

A RISK-AVERSE APPROACH FOR THE PLANNING OF A HYBRID RENEWABLE ENERGY SYSTEM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

By
Özlem Yılmaz
August 2017

A RISK-AVERSE APPROACH FOR THE PLANNING OF A
HYBRID RENEWABLE ENERGY SYSTEM

By Özlem Yılmaz

August 2017

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Master of Science.

Özlem Çavuş İyigün (Advisor)

Ayşe Selin Kocaman (Co-Advisor)

Emre Nadar

Melih Çelik

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

A RISK-AVERSE APPROACH FOR THE PLANNING OF A HYBRID RENEWABLE ENERGY SYSTEM

Özlem Yılmaz

M.S. in Industrial Engineering

Advisor: Özlem Çavuş İyigün

Co-Advisor: Ayşe Selin Kocaman

August 2017

We propose a risk-averse two-stage stochastic programming for a hybrid renewable energy system planning problem, where we model the risk-aversion using Conditional Value at Risk (CVaR). The aim of this study is to find the optimal capacities of the system components in a cost effective way while considering the risk-aversion of the decision maker. Renewable energy sources that are utilized in our hybrid system are solar and hydro, while the diesel fuel is used as a backup source. We assume that the water inflow to the reservoirs is uncertain, therefore, based on historical streamflow data for Mediterranean Region of Turkey, we generate scenarios for streamflow by using a modified k -nearest neighbor (k -NN) algorithm. We solve our model for different levels of risk-aversion and compare the optimal solutions. For models with large number of scenarios, we propose a multi-cut scenario-wise decomposition algorithm as an exact solution method. In order to evaluate the performance of our algorithm, we compare it with CPLEX. We conclude that, for a large number of scenarios, our algorithm is more efficient than CPLEX.

Keywords: Two-stage stochastic optimization, Conditional Value at Risk (CVaR), Scenario-wise decomposition, Scenario generation.

ÖZET

YENİLENEBİLİR HİBRİT ENERJİ SİSTEMİ PLANLAMASINA RİSKTEN KAÇINAN BİR YAKLAŞIM

Özlem Yılmaz

Endüstri Mühendisliği, Yüksek Lisans

Tez Danışmanı: Özlem Çavuş İyigün

Eş Tez Danışmanı: Ayşe Selin Kocaman

Ağustos 2017

Çalışmada yenilenebilir hibrit enerji sistemine riskten kaçınan iki aşamalı stokastik programlama modeli önerilmektedir. Riskten kaçınma, koşullu riske maruz değer ile modellenmiştir. Çalışmanın amacı, sistem bileşenlerinin optimum kapasitelerini en uygun maliyette riskten kaçınan tutum gözönüne alınarak belirlemektir. Ele aldığımız hibrit sistemdeki yenilenebilir enerji kaynakları güneş enerjisi ve hidroelektrik enerjisi iken dizel yakıt ise yedek kaynak olarak kullanılmaktadır. Modeldeki belirsizliğin rezervuarlara akan su miktarından kaynaklandığı ele alınmakta olup Türkiye'deki Akdeniz Bölgesi'ne ait su akış verileri baz alınarak modifiye edilmiş k en yakın komşu algoritması ile senaryolar üretilmektedir. Önerilen model farklı riskten kaçınma tutumları için çözülmekte ve elde edilen optimum çözümler kıyaslanmaktadır. Çok sayıda senaryoya sahip modeller için çoklu-kesitli senaryo tabanlı ayrıştırma algoritması kesin bir çözüm yöntemi olarak sunulmaktadır. Önerilen algoritmanın performansını değerlendirmek için algoritma CPLEX ile kıyaslanmaktadır. Önerilen algoritmanın CPLEX ile kıyaslandığında daha verimli olduğu sonucuna varılmıştır.

Anahtar sözcükler: İki aşamalı stokastik eniyileme, Koşullu Riske Maruz Değer, Senaryo tabanlı ayrıştırma, Senaryo üretme.



To my sister...

Acknowledgement

I wish to express my gratitude to Asst. Prof. Özlem Çavuş İyigün and Asst. Prof. Ayşe Selin Kocaman for their endless support, guidance and encouragement throughout my graduate study.

I sincerely thank to Asst. Prof. Emre Nadar and Asst. Prof. Melih Çelik for accepting to read and review my thesis, and their invaluable comments and suggestions.

I also would like to thank my family for their love and support. I would like to express my deepest gratitude to my sister Dr. Gülçin Yılmaz who is always being there for me.

Contents

1	Introduction	1
2	Literature Review	5
2.1	Energy Systems Planning	5
2.2	Risk-Averse Two-Stage Stochastic Optimization Models and Solution Methods	8
3	Problem Definition and Formulation	12
3.1	Problem Definition	12
3.2	Conditional Value at Risk (CVaR)	14
3.3	Problem Formulation	15
4	Solution Methodology	22
5	Computational Study	28
5.1	Scenario Generation	28

5.2	Data	30
5.3	Numerical Results	33
6	Conclusion	43
A	Numerical Results for Optimal First-Stage Decisions for the number of Scenarios 600, 700, 800, 900 and 1000	49
B	Numerical Results for the Percent Deviations from Risk Neutral Case for the number of scenarios 600, 700, 800, 900 and 1000	55
C	Numerical Results with Real Demand Data	61

List of Figures

5.1	Hourly Solar Radiation Data	31
5.2	Hourly Demand Data	31
5.3	Daily Historical Streamflows of Manavgat River for Years 1997 - 2015	32
5.4	50 Scenarios for Hourly Streamflows Generated Using Modified k -NN	32
5.5	Optimal First Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenarios	36
5.6	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenarios	42
A.1	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case	50
A.2	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case	51
A.3	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case	52

A.4	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case	53
A.5	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case	54
B.1	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case	56
B.2	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case	57
B.3	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case	58
B.4	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case	59
B.5	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case	60
C.1	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenario Case	62
C.2	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case	63
C.3	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case	64
C.4	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case	65
C.5	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case	66

C.6	Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case	67
C.7	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenario Case	68
C.8	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case	69
C.9	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case	70
C.10	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case	71
C.11	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case	72
C.12	% Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case	73

