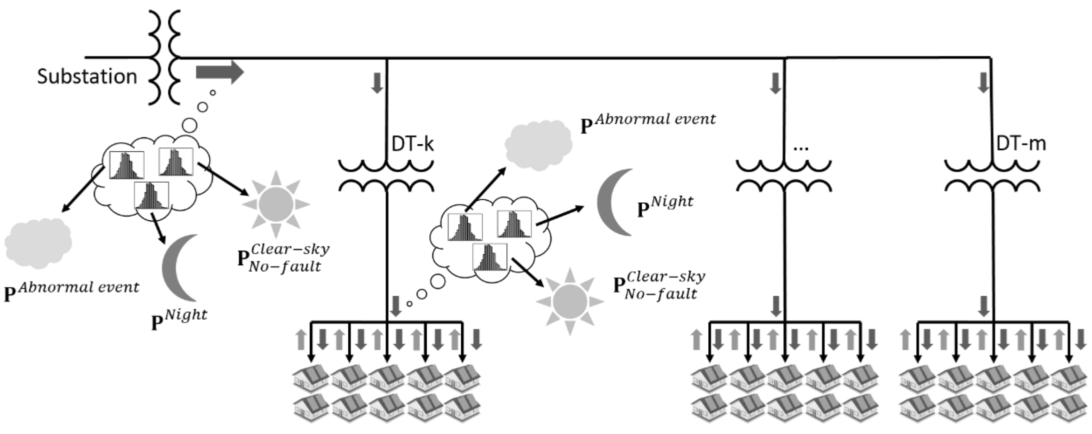

Figure 2.3 A typical service area

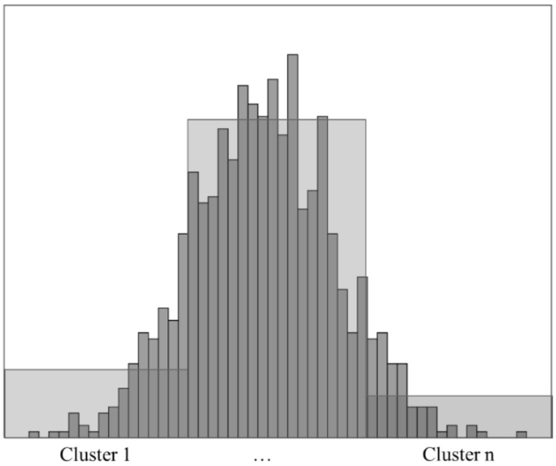

Figure 2.4 Scenario clustering

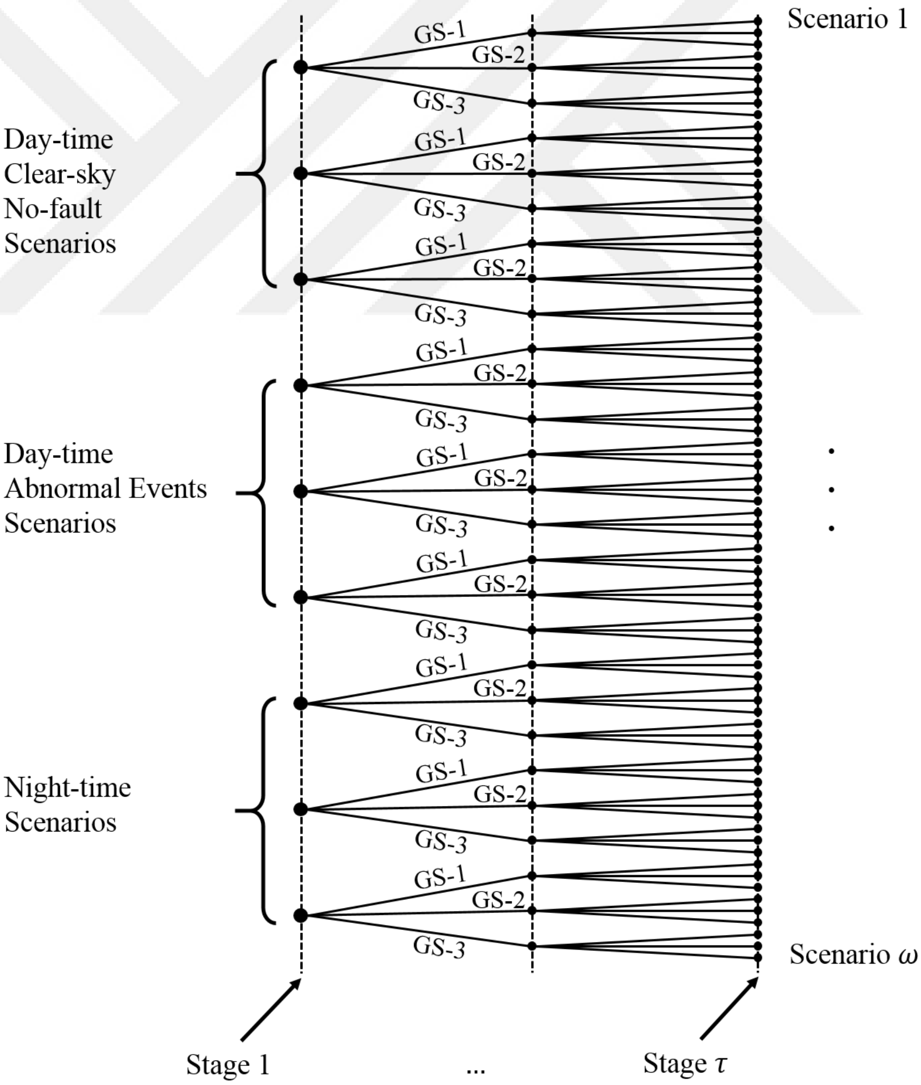

Figure 2.5 Scenario tree of stochastic dynamic SEP (GS: Growth Scenario)

The scenario tree of the problem comprises three stages, nine roots, and ω scenarios. As can be seen from Figure 2.5, some scenarios have common branches in the scenario tree at different stages. Investment decisions are made at the nodes of the scenario tree in stochastic dynamic programming (Høyland & Wallace, 2001). For a given scenario at Stage 1, there are three different investment decisions possible in Stage 2. After the second stage, the decision-maker again faces three different investment decisions in Stage 3 for a given scenario in Stage 2. By doing so, it is ensured that each root of the scenario tree of the problem experiences all the possible growth scenarios.

The discrete probability distribution sets for growth scenarios ($\mathcal{S}_G$), daytime injections under clear-sky and no-fault ($\mathcal{S}_{DCS}$), daytime injections under abnormal events ($\mathcal{S}_{DAE}$), nighttime injections ($\mathcal{S}_{NT}$), sun duration ($\mathcal{S}_{SD}$), and availability condition ($\mathcal{S}_{AC}$) are given as follows:

$$\mathcal{S}_G = \{(G^1, \rho_G^1), (G^2, \rho_G^2), \dots, (G^n, \rho_G^n)\} \tag{2.2}$$

$$\rho_G^1 + \rho_G^2 + \cdots + \rho_G^n = 1 \tag{2.3}$$

$$\mathcal{S}_{SD} = \{(SD^1, \rho_{SD}^1), (SD^2, \rho_{SD}^2)\} \tag{2.4}$$

$$\rho_{SD}^1 + \rho_{SD}^2 = 1 \tag{2.5}$$

$$\mathcal{S}_{AC} = \{(AC^1, \rho_{AC}^1), (AC^2, \rho_{AC}^2)\} \tag{2.6}$$

$$\rho_{AC}^1 + \rho_{AC}^2 = 1 \tag{2.7}$$

$$\mathcal{S}_{DCS} = \{(DCS^1, \rho_{SD}^1 \rho_{AC}^1 \rho_{DC}^1), \dots, (DCS^n, \rho_{SD}^1 \rho_{AC}^1 \rho_{DC}^n)\} \tag{2.8}$$

$$\rho_{DC}^1 + \rho_{DC}^2 + \cdots + \rho_{DC}^n = 1 \tag{2.9}$$

$$\mathcal{S}_{DAE} = \{(DAE^1, \rho_{SD}^1 \rho_{AC}^2 \rho_{DA}^1), \dots, (DAE^n, \rho_{SD}^1 \rho_{AC}^2 \rho_{DA}^n)\} \tag{2.10}$$

$$\rho_{DA}^1 + \rho_{DA}^2 + \cdots + \rho_{DA}^n = 1 \tag{2.11}$$

$$\mathcal{S}_{NT} = \{(NT^1, \rho_{SD}^2 \rho_N^1), \dots, (NT^n, \rho_{SD}^2 \rho_N^n)\} \tag{2.12}$$

$$\rho_N^1 + \rho_N^2 + \cdots + \rho_N^n = 1 \tag{2.13}$$

$$\rho_{DCS} = [\rho_{SD}^1 \rho_{AC}^1 \rho_{DC}^1, \dots, \rho_{SD}^1 \rho_{AC}^1 \rho_{DC}^n] \tag{2.14}$$

$$\rho_{DAE} = [\rho_{SD}^1 \rho_{AC}^2 \rho_{DA}^1, \dots, \rho_{SD}^1 \rho_{AC}^2 \rho_{DA}^n] \tag{2.15}$$

$$\rho_{NT} = [\rho_{SD}^2 \rho_N^1, \dots, \rho_{SD}^2 \rho_N^n] \tag{2.16}$$

$$\sum_n \rho_{DCS} + \sum_n \rho_{DAE} + \sum_n \rho_{NT} = 1 \tag{2.17}$$

$$\mathcal{S}_1 = \mathcal{S}_{DCS} \times \mathcal{S}_G \tag{2.18}$$

$$\mathcal{S}_2 = \mathcal{S}_{DAE} \times \mathcal{S}_G \tag{2.19}$$

$$\mathcal{S}_3 = \mathcal{S}_{NT} \times \mathcal{S}_G \tag{2.20}$$

$$\mathcal{S}_\omega = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \tag{2.21}$$

$$\sum_{\omega \in \mathcal{S}_\omega} \rho_{DCS} \rho_G + \sum_{\omega \in \mathcal{S}_\omega} \rho_{DAE} \rho_G + \sum_{\omega \in \mathcal{S}_\omega} \rho_{NT} \rho_G = 1 \tag{2.22}$$

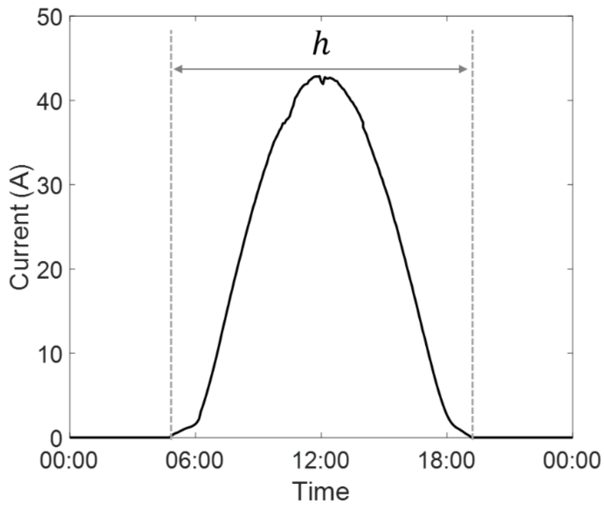

Figure 2.6 A typical current curve of a PV inverter

As a result, a set of possible scenarios $\mathcal{S}_\omega$ is obtained from the combinations of $\mathcal{S}_G$, $\mathcal{S}_{DCS}$, $\mathcal{S}_{DAE}$, $\mathcal{S}_{NT}$, $\mathcal{S}_{SD}$, and $\mathcal{S}_{AC}$. In (2.4), the superscripts of 1 and 2 represent daytime and nighttime, respectively. The probabilities of (2.5) are determined according to the daylight duration percentage. A typical current curve of a PV inverter is given in Figure 2.6. In the figure, h represents the duration that the PV inverter receives irradiance. Therefore, ρ_{SD}^{1} and ρ_{SD}^{2} are calculated as in (2.23) and (2.24), respectively.

$$\rho_{SD}^{1} = h/24 \tag{2.23}$$

$$\rho_{SD}^{2} = (24 - h)/24 \tag{2.24}$$

In (2.6), the superscripts of 1 and 2 represent clear-sky with no-fault conditions and abnormal events, respectively. The probabilities given in (2.7) are assigned by dealing with the frequency of abnormal events of the PV plant in the studied area. An example of five-day PV currents is illustrated in Figure 2.7. In the figure, k represents the duration that the PV plant experiences abnormal conditions. Thus, ρ_{AC}^{1} and ρ_{AC}^{2} is derived as in (2.25) and (2.26), respectively.