List of Tables

5.1	Parameters adapted from [3]	30
5.2	GCV Values for Different Number of Neighbors (k)	34
5.3	Comparison of CPU times for the algorithm and extensive formulation for 500 scenarios	35
5.4	Comparison of Total Expected Costs (\$) for 500 and 1000 scenarios	38
5.5	Comparison of Total Expected Costs (\$) for 500 and 700 scenarios	39
5.6	Comparison of Total Expected Costs (\$) for 700 and 1000 scenarios	39
5.7	Numerical Results for $\alpha = 0.7$ and $\lambda \in \{0.3, 0.5, 0.7\}$	40
5.8	Numerical Results for $\alpha = 0.8$ and $\lambda \in \{0.3, 0.5, 0.7\}$	40
5.9	Numerical Results for $\alpha = 0.9$ and $\lambda \in \{0.3, 0.5, 0.7\}$	41
C.1	Comparison of CPU times for the algorithm and Extensive Formulation for 500 scenarios	61

Chapter 1

Introduction

World's energy consumption highly depends on fossil fuels. International Energy Agency (IEA) declares that, in 2014, 81.1% of World's total energy consumption is based on fossil fuels and the distribution of this percentage is 28.6% coal, 21.2% natural gas, and 31.3% oil [1]. In 1971, the percent of fossil fuels as a primary energy source in Gigatons of Oil Equivalent (GTOE) was 86% [2]. Even though there is a 4.9% decrease in this amount from 1971 to 2014, energy consumption still significantly depends on fossil fuels. The consumption of these sources, however, releases toxic gases, such as CO_2 , to the atmosphere and this gives rise to the air pollution along with the global warming. CO_2 emissions resulted from the fossil fuel usage have an increasing trend over the years. Amount of CO_2 emissions from fossil fuel usage in 1870 is close to 0 in Gigatons of CO_2 (GtCO_2) and, in 2014, it is over 32 GtCO_2 [2]. Besides their harmful effects to the environment, fossil fuels are finite resources and, therefore, there is a growing need to use alternative renewable sources to satisfy the demand for energy.

Renewable energy sources that can be used as clean alternatives to fossil fuels are solar power, wind power, hydropower, geothermal energy, biofuel and biomass, hydrokinetic energy based on wave, tidal, and motion of water. Most commonly used ones are solar power, hydropower, and wind power. These sources can also be utilized in compositions such as in forms of hybrid renewable energy systems.

The main issue that has to be considered while using renewables as an energy source is that they are intermittent. Their intermittent characteristics make it harder to plan and satisfy the demand. Therefore, additional storage mechanisms are integrated into the hybrid system to make sure to satisfy the demand. Some of these storage devices are batteries, diesel generators, and pumped hydro storage. However, the construction of such hybrid energy systems requires large amount of investments and, therefore, a careful planning is essential.

A decision maker that aims to construct a hybrid renewable energy system to satisfy the demand for energy with minimum cost needs to consider the intermittent nature of the renewables. These sources are not always available as they needed. For instance, the amount of solar radiation in a certain location changes depending on time of the day (whether it is daytime or night) and clearness of the sky during daytime. Similarly, wind speed in a specific location depends on the geographic characteristics and climatic conditions of that place. Therefore, when the renewable energy sources are available and how much they can be utilized is uncertain. In this context, renewable energy systems planning and sizing problems are generally formulated as stochastic optimization models. The stochasticity in these problems can be related to streamflows, wind speed, solar radiation, and demand, depending on the types of renewable sources and problem frame.

In classical stochastic optimization problems, the objective function is to optimize the expected value of a random outcome that can be considered as a cost or profit. This approach however, implicitly assumes that the decision maker is risk-neutral. For a hybrid renewable energy system planning problem, a risk-averse approach will be more realistic considering the large amount of investment requirements and the intermittent characteristics of the renewables. In this way, the relation between risk associated with the uncertainties and the key decisions to be made by the decision makers with different attitudes towards risk can be observed.

A two-stage stochastic linear programming model for a hybrid renewable energy system planning problem is proposed by Kocaman et al. [3]. The hybrid

system is composed of solar panels, hydropower stations, transmission lines, and diesel generators. The aim in [3] is to find the optimal sizes of the system components with minimum cost while satisfying the demand. The source of stochasticity involved in the problem is related to the amount of water inflow to the reservoirs.

In this thesis, we introduce a risk-averse approach for the planning of the hybrid renewable energy system proposed in [3]. In that study, the objective function is minimizing the expected total cost and, therefore, decision makers are considered as risk-neutral. A risk-neutral decision maker will make investments as long as the expected random outcome is minimum (or maximum in profit case) and the fluctuations in these random outcomes will be disregarded. In other words, the possibility of large amounts of losses will be accepted if there is also a possibility of large amounts of gains. However, this situation will not be preferred by risk-averse decision makers, who tend to avoid extreme losses in any cases. Therefore, this approach is inadequate to capture the decisions made under risk-aversion.

In our proposed model, we incorporate a coherent risk measure, Conditional Value at Risk (CVaR), in order to model the risk-aversion. For computational experiments, we present a case study for Mediterranean Region in Turkey considering 19 years historical streamflow data of Manavgat River. In order to generate additional scenarios for streamflow, we use a modified k -NN (k -nearest neighbor) algorithm developed by Prairie et al. [4]. As a solution method, we propose a scenario-wise decomposition algorithm that is similar to the L-shaped method proposed by Van Slayke and Wets [5].

The structure of this thesis is as follows: In Chapter 2, we provide a literature review on energy system planning problems, risk-averse approaches for two-stage stochastic problems, and the solution methods. In Chapter 3, we define our problem, give preliminaries for the risk measures, and present our risk-averse formulation. Then, in Chapter 4, we propose a scenario-wise decomposition algorithm as an exact solution methodology. In Chapter 5, we first represent a modified k -NN algorithm that we use to generate additional scenarios. Then, we give the data for the application of our proposed model and solution method, and we provide the numerical results. In the last chapter, Chapter 6, we summarize

our study and provide the possible future work.



Chapter 2

Literature Review

This chapter consists of two sections representing a literature review on the studies concerned with energy systems planning (Section 2.1) and risk-averse two stage stochastic programming models and their solution methods (Section 2.2).

2.1 Energy Systems Planning

In the literature, energy systems planning problems generally include hybrid energy systems that consist of renewable and backup sources. Solar, wind and hydro power are mostly considered renewable sources while batteries and diesel generators are used as backup.

Ashok [6] studies a hybrid renewable energy planning problem consisting of wind power, micro-hydropower, solar power, diesel generator, and batteries. Combinations of these are used as energy sources to satisfy the demand in rural areas. The aim is to obtain the optimal combination of component sizes with minimum cost and diesel usage. This problem is modeled using deterministic non-linear constrained optimization and Quasi-Newton Method, which is an exact solution method, is used to solve the model.

Ekren et al. [7] study a hybrid energy system sizing problem which aims to determine the optimal sizes of the system components while minimizing the total cost. In that study, solar and wind power are the main sources to satisfy the demand and batteries are considered as backup sources. Solar radiation, wind speed and the demand are considered as sources of the uncertainty and a simulation approach is used to find the probability distributions related to these inputs. Response Surface Methodology (RSM) is used to solve the sizing problem. A break-even distance analysis is conducted to compare the cost efficiency of proposed hybrid system and another alternative energy source that is extension of transmission grid lines. They conclude that after a certain distance, it is more cost effective to use hybrid system than extending the grid lines.

Ekren and Ekren [8], different than [7], use a simulated annealing algorithm as a heuristic solution approach to find an acceptable solution for hybrid energy sizing problem. The results of these two studies, [7] and [8], are compared in terms of solution performance and it is concluded that the simulated annealing algorithm is better than RSM.

Senjyu et al. [9] consider a hybrid energy system with wind turbines, solar panels, batteries, and diesel generators to satisfy the demand for energy in isolated islands of Japan. The aim is to find the optimal configuration with minimum total cost that includes operating and fixed costs. Genetic algorithm-based-simulation optimization method is used to solve the problem. Simulations of the optimization model for three islands are presented. The method used in that study is capable of reducing the operational costs by 10% approximately compared to the case for which the diesel generators are the only energy sources.

Kuznia et al. [10] study a hybrid energy system planning problem to satisfy the demand in isolated areas. The system consists of wind turbines, storage device, thermal generators and transmission lines. They model the problem as a two-stage stochastic mixed integer programming with the aim of finding the optimal configuration with the minimum expected total cost. The randomness in that study is originated from the wind speed and the demand. Because of the complex nature of the model, they propose an exact solution algorithm based on

Benders decomposition.

Compared to the ones explained above, our study is more similar to the following studies.

Sun et al. [11] consider a hybrid energy system consisting of wind farms and transmission network to satisfy the demand for energy. A classic stochastic programming model in the form of a facility location and supply chain model and two different risk-averse models involving linear and non-linear risk measures are introduced for the problem. Demand and wind speed are the sources of uncertainty involved in these models. CVaR and Higher Moment Coherent Risk (HMCR) measure are used in order to model risk-aversion. As an exact solution method, a variation of Benders decomposition algorithm based on branch-and-bound is used to solve the constructed linear and non-linear mixed integer models. The model with HMCR involves a p -order cone constraint and in order to deal with this, polyhedral approximation is used in the proposed decomposition method. The proposed algorithms are applied to the problems with different number of scenarios ranging from 100 to 2000. It is concluded that these algorithms are more efficient than the CPLEX MIP solver.

Merzifonluoğlu and Uzgören [12] propose an energy system sizing problem that involves photovoltaic components, storage devices (batteries), and the grid lines in a campus area. They first propose a two-stage stochastic linear programming formulation for this problem and the aim is to find the optimum sizes of the components with minimum expected total cost while satisfying the demand. Stochasticity involved in the problem comes from the demand, solar radiation and the system performance. Then a risk-averse model is formulated involving CVaR as the measure of risk in the objective. These models are converted into mixed integer programming problems after including a payback period as a constraint. They propose a sample average approximation (SAA) method that is based on approximating the expected value of the objective function by using sample average estimates. They solve the problem through heuristic and exact solution methods. A numerical study is presented in order to evaluate the performance of the proposed methods. Scenarios for the demand data is created by

using Winters model while probability distribution of clearness indices is used to generate scenarios for solar radiation data. It is concluded that the proposed solution method is more efficient than CPLEX.

Kocaman et al. [3] consider a hybrid renewable energy system consisting of solar panels, hydropower stations, transmission lines and diesel generators. Primary sources of energy are solar power and hydropower while the diesel fuel is used as backup source. The aim is to determine the optimal sizes of the system components while satisfying the demand for energy with the minimum cost. Proposed model in that study is a two-stage stochastic programming model and the source of uncertainty is the amount of water inflow to the reservoirs. Proposed model is used in a case study for India.

In our study, we adopt the model proposed by Kocaman et al. [3] and develop a risk-averse two-stage stochastic programming model using CVaR as the measure of risk. The aim of our model different than [3] is to minimize the risk adjusted cost while satisfying the demand for energy. Different than the studies by Sun et al. [11] and Merzifonoğlu and Uzgören [12], the hybrid system that we consider consists of the hydropower stations, solar panels, transmission lines and the diesel generators. In contrast to [11] and [12], we present a risk-averse two-stage stochastic programming model and propose a multi-cut scenario-wise decomposition algorithm as an exact solution method.

2.2 Risk-Averse Two-Stage Stochastic Optimization Models and Solution Methods

In two-stage stochastic optimization problems, the objective is to minimize (maximize) expected total value of a random cost (profit). This approach implicitly assumes that decision makers are risk-neutral. These models have been first introduced by Dantzig [13] and Beale [14]. In order to solve these models, the L-shaped Method is proposed by Van Slayke and Wets [5] as an exact solution

method. It is based on decomposing a large sized problem into several small sized problems and effective on decreasing the solution times significantly.

However, the attitude towards risk can change depending on the definition of problems considered and the uncertainties included. Risk-averse approaches to the stochastic optimization problems are preferred if the aim is to observe how the randomness and the risk associated with this affect the decisions. In the following parts of this section, we present the literature review on risk-averse two-stage stochastic optimization models and the solution methods.

Risk-aversion can be modeled in different ways such as using chance constraints, concave utility functions, stochastic dominance, coherent risk measures, and mean-risk models. In this section, we present our review on models with mean-risk and coherent risk measures.