$$\rho_{AC}^{1} = (Total\ duration - k)/Total\ duration \tag{2.25}$$

$$\rho_{AC}^{2} = k/Total\ duration \tag{2.26}$$

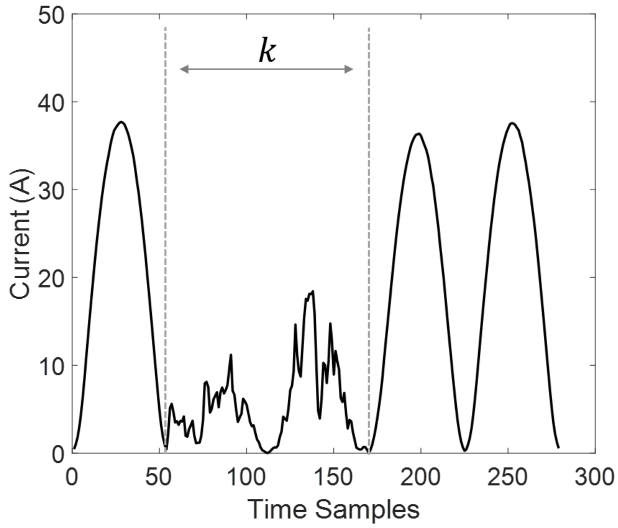

Figure 2.7 An example of five day PV currents

The probabilities in $\mathcal{S}_{SD}$ and $\mathcal{S}_{AC}$ sets can simply be adapted to any area and PV site with respect to the desired risk level on the consideration of the included abnormal events and sun duration without vitiating the applicability of the proposed model. The probabilities ρ_{DC}, ρ_{DA}, and ρ_N are the probabilities of daytime injections under clear-sky with no-fault, daytime injections with abnormal events, and nighttime injections, respectively. The injected powers for the first modeling state in Table 2.1 are calculated by using the joint pdfs, obtained from Copula Theory, of the load and the PV generation. On the other hand, when it comes to the period when the PV plants experience abnormal conditions, the injected powers are calculated by performing Monte Carlo simulation on the independent marginal pdfs of the load and PV generation. Since there is zero PV generation during nighttime the injected powers are obtained from the marginal pdf of the load. At that point, the probabilities ρ_{DC}, ρ_{DA}, and ρ_N are obtained by utilizing a straightforward clustering technique as in (Atwa & El-Saadany, 2011; Awad, EL-Fouly, & Salama, 2017; Esmaeeli et al., 2016; Saber & Venayagamoorthy, 2012) so that the summation of all probabilities are equal to 1, individually. Furthermore, the probabilities ρ_{DCS}, ρ_{DAE}, and ρ_{NT} are calculated as shown in (2.14), (2.15), and (2.16), respectively in order to map the plain power injections onto their corresponding modeling states. The scenarios obtained from the pdfs of different modeling states are crossed with the combinations of growth scenarios in a way that the summation of probabilities of all the resulting is equal to 1.

CHAPTER THREE

RISK-BASED STOCHASTIC DYNAMIC SUB-TRANSMISSION SUBSTATION EXPANSION PLANNING

In this section, formulations of the mathematical model for stochastic dynamic SEP decision-making problem are presented. The total cost of the scenario-based problem has three terms which are the variable cost of lines, the fixed cost of the installed substation, and the variable cost of installed substations as in (3.1). It should be noticed that due to the nature of scenario-based approaches, the cost function is not the same as the objective function of the problem. The cost function here is defined for each scenario separately. Afterward, each cost function is merged together in the weighted summation in the objective function and solved simultaneously with regard to the desired risk management strategy.

$$C_{total}^{\omega} = C_{line}^{\omega} + C_{fix\_sub}^{\omega} + C_{var\_sub}^{\omega} \tag{3.1}$$

The variable cost of lines is calculated as the multiplication of the feeder length and the injected power of the service area with a pricing coefficient as in (3.2). In the equation $X^{\omega,\tau}(i,j)$ is the binary decision variable of feeder connections. If the ith service area is fed by the jth substation it is 1, otherwise 0.

$$C_{line}^{\omega} = \sum_{\tau \epsilon \Omega_t} \sum_{i=1}^{Nl} \sum_{j=1}^{Ns} \varphi_L(i) X^{\omega,\tau}(i,j) D(i,j) S_L^{\omega,\tau}(i) \tag{3.2}$$

The fixed cost of installed substations can also be referred to as the land cost. The land cost is represented by $\varphi_S^f(j)$ as in (3.3) and assumed to be zero for the existing substations.

$$C_{fix\_sub}^{\omega} = \sum_{j=1}^{Ns} \varphi_S^f(j) X_S^{\omega,t_1}(j) \tag{3.3}$$

The variable cost of installed substations is the cost term that depends on the capacity of the substation at a stage as in (3.4). In the equation, $X_{exp}^{\omega,t}(j)$ is the integer decision variable of the expansion number. It represents the number of S_{cap} to be expanded in the jth substation.

$$C_{var\_sub}^{\omega} = \sum_{j=1}^{Ns} \varphi_S^v(j)\left[X_{exp}^{\omega,t_1}(j)S_{cap} - S_{exis}(j)\right] + \sum\nolimits_{\tau=t_2}^{t_3} \sum_{j=1}^{Ns} \varphi_S^v(j)\left[X_{exp}^{\omega,\tau}(j)S_{cap}\right] \tag{3.4}$$

Even though the traditional investment planning problems have an objective to minimize the total cost while maintaining some technical adequacy, introducing renewable energy production to those problems creates uncertainty in the total cost (Lima, Conejo, Langodan, Hoteit, & Knio, 2018). In scenario-based stochastic optimizations, each scenario contains a set of uncertain variables with their probabilities. It is clear that the input scenarios result in a set of cost scenarios with a discrete probability distribution function. In this thesis, these scenarios are introduced systematically and laboriously into the stochastic dynamic SEP problem. At that point, it is important to quantify the risk while minimizing cost function scenarios simultaneously in the problem. The coherent risk measure CVaR provides quantifying a fat tail in the cost distributions and including it in the optimization while maintaining the linearity of the problem. CVaR risk measure is commonly used in the literature for SEP problems (Esmaeeli et al., 2016; Karimi & Haghifam, 2017; Salyani et al., 2018) and it is defined as a risk assessment measure that quantifies the risk of the tail in an investment decision.

Furthermore, CVaR is the additional cost that the planner would need to pay if a worst-case scenario is realized in the future. In that point, CVaR is the calculated value of the weighted average of the worst-case costs of the tail in the distribution. The starting point for the tail of the distribution is described as Value-at-Risk (VaR) and is determined with respect to the confidence level (α) as illustrated in Figure 3.1.

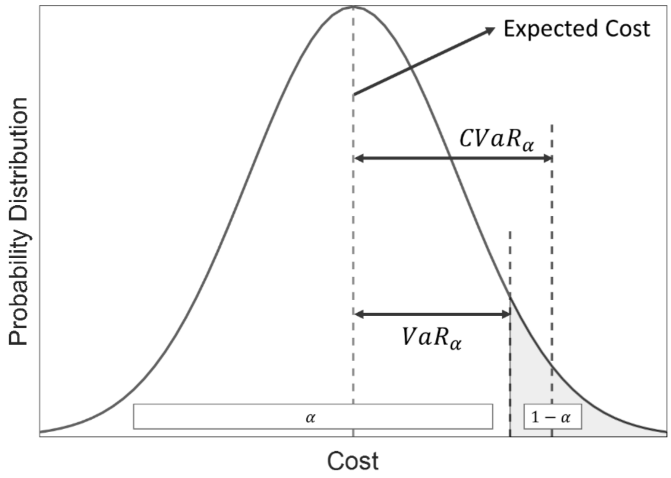

Figure 3.1 Illustration of CVaR on a cost distribution

Since CVaR does not include any nonlinear term, the risk management of stochastic dynamic SEP problem is modeled here by incorporating CVaR risk term and necessary constraints (Conejo, Carrión, & Morales, 2010), resulting in the ability to use a MILP framework (Seifi & Sepasian, 2011) in the scenario-based optimization. The objective function of the proposed model is given in (3.5). It can be seen that for a risk-based SEP problem the objective function is the expected value of the total cost function summed with the CVaR term.

$$\min \sum_{\omega_1}^{\omega_N} \rho_\omega [(1-\beta) \times C_{total}^{\omega}] - \beta \left(\xi - \frac{1}{1-\alpha} \mathrm{E}_\omega\{\eta_\omega\} \right) \qquad (3.5)$$

Subject to

$$X^{\omega,\tau}(i,j) D(i,j) \le D_{max} \qquad \forall\omega, \forall\tau, \forall i, \forall j \qquad (3.6)$$

$$\sum_{i=1}^{Nl} \sum_{\tau\epsilon\Omega_t} X^{\omega,\tau}(i,j) S_L^{\omega,\tau}(i) \le S_{max}(j) \qquad \forall\omega, \forall j \qquad (3.7)$$

$$\sum_{j=1}^{Ns} X^{\omega,\tau}(i,j) = 1 \qquad \forall\omega, \forall\tau, \forall i \qquad (3.8)$$

$$\sum_{i=1}^{Nl} X^{\omega,\tau}(i,j) \le X_S^{\omega,\tau}(j) Nl \qquad \forall\omega, \forall\tau, \forall j \qquad (3.9)$$

$$X^{\omega,t_2}(i,j) \geq X^{\omega,t_1}(i,j) \qquad \forall\omega,\forall i,\forall j \qquad (3.10)$$

$$X^{\omega,t_3}(i,j) \geq X^{\omega,t_2}(i,j) \qquad \forall\omega,\forall i,\forall j \qquad (3.11)$$

$$\xi + C_{total}^{\omega} \leq \eta_{\omega} \qquad \forall\omega \qquad (3.12)$$

$$\eta_{\omega} \geq 0 \qquad \forall\omega \qquad (3.13)$$

$$X^{\omega,\tau}(i,j) = X^{\widetilde{\omega},\tau}(i,j) \qquad \forall\omega,\forall\widetilde{\omega},\forall\tau,\forall i,\forall j \qquad (3.14)$$

If a service area is supplied from a substation that is located far from the service area, the voltage profile along the feeder has a possibility to be over or under the permissible value. Therefore, the voltage profile constraint can be described as in (3.6). Another constraint to be met is the substation maximum capacity constraint as in (3.7). Equation (3.8) represents the radiality constraint which indicates the requirement to feed a service area through only a single substation. Equation (3.9) describes the complementary constraint that is used in order to determine the value of $X_S^{\omega,t}(j)$ to be used in the fixed cost of the installed substation. Equations (3.10) and (3.11) are the dynamic programming constraints which exclude the possibility of shutting down an installed substation throughout the time stages. CVaR constraints are mandatory while incorporating the CVaR minimization approach. The necessary constraints in order to use CVaR risk measure can be seen in (3.12) and (3.13). In (3.5), (3.12), and (3.13), η_{ω} is the auxiliary non-negative continuous variable which describes the summation between VaR and the cost of scenario ω if this summation is non-negative, otherwise, it is equal to zero (Asensio & Contreras, 2016; Conejo, Baringo Morales, Kazempour, & Siddiqui, 2016; Conejo et al., 2010; Li, Yu, Yang, & Wang, 2018; Rockafellar & Uryasev, 2002). Therefore, the summation of the expected values of the scenario costs and the CVaR term refers to the mathematical formulation of the linguistic term which is minimizing the expected cost along with the cost of the worst-case scenarios (Asensio & Contreras, 2016; Conejo et al., 2016, 2010; Esmaeeli et al., 2016; Karimi & Haghifam, 2017; Li et al., 2018; Rockafellar & Uryasev, 2002). In dynamic scenario-based stochastic programming, it is not possible to distinguish between scenarios which have a shared history at a certain stage. It is not logical to obtain different results for scenarios when distinguishing between those scenarios is not possible. Therefore, in dynamic scenario-based stochastic programming, non-anticipativity constraints are used to force the results in a time stage to be the same for the scenarios that have a shared history, as can be seen in (3.14).

3.1 Case Study

In this section, the reason why inherent characteristics including abnormal conditions of PV plants should be integrated into stochastic SEP is revealed through the numerical results. Moreover, the effect of the risk management strategies on investment decisions at different points in time is also investigated. The problem is solved by using Coin-OR Branch and Cut (CBC) solver in Matlab environment. The parameters given in the following are explained in detail with their motives in Chapter 2. It is worthy to note that these parameters depend on the features of the analyzed site and subject to change at will. In the numerical study, the network contains one existing substation and four candidate substations as well as twenty-four service areas. A typical service area is assumed to have fifty distribution transformers such that each distribution transformer feeds approximately one hundred residential customers with a peak demand of 6 kW. A penetration level of 70% is assigned for each rooftop PV plant with regard to the physical limitations of the roof areas. The growth rates in the dynamic problem are given for three stages with their probabilities in Table 3.1. While abnormal events and clear-sky with no-fault conditions occurrence probabilities are conservatively considered equiprobable, daytime and nighttime probabilities are taken as 0.375 and 0.625, respectively. Moreover, in the clear-sky and no-fault conditions, the first and second shape parameters of the Beta distribution are selected as 4 and 2, respectively. On the other hand, in abnormal events, PV output power is modeled by assuming its input has a Gaussian pdf with a mean of a relatively low irradiance value, which is 250 W/m^2 and with a standard deviation of 25. Maximum capacities for substations are 3000 MW and the existing capacity in the network is 100 MW. Additionally, the variable unit cost of substations, land cost of substations, and unit cost of feeders are 2500 \$/pu, 17000\$, and 80 \$/km·pu, respectively.