Ahmed [15] considers mean-risk objective models in two-stage stochastic programming and examines if these objective functions conserve convexity. A cutting plane algorithm is proposed as an exact solution method for a linear stochastic program involving mean absolute semi-deviation in the objective. Cutting plane algorithm is based on iteratively calculating the subgradient of the objective function and obtaining a lower (or upper) approximation for the objective function. A parametric version of this algorithm is proposed to ease the parametric analysis for the weight involved in the objective function as a trade-off between the risk and expected value. The proposed parametric cutting plane method is used for several problems derived from [16] and it is concluded that this algorithm is more efficient in terms of parametric analysis than using the non-parametric version repeatedly for different parameters.

Ahmed [17], different than [15], considers two types of mean-risk measures involved in the objective function of a linear two-stage stochastic programming model. The first type considered is mean absolute semideviation and the second type is the mean quantile deviation. Deterministic equivalent models are presented for each risk-measure and two parametric cutting plane algorithms are proposed as a solution method for each case. These algorithms are applied to

solve the problems derived from [16] and it is concluded that the model involving mean quantile deviation allows more flexibility in the weight associated with the trade-off between risk and the expected value compared to the model involving mean absolute semideviation.

Kristoffersen [18] considers mean-risk models in two-stage stochastic linear programming. After examining the stability in terms of continuity, convexity and differentiability, three types of mean risk models, namely central deviation, semideviation, and expected excess of target are considered. As an exact solution method, a multicut decomposition algorithm, a variant of the L-shaped method, is proposed and used for linear relaxation of a scheduling problem with a different number of scenarios in a range between 100 and 1000. Contrary to [17], this algorithm is different in terms of the number of added cuts at each iteration and the information that these cuts convey. In our study, we use a solution approach similar to [18] for the proposed risk-averse model involving a different risk-measure, (CVaR).

Schultz and Tiedemann [19] present a two-stage stochastic mixed integer program including CVaR as the measure of risk in the objective. In that study, the second-stage decision variables are integer valued and this makes the objective function and therefore, the problem non-convex. In order to deal with this situation, they develop an algorithm that is based on the Lagrangian relaxation of nonanticipativity constraints and use this method for a problem related to a production planning system considering different number of scenarios ranging from 10 to 100.

Fábián [20] considers two-stage stochastic models that involve CVaR in the objective and in the constraints. A two-stage investment planning problem is considered with an objective function of minimizing the CVaR minus wealth. A decomposition algorithm that is based on a polyhedral representation is proposed as a solution method.

Miller and Ruszczyński [21] propose a formulation for a two-stage stochastic linear programming model with an objective that involves the composition

of conditional risk measures. In that formulation, after the second-stage, the uncertainty still continues. As an exact solution method, they propose a multi-cut decomposition algorithm that is based on the L-shaped method and use this algorithm for a portfolio optimization problem setting with two-stages. The performance of the proposed algorithm is compared to the classic decomposition methods and it is concluded that the proposed method is more efficient.

Noyan [22] presents a two-stage stochastic optimization problem with CVaR in the objective function. Two different decomposition algorithms that are variations of L-shaped method are proposed as a solution method. The first algorithm is based on the approximation of optimal mean-risk function of recourse rather than the expected value of recourse. The second algorithm is based on the sub-gradient based approach that is presented by Ahmed [17] for the mean quantile deviation. Two versions for these algorithms (single-cut and multi-cut) are introduced and their performances are evaluated for a disaster management problem. Formulation of the problem involves binary first-stage decision variables while the second-stage decision variables are continuous. Therefore, the proposed algorithms are used with a master problem in the form of a mixed integer program. In terms of performance, the multi-cut version requires fewer number of iterations compared to the single-cut version. However, the time required to solve the problem changes depending on the trade-off between the number of cuts and the iterations. Single-cut algorithm using a sub-gradient approach requires more optimality cuts compared to the others.

Chapter 3

Problem Definition and Formulation

This chapter consists of three sections. In Section 3.1, we give the problem definition and in Section 3.2, we briefly explain the risk measures and Conditional Value at Risk (CVaR) as preliminaries. We formulate our problem in Section 3.3.

3.1 Problem Definition

Fossil fuels are exhaustible resources and their usage as an energy source releases toxic gases to the atmosphere. Therefore, there has been a tendency towards renewable energy sources. Using these resources to generate electricity however, requires careful planning as they are intermittent. The construction of a renewable energy plant to satisfy the demand for energy necessitates large amount of investments. This situation, along with the intermittency of the renewable resources, will affect the investment decisions that are made by decision makers having different attitudes towards risk.

A hybrid renewable energy system with an aim of satisfying the demand for

energy is proposed by Kocaman et al. [3]. The hybrid system components in that study are hydropower stations, solar panels, diesel generators, and transmission lines. The demand for energy is satisfied through the hydropower stations, solar panels and diesel generators. Producing the energy through diesel generators is expensive and therefore they are used as a backup. In the system, solar panels are located at the demand points, and therefore, the energy generated using the solar panels is not transmitted through the transmission lines. Conventional hydropower stations are used in order to generate electricity using hydropower. Electricity generation mechanism for these systems is as follows: The potential energy stored in a certain amount of water turns into kinetic energy as it is released from a certain height. The movement of water is used to turn the large turbines and the generators are used in order to convert this kinetic energy to electrical energy. Then, generated electricity is distributed to the demand locations through the transmission lines. In the system, excess amount of water in reservoirs is allowed to be spilled.

The proposed model for this hybrid renewable energy system in [3] is a two-stage stochastic program and the source of the uncertainty is the amount of water inflow to the reservoirs. The first-stage decisions that are made before the realization of the uncertainty are related to the size of the reservoirs, solar panels and hydropower generators and the capacity of the transmission lines. Second-stage decisions that are made after the realization of the uncertainty are related to the mismatched demand for energy, transmitted energy and water levels at reservoirs.

The general form of the objective function in two-stage stochastic programming problems is to optimize the total expected value such as minimizing the expected total cost. Therefore, in these models, there is an implicit assumption that the decision maker is risk-neutral. However, the high cost of building a hybrid renewable energy system and the intermittency of the renewable resources make a risk-averse approach more realistic. In this context, considering the decision makers with different attitudes towards risk is crucial.

In our study, we adapt the model proposed by Kocaman et al. [3] and formulate a risk-averse two-stage stochastic linear optimization model by incorporating CVaR as the measure of risk-aversion in the objective function so that the actions of decision makers with different risk-aversion attitudes can be examined. In this way, we will be able to see how different attitudes towards risk will affect the investment decisions in the hybrid renewable energy system planning problem. In the following sections, we introduce the definition of CVaR and then we give our problem formulation.

3.2 Conditional Value at Risk (CVaR)

In this section, we briefly describe the coherent measures of risk and then give the definition of CVaR which is a coherent risk measure.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $X : \Omega \rightarrow \mathbb{R}$ that represents the uncertain outcome of costs. Let $\mathcal{Z} = \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, $p \in [1, \infty]$ be a space of p -integrable random costs. Then the function $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ is called a *coherent risk measure* if the following four axioms are satisfied [23].

(A1) Convexity: $\rho(\alpha X + (1 - \alpha)Y) \leq \alpha\rho(X) + (1 - \alpha)\rho(Y)$, $\forall X, Y \in \mathcal{Z}$ and $\alpha \in [0, 1]$.

(A2) Monotonicity: If $X, Y \in \mathcal{Z}$ and $X \succeq Y$, then $\rho(X) \geq \rho(Y)$

(A3) Translation Invariance: $\rho(X + a) = \rho(X) + a$, $\forall a \in \mathbb{R}, X \in \mathcal{Z}$

(A4) Positive Homogeneity: $\rho(tX) = t\rho(X)$ $\forall X \in \mathcal{Z}, t \geq 0$

Let $X, Y \in \mathcal{Z}$, then $X \succeq Y$ if $X(\omega) \geq Y(\omega)$ for all outcomes $\omega \in \Omega$

Conditional value at risk (CVaR) and its properties as a coherent risk measure are explained in detail in Rockafellar and Uryasev [24]. For atomless case, CVaR is a conditional expectation of the losses that are greater than or equal to the

Value at Risk (VaR) at level α .

For $Z \in \mathcal{Z}$ and $\alpha \in [0, 1)$, CVaR at level α is defined as follows:

$$\text{CVaR}_\alpha(Z) = \inf_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1 - \alpha} \mathbb{E}[(Z - \eta)_+] \right\}. \quad (3.1)$$

CVaR is closely related to the VaR, which is also a risk measure and corresponds to the α quantile of a loss distribution [25, 26]. In other words, it is a threshold that one can be at least α percent sure that the losses will not exceed that threshold. For $Z \in \mathcal{Z}$ and $\alpha \in [0, 1)$, VaR at level α is defined as follows:

$$\text{VaR}_\alpha(Z) = \min\{x : P\{Z \leq x\} \geq \alpha\}. \quad (3.2)$$

In equation (3.1), the infimum is attained at α percent quantile of Z . In other words, the minimizer is $\eta^* = \text{VaR}_\alpha[Z]$. For details, see Ogryczak and Ruszczyński [27]. Despite the close relationship between CVaR and VaR, one can say that CVaR is more sensitive to the extreme losses. For example, two distributions representing losses may have the same VaR value at a significance level α , even if one has a shorter upper tail (probability of having extreme losses is decreasing fast) while the other has a longer upper tail (probability of having extreme losses is decreasing slowly). However, CVaR value at level α for these two distributions will differ as it considers the expected extreme values. In addition to these, VaR is not a coherent risk measure as it does not satisfy (A1). For details, see Artzner et al. [23].

3.3 Problem Formulation

We propose a risk-averse two-stage stochastic optimization model to a hybrid renewable energy system planning problem. In our proposed model, we include the CVaR in the objective function in the following form;

$$\min \quad \lambda \mathbb{E}(Z) + (1 - \lambda) \text{CVaR}_\alpha(Z). \quad (3.3)$$

In this function, Z corresponds to the total cost and $\lambda \in [0, 1]$ is the weight given to the expected value while $(1 - \lambda)$ is the weight given to the CVaR value.

In this formula, when λ decreases, the decision maker gets more risk-averse and when $\lambda = 1$, then the decision maker is risk-neutral. The uncertainty involved in this model can be related to the amount of water inflow to the reservoirs, solar radiation and demand. Sets, parameters, and the decision variables of the proposed risk-averse model are as follows:

Sets

- I : The set of hydropower generation locations
- J : The set of demand (solar power generation) locations
- $\mathcal{T}$: The set of time periods $\mathcal{T} = \{1, 2, \dots, T\}$
- Ω : The set of scenarios

Parameters

- n : Length of the time period
- d_{hydro} : Used to express the investment cost on yearly basis, for hydropower stations, $d_{hydro} = i/(1 - (i + 1)^{-m_{hydro}})$, where m_{hydro} is the length of the planning horizon for hydropower.
- d_{solar} : Used to express the investment cost on yearly basis, for solar power stations, $d_{solar} = i/(1 - (i + 1)^{-m_{solar}})$, where m_{solar} is the length of the planning horizon for solar power.
- d_{tr} : Used to express the investment cost on yearly basis, for transmission lines, $d_{tr} = i/(1 - (i + 1)^{-m_{tr}})$, where m_{tr} is the the length of the planning horizon for transmission lines
- l_{ij} : Percentage of power loss while transmitting electricity from hydropower generation $i \in I$ to demand location $j \in J$
- g : Standard acceleration due to gravity
- h_i : Height of the reservoir in hydropower generation location $i \in I$
- β : Efficiency of hydropower station $\beta \in [0, 1]$
- γ : Efficiency of solar panels $\gamma \in [0, 1]$

- C_i^S : Unit cost of reservoir capacity for hydropower generation location $i \in I$
 C_i^{PG} : Unit cost of generator capacity for hydropower generation location $i \in I$
 C_j^M : Unit cost of solar energy at demand location $j \in J$
 C_{ij}^{TR} : Unit cost of transmission line capacity between hydropower generation location $i \in I$ and demand location $j \in J$
 μ_j : Unit cost of generating electricity using diesel generator (penalty for mismatched demand at demand location $j \in J$)
 p_ω : Probability of scenario $\omega \in \Omega$
 $W_{it\omega}$: Amount of water runoff to hydropower generation location $i \in I$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
 $N_{jt\omega}$: Amount of solar radiation at location $j \in J$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
 $D_{jt\omega}$: Amount of demand at location $j \in J$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$