Table 3.1 Possible growth scenarios

Stage 2		Stage 3	
Percentage	**Probability**	**Percentage**	**Probability**
15%	0.6	25%	0.6
20%	0.3	30%	0.3
30%	0.1	50%	0.1

3.1.1 Case 1: Integrating Inherent Characteristics of PV Plants Including Abnormal Events

In Case 1, the inputs of the stochastic dynamic SEP problem are modeled by using the hybrid model proposed in Chapter 2. Different risk management strategies with different confidence levels are considered in the proposed stochastic dynamic SEP model. In (3.5), β values control the risk management strategy which the decision-maker is willing to take. Risk management is conducted considering two different strategies such as risk-neutral and risk-averse which are formed by assigning the β value to zero and one, respectively. While the risk-neutral strategy indicates that the objective is to minimize the expected cost of the scenarios, the risk-averse strategy indicates that the objective is to minimize the summation of the expected cost and CVaR of the scenarios. In that manner, both risk management strategies here converged into a single infrastructure which is shown in Figure 3.2. There are two unselected substations among the candidates and service areas optimally connected to the corresponding substations. Despite there is an optimum infrastructure for the problem, the capacities of substations are expanded with respect to the level of risk taken. The optimum capacities for the scenarios corresponding to the expected cost are given in Table 3.2 for both strategies. It can clearly be seen that the expansion investments are made only when in need of more capacity throughout the planning horizon in the risk-neutral strategy. On the other hand, in the risk-averse strategy with a confidence level of 0.80, substations 1 and 2 are not expanded in Stage 2 since there is an over-investment in the substation capacities in Stage 1. Similarly, in the risk-averse strategy with a confidence level of 0.85, while substation 1 is not expanded in Stage 3, substation 2 is not expanded in Stage 2 and has a small expansion in Stage 3.

Likewise, substation 4 is not expanded after Stage 1. These are all due to taking a lower risk for what comes next in the future. Finally, in the risk-averse strategy with a confidence level of 0.90, while 87.28% of the overall investment is made in Stage 1, only 68.86% is made in the risk-neutral strategy.

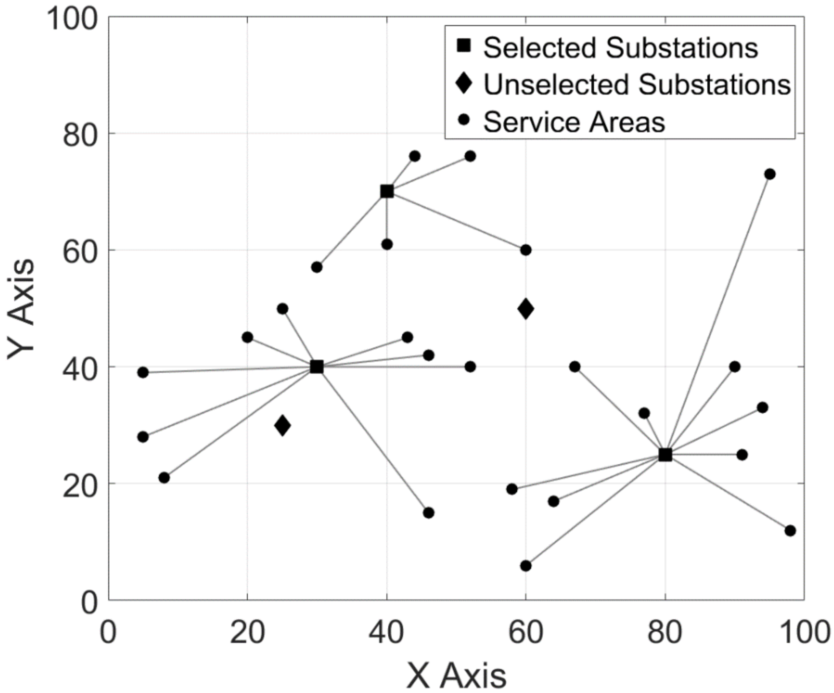

Figure 3.2 Optimum infrastructure for the problem

Furthermore, the optimum capacities for the scenario in which service areas draw the expected power from the network are given in Table 3.3. It is readily apparent that substation capacity investments are made when the existing capacity in a stage is not enough for the load of the next stage in the risk-neutral strategy. On the contrary, in the risk-averse strategy with a confidence level of 0.80, there is no need for any capacity expansion in Stage 2 because with a relatively low risk taken the investments are made beforehand. Moreover, while in the risk-neutral strategy the total amount of the capacity investment is 843 MW, in the risk-averse strategy with a confidence level of 0.85 and 0.90 the total amount of the capacity investment is 891 and 920 MW for the same scenario, respectively. Therefore, it is clear that risk aversion results in over-investments in the scenarios throughout the planning horizon.

Table 3.2 Optimum capacities for scenario corresponding to expected cost in Case 1

	Risk-neutral Strategy			**Risk-averse Strategy**								
				$\boldsymbol{\alpha = 0.80}$			$\boldsymbol{\alpha = 0.85}$			$\boldsymbol{\alpha = 0.90}$		
Subs.	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
1	421	55	19	421	0	74	189	119	0	293	146	0
2	163	102	10	234	0	41	105	0	9	243	0	1
4	325	204	21	468	61	21	333	0	0	486	2	0
Total	909	361	50	1123	61	136	627	119	9	1022	148	1

Table 3.3 Optimum capacities for scenarios in which service areas draw expected power in Case 1

	Risk-neutral Strategy			**Risk-averse Strategy**								
				$\boldsymbol{\alpha = 0.80}$			$\boldsymbol{\alpha = 0.85}$			$\boldsymbol{\alpha = 0.90}$		
Subs.	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
1	162	102	52	293	0	59	162	154	0	180	54	117
2	130	17	29	131	0	10	130	94	0	90	57	29
4	180	113	58	234	0	117	282	23	46	202	62	130
Total	472	232	139	658	0	185	574	271	46	472	172	276

In the interest of investigating the effect of risk management on scenarios as a whole, the probability distributions of the total expansion costs are given in Figure 3.3 when different risk management strategies are chosen. It can clearly be seen that the lowest points in the distributions increase with the decrease in the risk level taken. In the risk-neutral strategy, the scenario costs are distributed over a large interval with low frequencies resulting in higher risks than that of the others. On the other hand, the scenario costs are distributed over a narrower interval in risk-averse strategies, which indicates lower risk levels for the decision-maker.

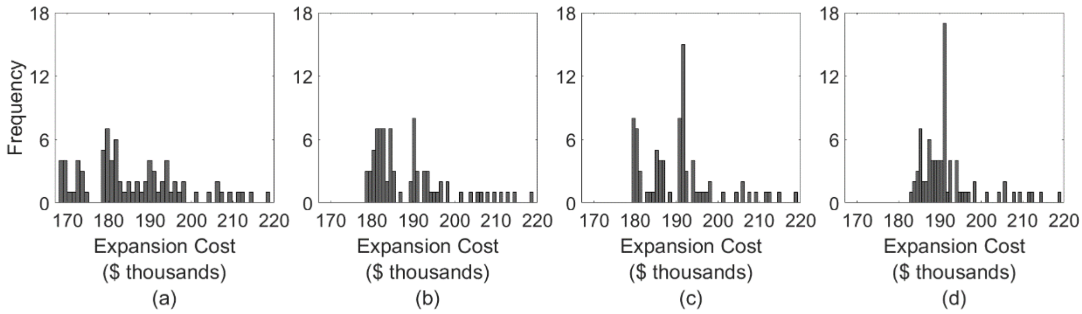

Figure 3.3 The cost probability distributions of Case 1 when risk-neutral (a) and risk-averse strategies with a confidence level of 0.80 (b), 0.85 (c), and 0.90 (d) are chosen

3.1.2 Case 2: Integrating Inherent Characteristics of PV Plants Excluding Abnormal Events

In Case 2, the inputs of the stochastic dynamic SEP problem are modeled by only considering the daytime and nighttime clear-sky with no-fault conditions. Thus, the difference in this case is that the absence of daytime abnormal events. The system converged into an optimum infrastructure as illustrated in Figure 3.2 for both risk management strategies similar to Case 1. However, the expansion capacities of substations are changed due to the distinct input modeling. The probability distributions of the total expansion costs are given in Figure 3.4 when different risk management strategies are chosen. The optimum capacities for the scenarios corresponding to the expected cost are given in Table 3.4. Since this case is studied in order to represent the differences between two different input modeling approaches, the risk management results are not discussed in detail. However, it is obvious that the same risk management characteristics are observed during this case as in Case 1.

Table 3.4 Optimum capacities for scenario corresponding to expected cost in Case 2

	Risk-neutral Strategy			Risk-averse Strategy								
				$\alpha = 0.80$			$\alpha = 0.85$			$\alpha = 0.90$		
Subs.	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
1	162	33	97	126	87	0	195	0	97	389	0	0
2	100	18	54	117	63	0	141	0	21	216	0	0
4	207	63	54	216	206	0	305	0	19	432	0	0
Total	469	114	205	459	356	0	641	0	137	1037	0	0

Table 3.5 Optimum capacities for scenarios in which service areas draw expected power in Case 2

	Risk-neutral Strategy			**Risk-averse Strategy**								
				$\alpha = 0.80$			$\alpha = 0.85$			$\alpha = 0.90$		
Subs.	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
1	162	102	0	260	0	12	195	16	53	324	0	0
2	100	56	0	147	0	6	141	12	0	104	0	113
4	180	113	0	301	0	13	305	0	25	270	81	0
Total	442	271	0	708	0	35	641	28	78	698	81	113

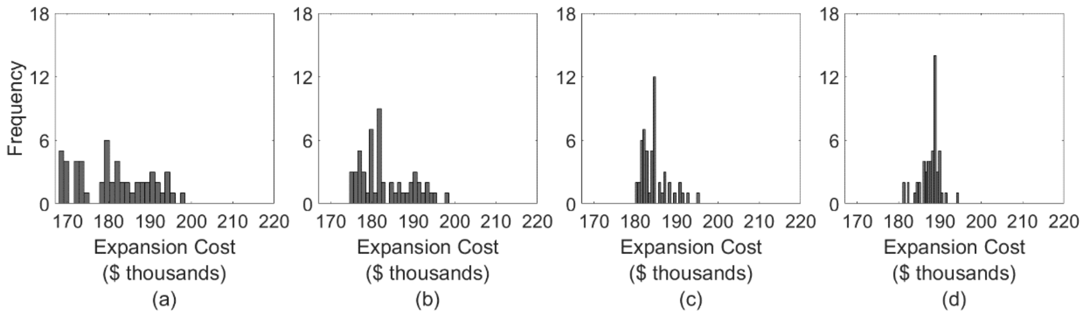

Figure 3.4 The cost probability distributions of Case 2 when risk-neutral (a) and risk-averse strategies with a confidence level of 0.80 (b), 0.85 (c), and 0.90 (d) are chosen.

From Table 3.4, it can clearly be seen that the total investment is significantly lower than that of the investments in Case 1 for scenarios corresponding to the expected cost. This is because daytime inputs are modeled without considering the possibility of abnormal events. In the absence of abnormal events, PV generation maintains its high performance and the injected power drawn from the grid is low. This results in an under-investment if the decision-maker chooses to select the budget for the planning considering the scenarios corresponding to the expected cost. Moreover, the optimum capacities for the scenario in which service areas draw expected power are given in Table 3.5. Similarly, the total investments for each risk management strategy are lower than that of the investments in Case 1.

Due to not considering abnormal events which cause higher injected power drawn from the grid, the planning is again resulted in under investments. As a result, in stochastic planning studies it is crucial to merge the load characteristics and the possibility for a PV plant to experience abnormal events. Otherwise, the investment plans lead to under-investment cases which results in inadequacy of the power system in the future. Moreover, risk management strategies in stochastic dynamic optimization change the diversity of the optimum costs of the scenarios in a way that higher risk levels lead to broader cost distributions and vice-versa. It is also observed that the frequencies of certain costs increase with the decrease in the risk level. Hence, lower risk levels make the scenario cost come closer to each other such that the probability of making under-investments throughout the scenario set decreases. The expected costs and CVaR risk measures of each risk management strategy with different confidence levels for Case 1 and Case 2 are presented in Table 3.6 and Table

3.7, respectively. It can be seen that while lower risk levels increase the expected cost, the CVaR risk measure decreases which indicates that the additional cost of the worst-case scenarios is decreased. By doing so, in the risk-averse strategy, scenarios which have relatively high costs with low probabilities are considered further. Additionally, the expected costs of Case 2 are lower than that of the costs of Case 1 due to lacking abnormal events in the modeling.

Table 3.6 Expected costs and CVaR risk measure of Case 1

	Risk-neutral Strategy	**Risk-averse Strategy**		
		$\boldsymbol{\alpha = 0.80}$	$\boldsymbol{\alpha = 0.85}$	$\boldsymbol{\alpha = 0.90}$
Cex ($)	184,290	185,520	186,870	190,680
CVaR ($)	-	7,848	7,260	5,839

Table 3.7 Expected costs and CVaR risk measure of Case 2

	Risk-neutral Strategy	**Risk-averse Strategy**		
		$\boldsymbol{\alpha = 0.80}$	$\boldsymbol{\alpha = 0.85}$	$\boldsymbol{\alpha = 0.90}$
Cex ($)	180,700	182,180	184,270	187,840
CVaR ($)	-	5,429	5,237	5,009

3.2 Conclusion

This chapter brings together stochastic, dynamic, and risk management approaches taking into account the inherent characteristics including abnormal conditions of PV plants in the service areas in SEP under the MILP framework. Risk management is conducted considering two different strategies such as risk-neutral and risk-averse in a dynamic approach which consists of three stages. It is observed that taking lower risks force decision-maker to make investments in the relatively early stages as well as over-investments for the scenarios in the risk management of stochastic dynamic optimization. Moreover, the inherent characteristics of the PV plants in the service areas are easily integrated along with load uncertainty by using a probabilistic scenario-based approach into the MILP framework. The results show that merging the characteristics of the load and possible abnormal events for the PV plants is essential to model an accurate stochastic dynamic SEP problem. In that manner, without integrating the abnormal events of the PV generation into the problem, the decision-maker is forced to make under-investments.