First Stage Decision Variables

- S_{max_i} : Size of active reservoir capacity in hydropower generation location $i \in I$
 M_j : Size of solar panels at demand location $j \in J$
 PG_{max_i} : Size of generator in hydropower generation location $i \in I$
 $TR_{max_{ij}}$: Amount of maximum energy transmitted from hydropower generation location $i \in I$ to demand location $j \in J$

Second Stage Decision Variables

- $S_{it\omega}$: Amount of water stored in the reservoir in hydropower generation location $i \in I$ under scenario $\omega \in \Omega$ in period $t \in \mathcal{T}$
- $Z_{jt\omega}$: Amount of mismatched demand at demand location $j \in J$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
- $TR_{ij t\omega}$: Amount of electricity sent from hydropower generation location $i \in I$ to demand location $j \in J$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
- $L_{it\omega}$: Amount of water spilled from the reservoir in hydropower generation location $i \in I$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$
- $R_{it\omega}$: Amount of water released from the reservoir in hydropower generation location $i \in I$ in period $t \in \mathcal{T}$ under scenario $\omega \in \Omega$

The extensive formulation of proposed risk-averse model is as follows:

$$\min \left(\sum_{i \in I} c_i^S S_{max_i} + \sum_{i \in I} c_i^{PG} PG_{max_i} \right) d_{hydro} + \sum_{j \in J} c_j^M M_j d_{solar} \quad (3.4)$$

$$+ \sum_{i \in I, j \in J} c_{ij}^{TR} TR_{max_{ij}} d_{tr} \quad (3.5)$$

$$+ \lambda \sum_{\omega \in \Omega} p_{\omega} \left(\sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j \right) \quad (3.6)$$

$$+ (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \sum_{\omega \in \Omega} p_{\omega} \left[\sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j - \eta \right]_+ \right) \quad (3.7)$$

s.t.

$$S_{itw} \leq S_{max_i}, \quad \forall i \in I, t \in \mathcal{T}, \omega \in \Omega \quad (3.8)$$

$$S_{itw} = S_{it-1w} + W_{itw} - R_{itw} - L_{itw}, \quad \forall i \in I, t \in \mathcal{T} \setminus \{1\}, \omega \in \Omega \quad (3.9)$$

$$S_{i1w} = S_{max_i} + W_{i1w} - R_{i1w} - L_{i1w}, \quad \forall i \in I, \omega \in \Omega \quad (3.10)$$

$$S_{iT\omega} = S_{max_i}, \quad \forall i \in I, \omega \in \Omega \quad (3.11)$$

$$R_{itw} g h_i \beta \leq PG_{max_i} n, \quad \forall i \in I, t \in \mathcal{T}, \omega \in \Omega \quad (3.12)$$

$$\sum_{j \in J} TR_{ijtw} = R_{itw} g h_i \beta, \quad \forall i \in I, t \in \mathcal{T}, \omega \in \Omega \quad (3.13)$$

$$TR_{ijtw} \leq TR_{max_{ij}} n, \quad \forall i \in I, j \in J, t \in \mathcal{T}, \omega \in \Omega \quad (3.14)$$

$$D_{jtw} \leq Z_{jtw} + N_{jtw} M_j \gamma + \sum_{i \in I} TR_{ijtw} (1 - l_{ij}), \quad \forall j \in J, t \in \mathcal{T}, \omega \in \Omega \quad (3.15)$$

$$S_{max_i}, PG_{max_i}, M_j, TR_{max_{ij}} \geq 0, \quad \forall i \in I, j \in J \quad (3.16)$$

$$S_{itw}, R_{itw}, L_{itw} \geq 0, \quad \forall i \in I, t \in \mathcal{T}, \omega \in \Omega \quad (3.17)$$

$$TR_{ijtw} \geq 0, \quad \forall i \in I, j \in J, t \in \mathcal{T}, \omega \in \Omega \quad (3.18)$$

$$Z_{jtw} \geq 0, \quad \forall j \in J, t \in \mathcal{T}, \omega \in \Omega \quad (3.19)$$

$$\eta \in \mathbb{R} \quad (3.20)$$

(3.4) - (3.7) correspond to the objective function of the risk-averse model. (3.4) and (3.5) are the annualized construction costs of the hybrid system. They include the annualized cost of constructing reservoirs, transmission lines, power

generators, and solar panels. We consider that each system has a useful life time and by using this information, we come up with a cost parameter that corresponds to an annual cost of construction for each system. The formula used for this parameter is $d = i/(1 - (i + 1)^{-m})$, where m is the planning horizon and i is the discount rate. (3.6) is the expected cost of the energy that is produced by the diesel generators and (3.7) is CVaR value of this cost at a significance level α . Constraint (3.8) is a capacity constraint and guarantees that the amount of water stored in the reservoir at any time period $t \in \mathcal{T}$ and under any scenario $\omega \in \Omega$ is less than the reservoir's capacity. Constraints (3.9) - (3.11) are the mass balance constraints at the hydropower stations. Here, we assume that at the first ($t = 1$) and at the last ($t = T$) time periods, the reservoirs are full. These constraints guarantee that the amount of water stored at a time period t is equal to the sum of the amount of water left from the previous time period $t - 1$ and amount of water inflow to the reservoir at time period t minus amount of water released and spilled from the reservoir at time period t . Constraint (3.12) guarantees that the amount of energy produced at a hydropower station cannot exceed its generators' capacity at any time period $t \in \mathcal{T}$ and under any scenario $\omega \in \Omega$. In (3.13), the amount of energy transmitted from a hydropower station to the demand points will be equal to the amount of energy produced at that station at any time period $t \in \mathcal{T}$ and under each scenario $\omega \in \Omega$. Constraint (3.14) guarantees that the amount of energy transmitted from a hydropower station to the demand points does not exceed the capacity of the transmission lines. In (3.15), at any time period $t \in \mathcal{T}$ and under each scenario $\omega \in \Omega$, the demand is satisfied from the sum of the energy produced through solar panels, energy transmitted from hydropower stations, and energy produced by diesel generators. (3.16) - (3.20) are domain constraints.

In this model, the objective function is non-linear due to the positive part that comes from the formulation of CVaR. Therefore, in order to linearize the model, we define a new random variable Y_ω and replace $\left[\sum_{j \in J, t \in \mathcal{T}} Z_{jt\omega} \mu_j - \eta \right]_+$ in the

objective function with Y_ω and add the following constraints to the model:

$$Y_\omega \geq \sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j - \eta, \quad \forall \omega \in \Omega \quad (3.21)$$

$$Y_\omega \geq 0, \quad \forall \omega \in \Omega. \quad (3.22)$$

Note that when the difference $(\sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j - \eta)$ is negative, then (3.22) will be an active constraint while (3.21) becomes a redundant constraint. Similarly, when the difference $(\sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j - \eta)$ is positive, then (3.21) will become an active constraint while (3.22) becomes redundant [24].

The risk-averse model after the linearization process is as follows:

$$\min \left(\sum_{i \in I} c_i^S S_{max_i} + \sum_{i \in I} c_i^{PG} PG_{max_i} \right) d_{hydro} + \sum_{j \in J} c_j^M M_j d_{solar} \quad (3.4)$$

$$+ \sum_{i \in I, j \in J} c_{ij}^{TR} TR_{max_{ij}} d_{tr} \quad (3.5)$$

$$+ \lambda \sum_{\omega \in \Omega} p_\omega \left(\sum_{j \in J, t \in \mathcal{T}} Z_{jtw} \mu_j \right) \quad (3.6)$$

$$+ (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \sum_{\omega \in \Omega} p_\omega Y_\omega \right) \quad (3.7)$$

s.t.

$$(3.8) - (3.22).$$

Chapter 4

Solution Methodology

We suggest a scenario-wise decomposition algorithm to solve our risk-averse two-stage stochastic programming problem described in Chapter 3. This algorithm is based on the L-shaped method, which is generally used to solve the two-stage stochastic programming problems. The general idea is to decrease the solution time by solving the several small sized problems instead of solving one large sized problem.

L-shaped method with a single-cut approach is proposed by Van Slyke and Wets [5], while a multi-cut version is proposed by Birge and Louveaux [28]. The difference between these algorithms is the number of optimality cuts added to the master problem at each iteration. In the single-cut version, in each iteration, only one cut (as an aggregated cut) is added to the master problem, while in the multi-cut version, the number of optimality cuts added changes depending on scenarios. Therefore, a larger number of iterations is needed to solve the problem with the single-cut version since the cut conveys less information to the master problem. On the other hand, the size of the master problem increases by a greater amount at each iteration in the multi-cut version. This indicates that there is a trade-off between the number of cuts added to the master problem and the number of iterations needed to solve the problem. The performance of these algorithms may be specific to the problem. However, it is stated in [28]

that the multi-cut version is expected to be more efficient than the single-cut version. Therefore, our proposed scenario-wise decomposition algorithm is based on a multi-cut approach.

In the rest of this chapter, starting from a general representation of the two-stage stochastic programming problem with CVaR, we introduce our reformulation of the master problem and then we give the proposed scenario-wise decomposition algorithm, where the general idea depends on the study by Kristoffersen et al. [18].

$$\min c^T x + \lambda \sum_{\omega \in \Omega} p_{\omega} q_{\omega}^T y_{\omega} \quad (4.1)$$

$$+ (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \sum_{\omega \in \Omega} p_{\omega} [q_{\omega}^T y_{\omega} - \eta]_+ \right) \quad (4.2)$$

s.t.

$$Ax = b \quad (4.3)$$

$$T_{\omega} x + W y_{\omega} = h_{\omega}, \quad \forall \omega \in \Omega \quad (4.4)$$

$$y_{\omega} \geq 0, \quad \forall \omega \in \Omega \quad (4.5)$$

$$x \geq 0 \quad (4.6)$$

$$\eta \in \mathbb{R}. \quad (4.7)$$

The problem in (4.1) - (4.7) corresponds to a general two-stage stochastic programming problem with CVaR. In this problem, $x \in \mathbb{R}^{n_1}$ is a vector of the first-stage decision variables, while $y_{\omega} \in \mathbb{R}^{n_2}$ represents the vector of second-stage decision variables for each scenario $\omega \in \Omega$. Parameters in this formulation are as follows: $c \in \mathbb{R}^{n_1}$ and $b \in \mathbb{R}^{m_1}$ are deterministic vectors while $A \in \mathbb{R}^{m_1 \times n_1}$ and $W \in \mathbb{R}^{m_2 \times n_2}$ are deterministic matrices. On the other hand, $T_{\omega} \in \mathbb{R}^{m_2 \times n_1}$, $h_{\omega} \in \mathbb{R}^{m_2}$ and $q_{\omega} \in \mathbb{R}^{n_2}$ may be random.

(4.1) in the objective consists of the cost associated with the first-stage decision variables and the expected cost related to the second-stage decision variables with a weight $\lambda \in [0, 1]$. (4.2) corresponds to the CVaR value of the second-stage cost at level α and the weight of this term is $(1 - \lambda)$.

The objective function of the problem in (4.1) - (4.7) is non-linear because of

the positive part that comes from the formulation of CVaR. In order to linearize this part and construct a linear master problem, we make several modifications to the objective function. The reader may refer to Kristoffersen et al. [18] for details. First we replace the positive part $[q_\omega^T y_\omega - \eta]_+$ in (4.2) with $(\max\{q_\omega^T y_\omega, \eta\} - \eta)$. With this modification, the function of the positive part is preserved. When $q_\omega^T y_\omega$ is greater than η , the expression $(\max\{q_\omega^T y_\omega, \eta\} - \eta)$ returns a positive value of $(q_\omega^T y_\omega - \eta)$. If η is greater than the value of $q_\omega^T y_\omega$, then the expression returns a value of 0. After this modification, the new problem becomes;

$$\min c^T x + \lambda \sum_{\omega \in \Omega} p_\omega q_\omega^T y_\omega \quad (4.1)$$

$$+ (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \left(\sum_{\omega \in \Omega} p_\omega \max\{q_\omega^T y_\omega, \eta\} - \eta \right) \right) \quad (4.8)$$

s.t.