CHAPTER FOUR

DISTRIBUTION SUBSTATION EXPANSION PLANNING CONSIDERING GEOGRAPHICAL CONFIGURATIONS

Making decisions on expansions of the power system as a result of the growth in electricity demand is essential in order to maintain a reliable network. Since substations take on the role of the intermediate link between electricity supply and demand, substation expansion planning is one of the most important planning studies. In this chapter, the SEP problem is merged with the different geographical configurations of the zones within a mixed-integer linear programming framework (Yurtseven & Karatepe, 2021). The voltage profile and power losses are integrated into the SEP problem by using the approximate lumped model of the zones. The effect of voltage constraints and the cost of feeder loss on the SEP problem is investigated and the comparison between the SEP problem with rectangular zones and with triangular zones is performed to show the impact of different geographical configurations on the investment decisions.

Maintaining a reliable network in an environment with an increase in demand necessitates the proper expansion of the power system. Due to the fact that substations form the intermediate system which links the power supplies and the zones, it is evident that substation expansion planning is an important subject in power system planning studies (Behal, 2019). In this context, SEP is utilized to determine the required expansions in capacities of the existing substations as well as to allocate and size the new substations. While doing so, the adequacy of the power system is maintained to supply all the zones within certain technical limitations (Jalali, Zare, & Hagh, 2014). In order to solve this complex problem, the SEP study can be conducted for transmission, sub-transmission, and distribution substations (Mazhari, Monsef, & Romero, 2015; Seifi & Sepasian, 2011). In that point, this chapter only focuses on the expansion planning of distribution substations. There are different optimization techniques, which can be collected under meta-heuristic and mathematical optimization, to solve the SEP problem. In the study of Fan et al. (2020), an enhanced evolutionary algorithm is proposed to deal with the SEP problem from a static point of view. In the study of Zare et al. (2018), after utilizing effective linearization techniques, the planning problem is solved by a mixed-integer linear programming model. Since the SEP is often a complex and large-scale problem, meta-heuristic approaches need to search an exceptionally large solution space. In order to reduce the computational time, the SEP problem here is formulated by using a mixed-integer linear programming framework including different geographical configurations of the zones.

Recent studies point out that zone modeling is an essential part of the power system planning (Mazhari et al., 2015). Because the SEP problem is naturally complex and highly depends on the modeling of the zones, it is inevitable that any effort to accurately model the zones is invaluable. Due to the increase in the growth of electricity demand, the effect of consumers on the voltage profile and power loss needs to be accurately included in the modeling of the zones. However, modeling a convenient input to represent the demand along with the topology of the zones in power system planning studies is still a challenge. It is crucial to establish a trade-off between detailed models and computational complexity in planning studies. In this context, it is possible to model the zones with approximate lumped models by using load distributions along the feeders without vitiating the applicability of the solution method (Kersting, 2017). Different load distributions along the feeder bring about the need for a suitable representation of voltage profiles and power losses in the problem. In this thesis, the aforementioned phenomena are handled by the integration of triangular and rectangular configurations into the SEP optimization problem under the MILP framework considering the voltage profile and power loss.

4.1 Problem Formulation of Distribution Substation Expansion Planning

The aim of the distribution SEP problem is to determine the necessary expansions of the power system in terms of the substations and adequate feeder connections between substations and zones while minimizing the total investment cost. The zones and a typical problem arrangement can be illustrated as in Figure 4.1.

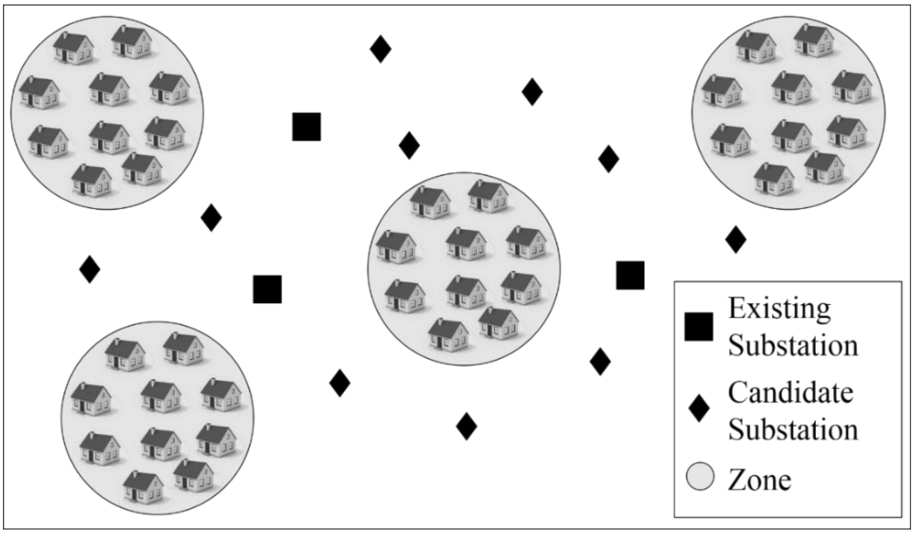

Figure 4.1 The zones and a typical SEP problem arrangement

In a typical SEP problem, there are existing and candidate substations in the network. The existing substations are involved in the problem only with their expansion capacities. On the other hand, the location is also the decision variable for candidate substations. Additionally, the zones are locations where several consumers exist and substations can feed several zones with the obligation that each zone is fed from only one substation and additional connections between zones are not allowed. In this section, the formulations of the distribution SEP problem with geographical configurations are presented under the MILP framework considering power loss and voltage drop. The objective function consists of four terms which are the cost of substations (C_{sub}), the land cost of substations (C_{land}), the cost of feeders (C_{feeder}), and the cost of feeder loss (C_{loss}) as in (4.1).

$$\min \sum_{j=1}^{Ns} \left[\varphi_S^v(j) \left(\sum_{i=1}^{Nl} X(i,j) S_L(i) - S_{exis}(j) \right) \right] + \sum_{j=1}^{Ns} \varphi_S^f(j) X_S(j)$$
$$+ \sum_{i=1}^{Nl} \sum_{j=1}^{Ns} \varphi_L(i) X(i,j) \big(D(i,j) + l(i) \big) S_L(i)$$
$$+ \gamma \sum_{i=1}^{Nl} \sum_{j=1}^{Ns} \varphi_{loss} \left(X(i,j) P_{loss\_geo}(i,j) \right) \tag{4.1}$$

Subject to

$$X(i,j)V_{drop}(i,j) \leq V_{drop\_max} \qquad \forall i, \forall j \qquad (4.2)$$

$$\sum_{i=1}^{Nl} X(i,j)S_L(i) \leq S_{max} \qquad \forall j \qquad (4.3)$$

$$\sum_{j=1}^{Ns} X(i,j) = 1 \qquad \forall i \qquad (4.4)$$

$$\sum_{i=1}^{Nl} X(i,j) \leq X_S(j)Nl \qquad \forall j \qquad (4.5)$$

In the formulation, Ns is the number of candidate and existing substations, Nl is the number of zones, j is the index of substations, i is the index of zones, φ_S^v is the unit cost of a substation, $X(i,j)$ is the binary variable of feeder connections, $X_S(j)$ is the binary variable of the presence of a substation, V_{drop} is the calculated voltage drop along the feeder with respect to the geographical configurations, V_{drop_max} is the maximum allowable voltage drop, S_L is the load magnitude of a zone, S_{exis} is the existing capacity of a substation, φ_S^f is the pre-defined land cost of a substation, X_S is the binary variable that represents the presence of a substation, γ is the total number of hours, φ_L is the cost of the feeder's unit length per unit power transfer capability, D is the distance between the zone and the substation, l is the length of a zone, φ_{loss} is the unit cost of power loss, and P_{loss_geo} is the power loss along the feeder which is specified for each geographical configurations.

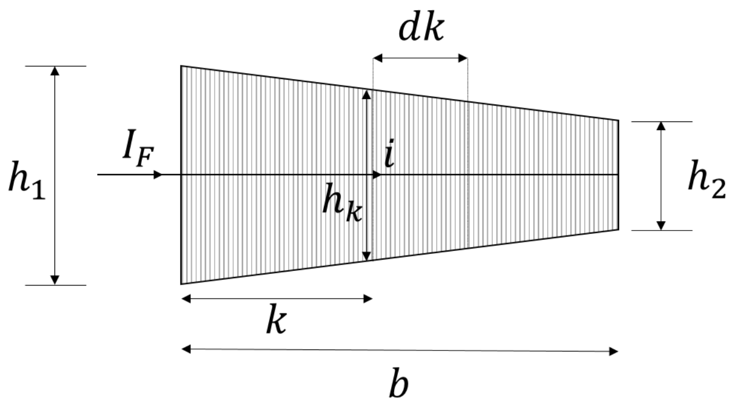

Figure 4.2 Geometric configuration of a feeder

The difficulties in gathering the detailed measured data for each feeder and a large number of complex calculations force planners to use approximate load models. In that point, feeders are often represented by geographical configurations of the load such as rectangles and triangles (Bogdan, Bogdan, Garkavyi, & Netrebko, 2021; Wang & Nehrir, 2004). By doing so, the approximate calculation of power losses and voltage drops along the feeder is made analytically (Gonen, 2014; Kersting, 2017). In order to derive analytical formulations for approximate power losses and voltage drops, a geometric configuration shown in Figure 4.2 is considered. The feeder current at the beginning is calculated as in (4.6).

$$I_F = \frac{Load\ Density \times \frac{1}{2}(h_1 + h_2) \times b}{\sqrt{3} \times kV_{LL}} \tag{4.6}$$

Afterward, the current that is delivered up to k of the total feeder is calculated as in (4.7).

$$I_k = \frac{Load\ Density \times \frac{1}{2}(h_k + h_1) \times k}{\sqrt{3} \times kV_{LL}} \tag{4.7}$$

Hence, a relation between I_F and I_k is derived as in (4.8).

$$I_k = \frac{\frac{1}{2}(h_k + h_1) \times k}{\frac{1}{2}(h_1 + h_2) \times b} \times I_F \tag{4.8}$$

Then, the main goal is to find an expression for the current i which enters the incremental line segment without the use of the unknown value h_k. In that point, the term h_k is expressed by the known parameters as in (4.9).

$$h_k = h_1\left(1 - \frac{k}{b}\right) + \frac{k}{b}h_2 \tag{4.9}$$

Therefore, the current i is calculated as in (4.10) to (4.14).

$$i = I_F - I_k \tag{4.10}$$

$$i = I_F\left(1 - \frac{\frac{1}{2}(h_k + h_1) \times k}{\frac{1}{2}(h_1 + h_2) \times b}\right) \tag{4.11}$$

$$i = I_F\left[1 - \frac{\frac{1}{2}\left(h_1\left(2 - \frac{k}{b}\right) + \frac{k}{b}h_2\right) \times k}{\frac{1}{2}(h_1 + h_2) \times b}\right] \tag{4.12}$$

$$i = \frac{I_F}{(h_1 + h_2)b}\left[(h_1 + h_2)b - 2h_1k + h_1\frac{k^2}{b} - h_2\frac{k^2}{b}\right] \tag{4.13}$$

$$i = \frac{I_F}{(h_1 + h_2)b}\left[\left(b - 2k + \frac{k^2}{b}\right)h_1 + \left(b - \frac{k^2}{b}\right)h_2\right] \tag{4.14}$$

Thus, the voltage drop and the power loss in the incremental line segment is calculated as in (4.11) and (4.12), respectively.

$$dv = Re\{z \times i \times dk\} \tag{4.11}$$

$$dp = r \times i^2 \times dk \tag{4.12}$$

In that point, the voltage drop and the power loss along the feeder is calculated by evaluating the integrals as in (4.13) to (4.14) and (4.15) to (4.16), respectively.

$$\begin{aligned} V_{drop} = \int_0^b dv = \int_0^b Re\left\{\frac{z \times I_F}{(h_1 + h_2)b}\left[\left(b - 2k + \frac{k^2}{b}\right)h_1\right.\right. \\ \left.\left. + \left(b - \frac{k^2}{b}\right)h_2\right]dk\right\} \end{aligned} \tag{4.13}$$

$$V_{drop} = Re\left\{Z \times I_F \times \left(\frac{h_1 + 2h_2}{3(h_1 + h_2)}\right)\right\} \tag{4.14}$$

$$P_{loss} = \int_0^b dp = r \int_0^b \left[\frac{I_F}{(h_1 + h_2)b} \left[\left(b - 2k + \frac{k^2}{b} \right) h_1 + \left(b - \frac{k^2}{b} \right) h_2 \right] \right]^2 dk \tag{4.15}$$

$$P_{loss} = R \times I_F^2 \times \left[\frac{8h_2^2 + 9h_1 h_2 + 3h_1^2}{15(h_1 + h_2)^2} \right] \tag{4.16}$$

In the rectangular feeder area, it is assumed that the load has constant density and uniformly distributed along the feeder, as illustrated in Figure 4.3.

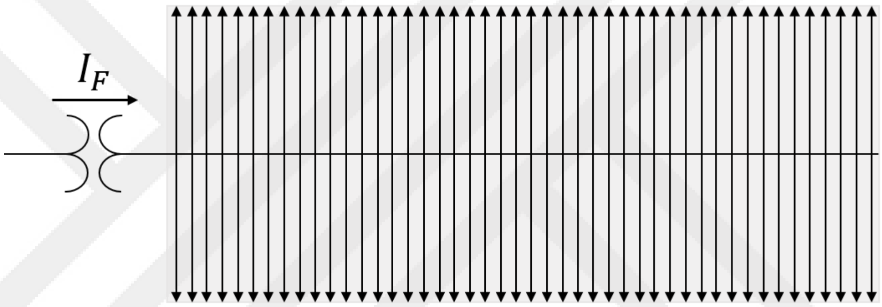

Figure 4.3 Rectangular feeder area

The power loss along the feeder when the zones are rectangular areas is calculated by assigning an equality between $h_1$and h_2 as in (4.17). r is the resistance of the feeder and I_T is the total injected current for the lumped load. Hence, if the feeder area is rectangular, the power loss along the feeder is calculated by either assuming the total load is located at its one-third length of the feeder, or one-third of the total load is at the end of the feeder.

$$P_{loss\_geo}^{rect} = \frac{1}{3} \times r \times l \times I_F^2 \tag{4.17}$$

The distribution of the loads along the feeder does not necessarily have to be uniform. In the triangular feeder area, it is assumed that the load has increasing distribution along the feeder, as depicted in Figure 4.4.

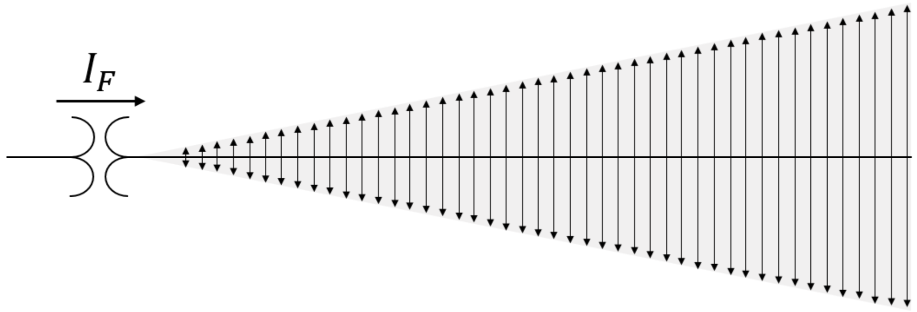

Figure 4.4 Triangular feeder area

The power loss along the feeder with the triangular area is calculated by assigning $h_1$value to 0 as in (4.18).

$$P_{loss\_geo}^{tri} = \frac{8}{15} \times r \times l \times I_F^2 \tag{4.18}$$

The voltage drop along the feeder with uniformly distributed load is calculated by lumping all the load at the centroid of the feeder as in (4.19). In the equation, z is the line impedance per unit length.

$$V_{drop}^{rect} = Re\left(z \times I_F \times \frac{1}{2} \times l\right) \tag{4.19}$$

The voltage drop along the feeder with a triangular area is calculated by lumping all the total triangular load at two-thirds of the length of the feeder as in (4.20).

$$V_{drop}^{tri} = Re\left(z \times I_F \times \frac{2}{3} \times l\right) \tag{4.20}$$

After these calculations, the voltage drop along the feeder is constrained to be lower than a pre-defined value as in (4.2). Another constraint to be met is the substation maximum capacity constraint as given in (4.3). S_{max} is the maximum capacity of a substation. Radiality constraints (Kiani Rad & Moravej, 2017; Mazhari et al., 2015) express the requirement to feed a zone through only one substation as given in (4.4). The last constraint to be met is called the complementary constraint (Seifi & Sepasian, 2011). It is used in order to determine the value of X_S to be either one or zero as in (4.5).

4.2 Numerical Results of Distribution Substation Expansion Planning

In this section, firstly a comparison between the effect of different geographical configurations in the SEP problem is presented. In order to make a fair comparison, the load densities of the triangular areas are assigned to twice the load densities in rectangular areas such that the load values are the same for both geographical configurations. Secondly, the effect of the absence of voltage constraints and feeder loss on the SEP problem is demonstrated. The total loads of the zones are given in Table 4.1. The power factor is assigned to 0.9 lagging. The network in the study consists of 43 zones and 32 candidate substations in which three of them have existing capacities. The resistance and reactance of feeder per km are taken as 0.002 pu and j0.005 pu, respectively. The unit cost of feeder installation and power loss are 1,000$/km·pu and 800$/km·pu, respectively. The length of the zones is considered to be 10 km and equal for each geographical configuration. The distance *D* is calculated by using X and Y values in Table 4.1 and Table 4.2. Additionally, the maximum allowable voltage drop is 5%. The unit cost of substations and land costs depend on the location and the type of substation. Substations can be installed in different types. In this chapter, two different kinds of substations are considered which are outdoor substations and underground substations. In outdoor substations the equipment is installed outdoor resulting in higher maximum capacity than other types of substations. On the other hand, at the locations where there are several zones nearby the available area is limited for an outdoor substation and the land cost is high. Under such conditions, the substation equipment is installed underground. The maximum capacity for underground substations is lower than other types of substations (Mehta & Mehta, 2005). The locations, existing capacities in pu, and the types of the substations are given in Table 4.2. In the table OA, ON, and U stand for outdoor substation away from zones, outdoor substation near zones, and underground substation, respectively. The unit cost of substations and land costs for different types of substations are given with respect to their locations in Table 4.3.

Table 4.1 Locations and total load of zones

Zone	X	Y	S_L [pu]
1	57	46	0.87
2	96	65	0.81
3	48	23	0.59
4	44	84	0.69
5	96	56	0.62
6	18	48	0.59
7	82	64	1.00
8	21	8	0.85
9	43	68	0.80
10	65	67	0.57
11	26	24	0.68
12	86	45	0.49
13	69	47	0.36
14	61	38	0.70
15	52	82	0.93
16	91	7	0.50
17	78	14	0.74
18	42	42	0.52
19	62	9	0.62
20	85	40	0.43
21	48	56	0.64
22	53	44	0.57
23	83	74	0.73
24	70	26	0.32
25	97	24	0.86
26	38	49	0.61
27	56	65	0.56
28	65	40	0.83
29	61	81	0.84

Table 4.1 continues

30	87	35	0.58
31	25	69	0.53
32	52	35	0.64
33	60	28	0.68
34	50	39	0.86
35	99	46	0.57
36	25	15	0.39
37	32	35	0.51
38	28	55	0.43
39	30	49	0.69
40	53	7	0.52
41	74	56	0.47
42	43	8	0.87
43	74	46	0.64

Table 4.2 Substation location and capacities

Subs.	**X**	**Y**	$\boldsymbol{S_{exis}}$	**Type**
1	30	75	1	OA
2	40	86	1	OA
3	67	36	1	U
4	95	17	0	OA
5	17	74	0	OA
6	95	75	0	OA
7	75	24	0	U
8	26	33	0	ON
9	64	56	0	U
10	5	42	0	OA
11	25	91	0	OA
12	24	72	0	OA
13	8	3	0	OA
14	61	23	0	U

Table 4.2 continues

15	91	87	0	OA
16	52	22	0	ON
17	48	94	0	OA
18	33	82	0	OA
19	100	40	0	U
20	21	16	0	U
21	54	9	0	U
22	79	55	0	U
23	3	50	0	OA
24	83	14	0	OA
25	39	15	0	OA
26	46	41	0	U
27	49	64	0	ON
28	45	19	0	U
29	41	20	0	U
30	35	63	0	U
31	90	60	0	U
32	95	40	0	U

Table 4.3 The unit cost of substations and the land costs

Substation Type	$\boldsymbol{\varphi_S^f}$**($)**	$\boldsymbol{\varphi_S^v}$**($/pu)**	$\boldsymbol{S_{max}}$**(pu)**
Outdoor (away from zones)	115,000	17,000	6
Outdoor (near zones)	460,000	17,000	4
Underground	230,000	34,000	1.5

4.2.1 Effect of The Geographical Configurations

In this section, a comparison between the effect of different geographical configurations on the SEP problem is presented. In the first part, the zones are considered as the rectangular area along the installed feeder. In this context, the voltage drop and power loss in the feeder are integrated into the optimization problem. The summary of the investment decisions is listed in Table 4.4 and the resulting network is illustrated in Figure 4.5. It can be seen from the table that the existing substations 1, 2, and 3 are expanded 1.9, 1.5, and 0.5 pu along with the newly installed candidate substations. There are six outdoor substations away from the zones, one outdoor substation near zones, and five underground substations one of which is an existing substation in the network.

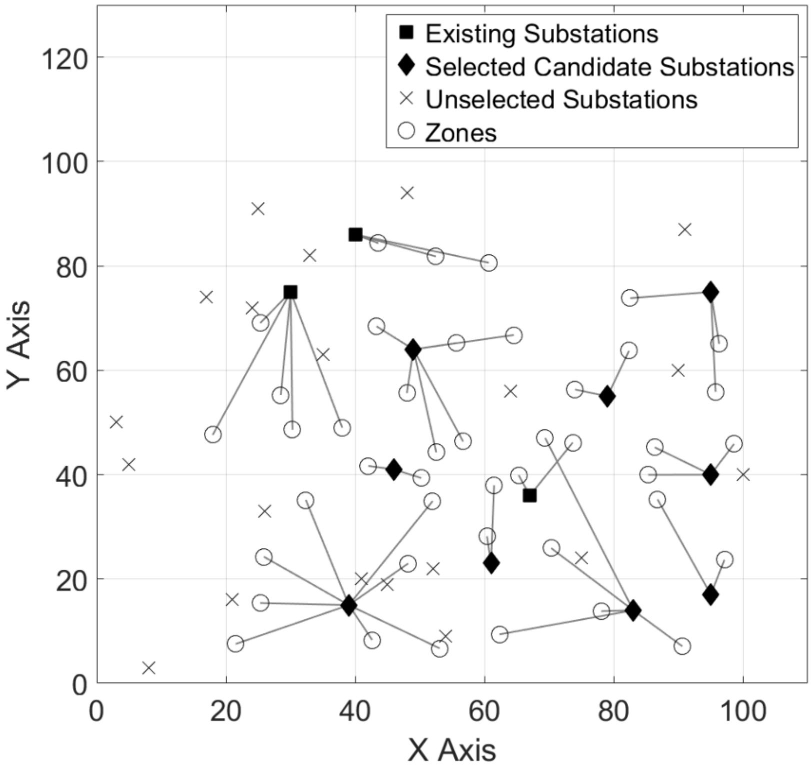

Figure 4.5 The resulting network for rectangular zones

Table 4.4 Investment decisions with rectangular zones

Substation	Connected Zones	Optimal Capacity
1	6, 26, 31, 38, 39	2.9
2	4, 15, 29	2.5
3	28, 43	1.5
4	25, 30	1.4
6	2, 5, 23	2.2

Table 4.4 continues

14	14, 33	1.4
22	7, 41	1.5
24	13, 16, 17, 19, 24	2.6
25	3, 8, 11 ,32, 36, 37 ,40, 42	5.1
26	18, 34	1.4
27	1, 9, 10, 21, 22, 27	4
32	12, 20, 35	1.5

In the second part, the zones are considered triangular areas. The summary of the investment decisions is listed in Table 4.5 and the resulting network is given in Figure 4.6. It can be seen from the table that the existing substations 1, 2, and 3 are expanded 0.4, 0.2 and 0.5 pu along with the newly installed candidate substations. There are nine outdoor substations away from the zones and six underground substations one of which is an existing substation in the network.

Table 4.5 Investment decisions with triangular zones

Substation	**Connected Zones**	**Optimal Capacity**
1	9, 21	1.4
2	4, 27	1.2
3	28, 43	1.5
4	25, 30	1.4
6	2, 5, 23	2.2
9	1, 10	1.4
12	6, 26, 31, 38, 39	2.9
13	8	0.9
14	14, 33	1.4
17	15, 29	1.8
22	7, 41	1.5
24	13, 16, 17, 19, 24	2.6
25	3, 11, 18, 32, 36, 37, 40, 42	4.7
26	22, 34	1.4
32	12, 20, 35	1.5

The summary of the costs is given in Table 4.6. It can be seen that the costs of substations are lower in the SEP problem with rectangular zones since 20.7 pu and 7.3 pu are fed by outdoor and underground substations respectively for rectangular zones while 18.2 pu and 8.7 pu are fed by outdoor and underground substations respectively for triangular zones. It is observed that the total land cost is higher in the SEP problem with the triangular zones.

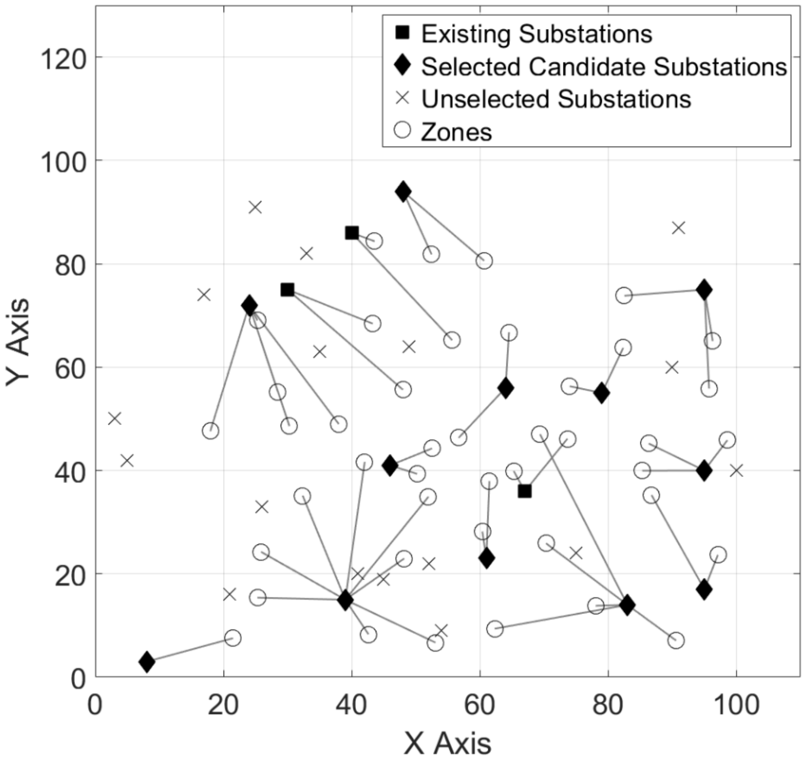

Figure 4.6 The resulting network for triangular zones

Likewise, the cost of the feeders is higher when the zones have triangular load distribution. The reason for this difference is that with triangular zones substation 27 is not installed as in for rectangular zones resulting in the installation of substation 9. Since substation 9 is an underground substation, its maximum capacity is low. Due to that, the optimization is forced to install several outdoor substations which necessitates the installation of longer feeders. Consequently, since there are longer feeders and more severe power loss for the triangular zones, the feeder losses are higher. Finally, the total cost is also higher in the case when the zones are considered to have triangular load distributions.

Table 4.6 Summary of the costs ($ thousands)

Zone	C_{sub}	C_{land}	C_{feeder}	C_{loss}	C_{total}
Rectangle	525.58	1840	601.37	4362.20	7329.10
Triangle	550.78	1955	622.15	4937.40	8065.30

4.2.2 Effect of The Voltage Drop Constraint and Feeder Loss Term

In this section, the effect of the absence of voltage constraints and power loss term on the SEP problem is demonstrated. The zones are assumed to have triangular load distributions for both simulations in this section. In the first part, the SEP problem without the voltage constraint is solved in order to observe its effect on the solution of the problem. The summary of the investment decisions is listed in Table 4.7 and the resulting network is given in Figure 4.7. It can be seen from the table that the existing substations 1, 2, and 3 are expanded 1.9, 1.5, and 0.5 pu along with the newly installed candidate substations. There are six outdoor substations away from the zones, one outdoor substation near zones, and four underground substations one of which is an existing substation in the network.

Table 4.7 Investment decisions without voltage drop constraints

Substation	Connected Zones	Optimal Capacity
1	6, 26, 31, 38, 39	2.9
2	4, 15, 29	2.5
3	28, 43	1.5
4	25, 30	1.4
6	2, 5, 7, 23, 41	3.6
14	14, 33	1.4
24	13, 16, 17, 19, 24	2.6
25	3, 8, 11, 32, 36, 37, 40, 42	5.1
26	18, 34	1.4
27	1, 9, 10, 21, 22, 27	4
32	12, 20, 35	1.5

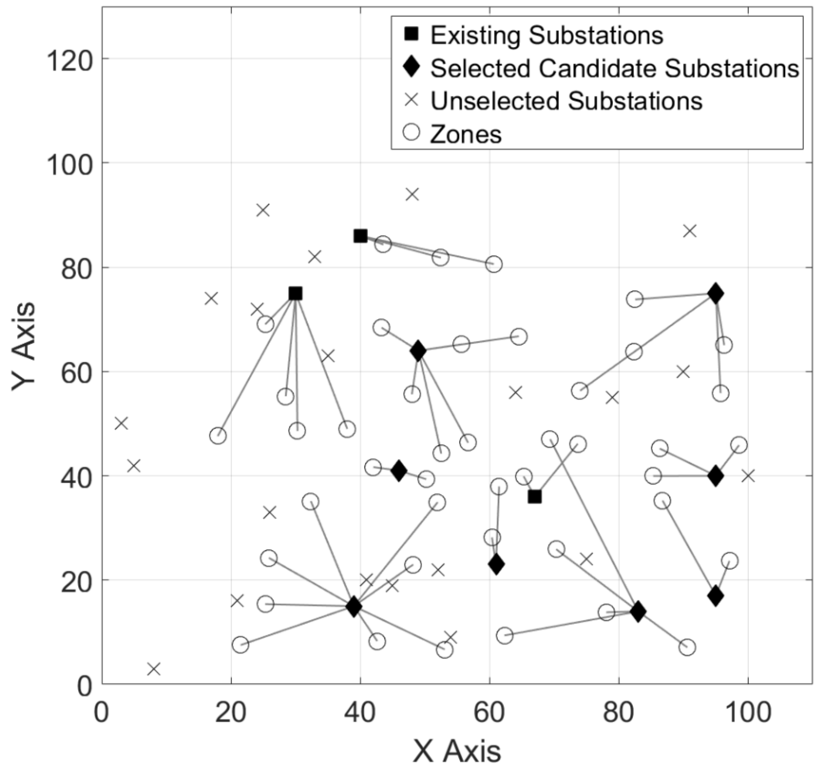

Figure 4.7 The resulting network without the voltage drop constraints

In the second part, the SEP problem without the cost of feeder loss is solved to investigate its effect on the SEP problem. The summary of the investment decisions is listed in Table 4.8 and the resulting network is given in Figure 4.8. It can be seen from the table that the existing substations 1, 2, and 3 are expanded 2.6, 0.8, and 0.5 pu along with the newly installed candidate substations. There are seven outdoor substations away from the zones and five underground substations one of which is an existing substation in the network.

Table 4.8 Investment decision without the cost of feeder loss

Substation	**Connected Zones**	**Optimal Capacity**
1	6, 9, 21, 26, 31, 38	3.6
2	4, 10, 27	1.8
3	34, 43	1.5
6	2, 5, 12, 23, 35	3.2

Table 4.8 continues

7	24, 28	1.2
9	1, 13	1.2
13	8	0.9
17	15, 29	1.8
22	7, 41	1.5
24	16, 17, 19, 20, 25, 30	3.7
25	3, 11, 18, 22, 32, 33, 36, 37, 40, 42	6
26	14, 39	1.4

In the third part, the SEP problem without both the voltage constraint and the cost of feeder loss is solved. The summary of the investment decisions is listed in Table 4.9 and the resulting network is given in Figure 4.9. It can be seen from the table that the existing substations 1, 2 and 3 are expanded 4.5, 3.4 and 0.4 pu along with the newly installed candidate substations. There are five outdoor substations away from the zones and one underground substation which is an existing substation in the network. It is clear that all the substations are installed near their maximum capacity due to the absence of voltage constraints and power loss term.

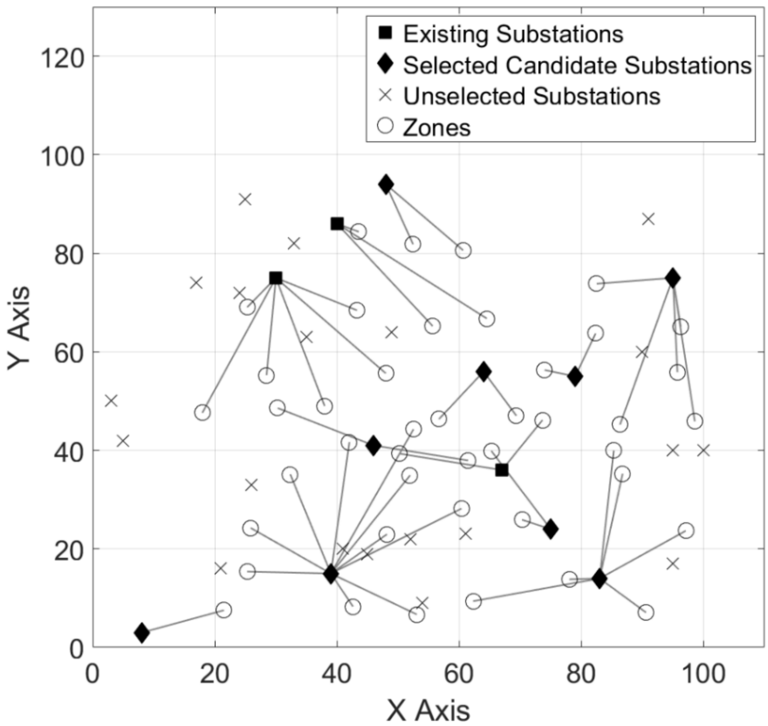

Figure 4.8 The resulting network without the cost of feeder loss

Table 4.9 Investment decisions without voltage drop constraints and cost of feeder loss

Substation	**Connected Zones**	**Optimal Capacity**
1	1, 6, 18, 21, 22, 26, 31, 38, 39	5.5
2	4, 9, 10, 15, 27, 29	4.4
3	13, 14, 24	1.4
6	2, 5, 7, 12, 23 ,35, 41	4.7
24	16, 17, 19, 20, 25, 28, 30, 33, 43	5.9
25	3, 8, 11, 32, 34, 36, 37 ,40, 42	5.9

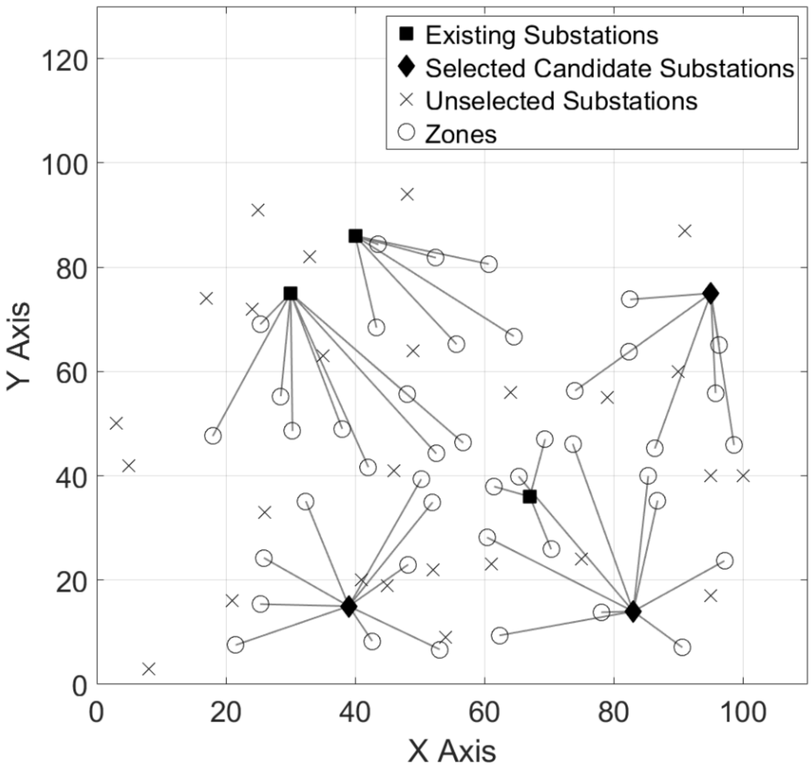

Figure 4.9 The resulting network without the voltage drop constraints and the cost of feeder loss

The summary of the costs is given in Table 4.10. It can clearly be seen that the highest cost of substations belongs to the result without feeder loss term. Likewise, the highest total land cost belongs to the result of the SEP problem without voltage drop constraint. Moreover, the highest feeder cost is obtained in the SEP problem without both the voltage constraint and the feeder loss, as expected which indicates longer feeders since there is no voltage drop limitation and feeder losses are neglected. In that point, longer feeders cause the installation of the minimum number of substations. Because the longest feeders are installed in the SEP problem without both the voltage constraint and the feeder loss term, feeder loss is at its highest value among other results. Finally, the total cost is the lowest in the SEP problem without the voltage drops which indicates that voltage drop is more dominant over that of the feeder loss in the increase of the total investment cost.

Table 4.10 Summary of the costs ($ thousands)

	C_{sub}	C_{land}	C_{feeder}	C_{loss}	C_{total}
without voltage drop	500.57	1610	631.91	5071.90	7814.40
without feeder loss	517.84	1495	754.15	5814.30	8581.30
without both	426.80	345	887.92	6800.40	8460.10

4.3 Conclusion

Modeling a convenient input to represent the demand along with the topology of the zones remains to be a challenge in power system planning studies. Results show that the most expensive overall cost belongs to the case when the zones have triangular load distribution. The case when the zones with rectangular load distribution are less expensive in terms of each cost term. Additionally, the effect of the voltage constraint and the cost of feeder loss is also investigated. Results indicate that the voltage drop is more dominant over that of the feeder loss in the increase of the total investment cost. Moreover, in the SEP problem without both the voltage constraint and the cost of feeder loss, the length of feeders increased which results in a few numbers of installed substations in the network. Consequently, it is observed that if the geographical configuration along the feeders is known or pre-planned, the investment decisions change accordingly.

CHAPTER FIVE
DISCUSSION

The stochastic nature of the players in the power system encumbers the ability of planners in decision-making drastically, resulting in the need for seeking out a wide range of alternative planning strategies. At that point, the progress in the research on this subject is encouraging. In order to contribute to the knowledge in the ongoing literature, I have introduced a methodology to systematically integrate the inherent characteristics of PV plants into the stochastic dynamic SEP from the risk management perspective in Chapter 2. This point of view enables the ability to quantify the risk while optimizing the range of possible investment costs, simultaneously. In the presence of a number of possible input scenarios, the risk is defined as the range of the possible investment costs such that high-risk corresponds to a wide range of investment costs and vice-versa. It is observed that the possible investment costs are scattered over a wide range with low frequencies in the risk-neutral strategy. On the contrary, the possible investment costs are observed to cover a narrow range with high frequencies in the risk-averse strategy. What outcome is revealed by taking this study into account is that suitable service area modeling in SEP leads to avoiding the poor usage of risk management strategies and consequently inadequate investment decisions.

PV power generation had a massive growth worldwide over the past years and it is expected to continue growing in the future. The main advantage of PV power generation comes from its unlimited and universally available source, the Sun; however, this advantage often creates a burden on the PV plants such as depending too much on environmental and weather conditions. In addition, PV plants are built by connecting so many interconnected units together to make them operate in harmony with each other which results in a high probability of faults in the PV plants. A number of researchers are currently searching for ways to shorten the period of most of the abnormal operation conditions by developing detection algorithms; however, this is still an ongoing research area and there is no denying that these conditions hold a significant portion of the operation duration (Klise et al., 2018; Köntges et al., 2014;

Yurtseven et al., 2021). As power system planners, instead of turning our back on this widely accepted phenomenon, we need to figure out how to accurately integrate this stochastic nature of the PV plant operation into planning problems. This thesis is written with the motivation of paving the way for incorporating the distinctive characteristics of PV plants with the power system planning studies. It goes without saying that the proposed hybrid model has its limitations and open to development for future research.

Substation expansion planning problem under uncertainties is handled by using either mathematical or meta-heuristic optimization. In mathematical optimization, there are several ways to model the optimization problem such as linear programming, nonlinear programming, dynamic programming, mixed-integer programming, etc. Mathematical optimization is known for its fast convergence (Nemati, Braun, & Tenbohlen, 2018). In that point, in order to conduct a risk management study, using CVaR as the risk measure is more than convenient due to its availability to have a straightforward linear formulation. The proposed model is a pioneered step in the literature to integrate inherent characteristics including abnormal events in PV plants into the planning studies through a systematical scenario tree generation. Moreover, this thesis also combines the prosumer-based risk-constrained stochastic dynamic SEP problem and conventional MILP framework under one roof in order to facilitate simple implementation. It is worthy to note that since real-world systems often tend to be complex problems, practical applications necessitate establishing a trade-off between modeling accuracy and computational complexity (Castillo, Conejo, Pedregal, Garciá, & Alguacil, 2001).

In the SEP problem, several authors conducted studies to obtain the general layout of the optimally expanded network without considering the geographical or electrical constraints (Gonzalez-Sotres, Mateo Domingo, Sanchez-Miralles, & Alvar Miro, 2013). Likewise, the voltage drop constraints are not included in (Bosisio et al., 2021; Alessandro Bosisio, Berizzi, Amaldi, Bovo, & Andy Sun, 2020). In the study of Alessandro Bosisio et al. (2020), due to the establishment of the trade-off between computational complexity and model accuracy, a mixed-integer programming approach is used to solve a practical problem that occurred because of the necessity of optimally integrating two different networks in Milano, Italy. In another study of Bosisio et al. (2021), a geographic information system based expansion planning problem is solved with a MILP framework including a risk index, which depends on the length of the feeder, based on regulations to keep the model linear. Moreover, a MILP-based SEP model is constructed in the study of Seifi & Sepasian (2011) such that the voltage drop along the feeders is constrained by simply limiting the distance between the service areas and the substations. In a similar manner, the distance from the substations to the service areas is limited according to Bolivia's regulations in order to avoid high voltage drops in the substation allocation problem in the study of Corigliano, Rosato, Ortiz Dominguez, & Merlo (2021). In that manner, it is adequate to use the supply radius of substations point-of-view in order to simply prevent the unrealistic solutions in terms of technical constraints as in (Bosisio et al., 2021; Alessandro Bosisio et al., 2020; Corigliano et al., 2021; Ge, Wang, Lu, & Liu, 2015; Gonzalez-Sotres et al., 2013; Lu, Wang, Ge, & Wang, 2014; Seifi & Sepasian, 2011). Similarly, regarding the (3.6), voltage profile constraints are formulated and implemented to the problem without conducting power flow analysis. The voltage profile along a feeder can be constrained by simply constraining its length under the MILP framework. This is done with the aim of encouraging practical and easy-to-use problem solving and not be burdened by unnecessary calculations as a first choice. That being said, it is obvious that in order to obtain the more detailed and tailor-made solutions, this point of view leaves its place to more complex calculations.

Integrating the systematic generation of suitable scenario sets that reflect the interrelated behaviors of the load and PV generation into the planning problem is still a challenge. In spite of the fact that the modern power system planning studies are based on the modeling of the stochastic behavior of the load and the PV generation, the development of the model of a suitable scenario set fell behind the growth of PV generation in the world. The proposed MILP-based method is a promising step towards addressing this gap by introducing a hybrid model that enables the ability for the inherent characteristics of PV plants to be integrated into the stochastic dynamic SEP through systematic probabilistic scenario tree generation. Needless to say, power system planning is still an ongoing research area and it is obvious that further work is necessary in terms of the decision-making in long-term planning under uncertainties.

CHAPTER SIX
CONCLUSION

This thesis focuses on two main parts. The first one is the siting, sizing, and timing of sub-transmission substations taking into account the inherent characteristics of PV plants under a risk-based dynamic stochastic mixed-integer linear programming framework. The second one is the allocation of distribution substations with the integration of triangular and rectangular configurations into the SEP optimization problem under the mixed-integer linear programming framework . In that manner, the general overview of this thesis and concluding remarks are presented as follows.

In Chapter 2, stochastic service area modeling and systematic scenario tree generation are presented. Firstly, the reason why the inherent characteristics including abnormal conditions of PV plants should be integrated into the stochastic dynamic SEP problem is discussed. At that point, three different modeling states are introduced to adequately integrate the inherent characteristics of PV and the load into the SEP problem. The first modeling state is formed by using Copula functions to describe the correlated nature between the PV generation and the load. The second modeling state considers the PV generation and the load as independent random variables and models scenarios when PV plants work under abnormal conditions. The third modeling state describes the nighttime characteristics of the load. Afterward, a method is proposed to create a suitable set of probabilistic scenarios and systematically generate the scenario tree of the stochastic dynamic SEP problem under the MILP framework.

In Chapter 3, the problem formulation of the proposed stochastic dynamic SEP problem under the MILP framework and case studies are presented. It is worthy to note that the objective function is different from the total cost function in scenario-based stochastic optimization since the probability of each scenario has to be included in the optimization problem. Hence, the weighted sum of all the cost functions is used as the objective function in the problem. At that point, the risk level which the decision-maker is willing to take becomes an important part. By using a risk weighting parameter, the risk level is quantified in the stochastic dynamic SEP with the CVaR risk measure. The CVaR risk measure is integrated into the MILP framework to manage the risk of uncertainties introduced in the service areas. Moreover, two case studies are conducted in this chapter. Case 1 is conducted in order to investigate the effect of the risk management strategies on investment decisions at different points in time. The results showed that the risk level corresponds to the range of possible cost scenarios that are used by the decision-makers while arranging the budget of the planning process. In this context, the higher risk level results in a wider range of possible cost scenarios, whereas the lower risk level results in a narrower range. In addition, when the decision-maker chooses the risk-averse strategy, the over investments in the early stages of the planning process are encountered as expected. Furthermore, Case 2 is conducted in order to reveal the reason why inherent characteristics including abnormal conditions of PV plants should be integrated into the stochastic dynamic SEP problem. Results show that forming the planning problem as if the PV plants always operate smoothly results in misleading outcomes. The presence of the inherent characteristics including abnormal conditions of PV plants in the stochastic dynamic SEP problem prevents the possible under investments in the planning process which can eventually lead to the inadequacy of the power system.

In Chapter 4, the distribution substation expansion planning problem is merged with the different geographical configurations of the zones within a mixed-integer linear programming framework. The voltage profile and power losses are integrated into the SEP problem by using the approximate lumped model of the zones. The effect of voltage constraints and the cost of feeder loss on the SEP problem is investigated and the comparison between the SEP problem with rectangular zones and with triangular zones is performed to show the impact of different geographical configurations on the investment decisions.

REFERENCES

Abedi, M. H., Hosseini, H., & Jalilvand, A. (2019). Sub-transmission substation expansion planning considering load center uncertainties of size and location. *International Journal of Electrical Power & Energy Systems*, *109*, 413–422.

Abedi, M. H., Hosseini, H., & Jalilvand, A. (2020). Sub-transmission substation expansion planning (SSEP) considering load uncertainties using a new clustering method. *International Transactions on Electrical Energy Systems*, *30 (7),* 12421.

Anukoolthamchote, P. C., Assané, D., & Konan, D. E. (2020). Net electricity load profiles: shape and variability considering customer-mix at transformers on the island of Oahu, Hawai'i. *Energy Policy*, *147*, 111732.

Arias, N. B., Tabares, A., Franco, J. F., Lavorato, M., & Romero, R. (2018). Robust joint expansion planning of electrical distribution systems and EV charging stations. *IEEE Transactions on Sustainable Energy*, *9*, 884–894.

Asensio, M., & Contreras, J. (2016). Stochastic unit commitment in isolated systems with renewable penetration under CVaR assessment. *IEEE Transactions on Smart Grid*, *7*, 1356–1367.

Atwa, Y. M., & El-Saadany, E. F. (2011). Probabilistic approach for optimal allocation of wind-based distributed generation in distribution systems. *IET Renewable Power Generation*, *5*, 79.

Avilés, J. P., Mayo-Maldonado, J. C., & Micheloud, O. (2020). A multi-objective evolutionary approach for planning and optimal condition restoration of secondary distribution networks. *Applied Soft Computing*, *90*, 106182.

Awad, A. S. A., EL-Fouly, T. H. M., & Salama, M. M. A. (2017). Optimal ESS allocation for benefit maximization in distribution networks. *IEEE Transactions on Smart Grid*, *8*, 1668–1678.

Behal, G. (2019). Substation innovations: the heart and brain of the grid. *IEEE Power and Energy Magazine*, *17*, 16–22.

Bogdan, A. V., Bogdan, V. A., Garkavyi, K. A., & Netrebko, D. S. (2021). The influence of load distribution along a line on power and place of connection of a photovoltaic system. *IOP Conference Series: Materials Science and Engineering*, *1019*, 012011.

Bosisio, A., Berizzi, A., Amaldi, E., Bovo, C., Morotti, A., Greco, B., & Iannarelli, G. (2021). A GIS-based approach for high-level distribution networks expansion planning in normal and contingency operation considering reliability. *Electric Power Systems Research*, *190*, 106684.

Bosisio, Alessandro, Berizzi, A., Amaldi, E., Bovo, C., & Andy Sun, X. (2020). Optimal feeder routing in urban distribution networks planning with layout constraints and losses. *Journal of Modern Power Systems and Clean Energy*, *8*, 1005–1013.

Cai, D., Shi, D., & Chen, J. (2014). Probabilistic load flow computation using Copula and Latin hypercube sampling. *IET Generation, Transmission & Distribution*, *8*, 1539–1549.

Castillo, E., Conejo, A. J., Pedregal, P., Garciá, R., & Alguacil, N. (2001). B*uilding and solving mathematical programming models in engineering and science*. New Jersey: John Wiley & Sons, Inc.

Chapaloglou, S., Nesiadis, A., Iliadis, P., Atsonios, K., Nikolopoulos, N., Grammelis, P., Kakaras, E. (2019). Smart energy management algorithm for load smoothing and peak shaving based on load forecasting of an island's power system. *Applied Energy*, *238*, 627–642.

Climatic Atlas of Clouds Over Land and Ocean. (n.d.). Retrieved February 17, 2021, from https://atmos.uw.edu/CloudMap/

Conejo, A. J., Baringo Morales, L., Kazempour, S. J., & Siddiqui, A. S. (2016). I*nvestment in electricity generation and transmission*. Switzerland: Springer International Publishing.

Conejo, A. J., Carrión, M., & Morales, J. M. (2010). D*ecision making under uncertainty in electricity markets*. Boston, MA: Springer US.

Corigliano, S., Rosato, F., Ortiz Dominguez, C., & Merlo, M. (2021). Clustering techniques for secondary substations siting. *Energies*, *14*, 1028.

Dadkhah, M., & Venkatesh, B. (2012). Cumulant based stochastic reactive power planning method for distribution systems with wind generators. *IEEE Transactions on Power Systems*, *27*, 2351–2359.

Dukpa, A., Venkatesh, B., & Chang, L. (2011). Fuzzy stochastic programming method: capacitor planning in distribution systems with wind generators. *IEEE Transactions on Power Systems*, *26*, 1971–1979.

Engerer, N. A., & Mills, F. P. (2014). KPV: A clear-sky index for photovoltaics. *Solar Energy*, *105*, 679–693.

Esmaeeli, M., Kazemi, A., Shayanfar, H., Haghifam, M.-R., & Siano, P. (2016). Risk-based planning of distribution substation considering technical and economic uncertainties. *Electric Power Systems Research*, *135*, 18–26.

European Commission. Joint Research Centre. (2019). *Distribution system operators observatory 2018: overview of the electricity distribution system in Europe.* Luxembourg: Publications Office of the European Union.

European Commission. Joint Research Centre. (2020). *Preparatory study for solar photovoltaic modules, inverters and systems: final report.* Luxembourg: Publications Office of the European Union.

Evangelopoulos, V. A., & Georgilakis, P. S. (2014). Optimal distributed generation placement under uncertainties based on point estimate method embedded genetic algorithm. *IET Generation, Transmission & Distribution*, *8*, 389–400.

Fan, V. H., Dong, Z., & Meng, K. (2020). Integrated distribution expansion planning considering stochastic renewable energy resources and electric vehicles. *Applied Energy*, *278*, 115720.

Franco, J. F., Rider, M. J., & Romero, R. (2016). Robust multi-stage substation expansion planning considering stochastic demand. *IEEE Transactions on Power Systems*, *31*, 2125–2134.

Ge, S., Wang, S., Lu, Z., & Liu, H. (2015). Substation planning method in an active distribution network under low-carbon economy. *Journal of Modern Power Systems and Clean Energy*, *3*, 468–474.

Geoffrey T. Klise, Roger R. Hill, Colin J. Hamman, Peter H. Kobos, Vipin Gupta, Benjamin B. Yang, & Nadav Enbar. (2014). *PV reliability operations and maintenance (PVROM) database initiative: 2014 progress report* (No. SAND2014-20612). Albuquerque, US: Sandia National Laboratories.

Gonen, T. (2014). *Electric power distribution engineering* (3rd edition). Florida, US: CRC Press.

Gonzalez-Sotres, L., Mateo Domingo, C., Sanchez-Miralles, A., & Alvar Miro, M. (2013). Large-scale MV/LV transformer substation planning considering network costs and flexible area decomposition. *IEEE Transactions on Power Delivery*, *28*, 2245–2253.

Goodrich, A., James, T., & Woodhouse, M. (2012). *Residential, commercial, and utility-scale photovoltaic (PV) system prices in the United States: current drivers and cost-reduction opportunities* (No. NREL/TP-6A20-53347, 1036048; p. NREL/TP-6A20-53347, 1036048). Colorado, US: National Renewable Energy Laboratory.

Haghifam, M. R., & Shahabi, M. (2002). Optimal location and sizing of HV/MV substations in uncertainty load environment using genetic algorithm. *Electric Power Systems Research*, *63*, 37–50.

Hemmati, R., Hooshmand, R.-A., & Taheri, N. (2015). Distribution network expansion planning and DG placement in the presence of uncertainties. *International Journal of Electrical Power & Energy Systems*, *73*, 665–673.

Hincapie I., R. A., Gallego R., R. A., & Mantovani, J. R. S. (2019). A decomposition approach for integrated planning of primary and secondary distribution networks considering distributed generation. *International Journal of Electrical Power & Energy Systems*, *106*, 146–157.

Hou, Q., Zhang, N., Du, E., Miao, M., Peng, F., & Kang, C. (2019). Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China. *Applied Energy*, *242*, 205–215.

Høyland, K., & Wallace, S. W. (2001). Generating scenario trees for multistage decision problems. *Management Science*, *47*, 295–307.

Jaeger-Waldau, A. (2019). *PV status report 2019* (EUR 29938 EN). Luxembourg: Publications Office of the European Union. JRC118058 (JRC118058).

Jalali, M., Zare, K., & Hagh, M. T. (2014). A multi-stage MINLP-based model for sub-transmission system expansion planning considering the placement of DG units. *International Journal of Electrical Power & Energy Systems*, *63*, 8–16.

Karimi, M., & Haghifam, M. R. (2017). Risk based multi-objective dynamic expansion planning of sub-transmission network in order to have eco-reliability, environmental friendly network with higher power quality. *IET Generation, Transmission & Distribution*, *11*, 261–271.

Kersting, W. H. (2017). *Distribution system modeling and analysis* (4th ed.). Florida, US: CRC Press.

Kiani Rad, H., & Moravej, Z. (2017). An approach for simultaneous distribution, sub-transmission, and transmission networks expansion planning. *International Journal of Electrical Power & Energy Systems*, *91*, 166–182.

Klise, G. T., Lavrova, O., & Gooding, R. L. (2018). *PV system component fault and failure compilation and analysis.* (No. SAND--2018-1743, 1424887; p. SAND--2018-1743, 1424887). Albuquerque, US: Sandia National Lab.

Köntges, M., Kurtz, S., Packard, C., Jahn, U., Berger, K., Kato, K., Friesen, G. (2014). *Review of failures of photovoltaic modules* (IEA PVPS Task 13 No. IEA-PVPS T13-01:2014). IEA-PVPS: International Energy Agency.

Laukamp, H. (2002). *Reliability study of grid connected PV systems* (No. IEA-PVPS T7-08: 2002). Freiburg, Germany: The International Energy Agency.

Li, P., Yu, D., Yang, M., & Wang, J. (2018). Flexible look-ahead dispatch realized by robust optimization considering CVaR of wind power. *IEEE Transactions on Power Systems*, *33*, 5330–5340.

Li, Y., Du, M., Xie, W., Yang, B., Fang, C., Zhang, Y., & Wang, S. (2018). Method for division of urban load power supply district based on cluster analysis. *IET Generation, Transmission & Distribution*, *12*, 4577–4581.

Lima, R. M., Conejo, A. J., Langodan, S., Hoteit, I., & Knio, O. M. (2018). Risk-averse formulations and methods for a virtual power plant. *Computers & Operations Research*, *96*, 350–373.

Liu, Z., Wen, F., & Ledwich, G. (2011). Optimal siting and sizing of distributed generators in distribution systems considering uncertainties. *IEEE Transactions on Power Delivery*, *26*, 2541–2551.

Long, C. N., & Ackerman, T. P. (2000). Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. *Journal of Geophysical Research: Atmospheres*, *105*, 15609–15626.

Lu, Z., Wang, S., Ge, S., & Wang, C. (2014). Substation planning method based on the weighted Voronoi diagram using an intelligent optimisation algorithm. *IET Generation, Transmission & Distribution*, *8*, 2173–2182.

Mazhari, S. M., Monsef, H., & Romero, R. (2015). A hybrid heuristic and evolutionary algorithm for distribution substation planning. *IEEE Systems Journal*, *9*, 1396–1408.

Mazhari, Seyed Mahdi, & Monsef, H. (2013). Dynamic sub-transmission substation expansion planning using learning automata. *Electric Power Systems Research*, *96*, 255–266.

Mehta, V. K., & Mehta, R. (2005). *Principles of power system*. India: Chand (S.) & Co Ltd.

Melgar-Dominguez, O. D., Pourakbari-Kasmaei, M., & Mantovani, J. R. S. (2019). Adaptive robust short-term planning of electrical distribution systems considering siting and sizing of renewable energy based DG units. *IEEE Transactions on Sustainable Energy*, *10*, 158–169.

Nemati, M., Braun, M., & Tenbohlen, S. (2018). Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming. *Applied Energy*, *210*, 944–963.

Othman, M. M., Abdelaziz, A. Y., Hegazi, Y. G., & El-Khattam, W. (2015). Approach for modelling stochastically dependent renewable energy-based generators using diagonal band copula. *IET Renewable Power Generation*, *9*, 809–820.

Palanichamy, C., & Babu, N. S. (2002). Day-night weather-based economic power dispatch. *IEEE Transactions on Power Systems*, *17*, 469–475.

Papaefthymiou, G., & Kurowicka, D. (2009). Using copulas for modeling stochastic dependence in power system uncertainty analysis. *IEEE Transactions on Power Systems*, *24*, 40–49.

Park, H., & Baldick, R. (2020). Optimal capacity planning of generation system integrating uncertain solar and wind energy with seasonal variability. *Electric Power Systems Research*, *180*, 106072.

Reno, M. J., & Hansen, C. W. (2016). Identification of periods of clear sky irradiance in time series of GHI measurements. *Renewable Energy*, *90*, 520–531.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, *26*, 1443–1471.

Saber, A. Y., & Venayagamoorthy, G. K. (2012). Resource scheduling under uncertainty in a smart grid with renewables and plug-in vehicles. *IEEE Systems Journal*, *6*, 103–109.

Salyani, P., Salehi, J., & Samadi Gazijahani, F. (2018). Chance constrained simultaneous optimization of substations, feeders, renewable and non-renewable distributed generations in distribution network. *Electric Power Systems Research*, *158*, 56–69.

Seifi, H., & Sepasian, M. S. (2011). *Electric power system planning*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sepasian, M. S., Seifi, H., Foroud, A. A., Hosseini, S. H., & Kabir, E. M. (2006). A new approach for substation expansion planning. *IEEE Transactions on Power Systems*, *21*, 997–1004.

Silvestre, S., Chouder, A., & Karatepe, E. (2013). Automatic fault detection in grid connected PV systems. *Solar Energy*, *94*, 119–127.

Silvestre, S., Kichou, S., Chouder, A., Nofuentes, G., & Karatepe, E. (2015). Analysis of current and voltage indicators in grid connected PV (photovoltaic) systems working in faulty and partial shading conditions. *Energy*, *86*, 42–50.

Sklar, M. (1959). Fonctions de repartition a n dimensions et leurs marges. *Paris Institute of Statistics*, *8*, 229–231.

Soroudi, A., & Afrasiab, M. (2012). Binary PSO-based dynamic multi-objective model for distributed generation planning under uncertainty. *IET Renewable Power Generation*, *6*, 67.

Temraz, H. K., & Salama, M. M. A. (2002). A planning model for siting, sizing and timing of distribution substations and defining the associated service area. *Electric Power Systems Research*, *62*, 145–151.

Ugranlı, F. (2020). Probabilistic distribution planning: including the interactions between chance constraints and renewable energy. *Sustainable Energy, Grids and Networks*, *23*, 100372.

Wang, C., & Nehrir, M. H. (2004). Analytical approaches for optimal placement of distributed generation sources in power systems. *IEEE Transactions on Power Systems*, *19*, 2068–2076.

Warne, D. F. (2005). *Newnes electrical power engineer's handbook*. Oxford: Elsevier.

Xue Lin, Yanzhi Wang, Pedram, M., Jaemin Kim, & Naehyuck Chang. (2014). Designing fault-tolerant photovoltaic systems. *IEEE Design & Test*, *31*, 76–84.

Youcef Ettoumi, F., Mefti, A., Adane, A., & Bouroubi, M. Y. (2002). Statistical analysis of solar measurements in Algeria using beta distributions. *Renewable Energy*, *26*, 47–67.

Yurtseven, K., & Karatepe, E. (2021). Distribution substation expansion planning considering different geographical configurations. *2021 12th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, 1–6. Bucharest, Romania: IEEE.

Yurtseven, K., Karatepe, E., & Deniz, E. (2021). Sensorless fault detection method for photovoltaic systems through mapping the inherent characteristics of PV plant site: Simple and practical. *Solar Energy*, *216*, 96–110.

Zare, A., Chung, C. Y., Zhan, J., & Faried, S. O. (2018). A distributionally robust chance-constrained MILP model for multistage distribution system planning with uncertain renewables and loads. *IEEE Transactions on Power Systems*, *33*, 5248–5262.

APPENDICES

NOMENCLATURE

Ω_{t}	Set of time stages
ω	Set of scenarios
i	Index of service areas
j	Index of substations
Nl	Number of service areas
Ns	Number of the candidate and existing substations
$S_L^{\omega,t}(i)$	Load magnitude of the ith service area at stage t under scenario ω (MW)
n	Number of injected power clusters
m	Number of growth scenarios
S_{cap}	Expandable transformer unit rated power (MW)
$S_{exis}(j)$	The existing capacity of jth substation (MW)
$D(i,j)$	The distance between the ith service area and the jth substation (km)
φ_L	The cost of the feeder's unit length per unit power transfer capability ($/km·MW)
φ_S^f	Fixed cost (land cost) of a substation ($)
φ_S^v	The variable cost of a substation ($/MW)
β	Risk weighting parameter
α	Confidence level
ρ_ω	Probability of the scenario ω
C_{line}	The variable cost of lines ($)
C_{fix_sub}	Fixed cost of installed substations ($)

C_{var_sub}	The variable cost of installed substations ($)
ξ	The auxiliary continuous variable that computes the Value-at-Risk ($)
η_{ω}	Auxiliary continuous variable equal to the summation of the VaR and the system cost under scenario ω ($)
$X^{\omega,t}(i,j)$	The binary variable of feeder connections.
$X_S^{\omega,t}(j)$	The binary variable of presence of substations.
$X_{exp}^{\omega,t}(j)$	Integer variable of expansion number.
$\mathcal{S}_G$	Set of growth scenarios
$\mathcal{S}_{DCS}$	Set of daytime injections with clear-sky and no-fault
$\mathcal{S}_{DAE}$	Set of daytime injections under abnormal events
$\mathcal{S}_{NT}$	Set of nighttime injections
$\mathcal{S}_{SD}$	Set of sun duration
$\mathcal{S}_{AC}$	Set of availability condition
$\mathcal{S}_{\omega}$	Set of all possible scenarios
ρ_G	Probability of the growth scenario set
ρ_{SD}	Probability of sun duration scenario set
ρ_{AC}	Probability of availability condition scenario set
ρ_{DC}	Probability of injections in daytime with clear-sky and no-fault cluster
ρ_{DA}	Probability of injections in daytime with abnormal events cluster
ρ_N	Probability of injections in nighttime cluster
ρ_{DCS}	Probability of injections in daytime with clear-sky and no-fault scenario set
ρ_{DAE}	Probability of injections in daytime with abnormal events scenario set
ρ_{NT}	Probability of injections in nighttime scenario set