$$(4.3) - (4.7).$$

As a second modification we replace $q_\omega^T y_\omega$ with a new decision variable of Φ_ω and we replace the maximum function with an another decision variable of θ_ω . Then we add the following constraints to the problem:

$$\Phi_\omega \geq q_\omega^T y_\omega, \quad \forall \omega \in \Omega \quad (4.9)$$

$$\theta_\omega \geq \Phi_\omega, \quad \forall \omega \in \Omega \quad (4.10)$$

$$\theta_\omega \geq \eta, \quad \forall \omega \in \Omega. \quad (4.11)$$

After these modifications, the new problem is as follows;

$$\min c^T x + \lambda \sum_{\omega \in \Omega} p_\omega \Phi_\omega \quad (4.12)$$

$$+ (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \left(\sum_{\omega \in \Omega} p_\omega \theta_\omega - \eta \right) \right) \quad (4.13)$$

s.t.

$$(4.3) - (4.7)$$

$$(4.9) - (4.11)$$

Constraint (4.9) combined with objective function forces Φ_ω to be equal to $q_\omega^T y_\omega$. Similarly constraints (4.10) and (4.11) force θ_ω to be equal to $\max\{\Phi_\omega, \eta\}$.

After completing the modifications, the master problem that will be used in the scenario-wise decomposition algorithm is as follows:

$$\min c^T x + \lambda \sum_{\omega \in \Omega} p_{\omega} \Phi_{\omega} + (1 - \lambda) \left(\eta + \frac{1}{1 - \alpha} \left(\sum_{\omega \in \Omega} p_{\omega} \theta_{\omega} - \eta \right) \right) \quad (4.14)$$

s.t.

$$\Phi_{\omega} \geq \pi_{m(\omega)}^T (h_{\omega} - T_{\omega} x), \quad m(\omega) = 1, \dots, l_{\omega}, \omega \in \Omega \quad (4.15)$$

$$\sigma_m^T (h_{\omega} - T_{\omega} x) \leq 0, \quad m = 1, \dots, k, \omega \in \Omega \quad (4.16)$$

$$\theta_{\omega} \geq \Phi_{\omega}, \quad \forall \omega \in \Omega \quad (4.17)$$

$$\theta_{\omega} \geq \eta, \quad \forall \omega \in \Omega \quad (4.18)$$

$$Ax = b \quad (4.19)$$

$$x \geq 0 \quad (4.20)$$

$$\Phi_{\omega}, \theta_{\omega}, \eta \in \mathbb{R}, \quad \forall \omega \in \Omega. \quad (4.21)$$

Constraint (4.15) corresponds to the optimality cuts and they provide a lower approximation for Φ_{ω} in search for the optimal solution. $\pi_{m(\omega)}$ in (4.15), is the dual vector corresponding to the optimal basis of the problem (4.27) - (4.29). Note that the constraints (4.4) and (4.5) that are included in the other models stated before are not present in the master problem. That is because we include the feasibility cuts (4.16) in the master problem and these cuts are added to assure that there are nonnegative y values that satisfy the constraint (4.4) for a given x . In (4.16), σ_m is the dual vector corresponding to the optimal basis of the problem (4.22) - (4.24).

The algorithm that we propose solves the master problem at each iteration and obtains a candidate solution for the first-stage decision variables, x . Then, feasibility cuts are used in order to guarantee that, for given x , there are non-negative second-stage decision variables y . If this is not the case for a particular x , then a feasibility cut is added to the master problem to cut the solution x . Then, optimality cuts are used to make an approximation for Φ_{ω} in search for the optimal solution.

Feasibility cuts are omitted in a case of complete recourse, which means, there are always nonnegative second-stage decision variables regardless of the value of x

[29]. In our model, for any solution x , we have nonnegative second-stage decision variables. For instance, if x is equal to zero, then the demand will be satisfied using only the diesel generators. If x has a nonnegative value, then the demand may be satisfied through the renewables and the diesel generators. Therefore, in our algorithm, we omit the feasibility cuts.

The proposed scenario-wise decomposition algorithm for general representation of the two-stage stochastic programming problem with CVaR is represented below.

Scenario-Wise Decomposition Algorithm

Step 0: (Initialization) $k = p = 0$, $l_\omega = 0$, $\omega \in \Omega$

Step 1: Set $p = p+1$ and solve the linear programming problem (master problem)

$$\begin{aligned} & \min (4.14) \\ & s.t. \\ & (4.15) - (4.21) \end{aligned}$$

When there is no optimality cut for some $\omega \in \Omega$, then ignore Φ_ω , θ_ω in the computation. Let x^p be an optimal solution.

Omit **Step 2** in the case of complete recourse.

Step 2: For all $\omega \in \Omega$ solve the following optimization problem (4.22) - (4.24) to check the feasibility of the solution $x = x^p$.

$$\min e^T v^+ + e^T v^- \tag{4.22}$$

s.t.

$$Wy + v^+ + v^- = h_\omega - T_\omega x^p \tag{4.23}$$

$$y \geq 0, v^+ \geq 0, v^- \geq 0 \tag{4.24}$$

This problem with a zero optimal value ensures that there are non-negative second-stage decision variables y , that satisfies the constraint

$$Wy = h_\omega - T_\omega x^p \tag{4.25}$$

and indicating that x^p is a feasible solution. On the other hand, if the optimum value is greater than zero, then the constraint (4.25) is not satisfied indicating the solution x^p is not feasible.

If the optimal value of this problem is zero for all $\omega \in \Omega$ then go to **Step 3**. If the optimal value is greater than zero, then add the feasibility cut (4.26) to the master problem to cut off the solution $x = x^p$.

$$\sigma_m^T(h_\omega - T_\omega x) \leq 0 \quad (4.26)$$

In (4.26), σ_m is the dual vector corresponding to the optimal basis of the problem (4.22) - (4.24). Set $k = k + 1$ and go to **Step 1**. Repeat **Step 1** and **Step 2** until for all ω the optimum value of the problem (4.22) - (4.24) is zero. Then go to **Step 3**.

Step 3: For all $\omega \in \Omega$ solve the linear programming problems (4.27) - (4.29).

$$\min q_\omega^T y_\omega \quad (4.27)$$

s.t.

$$W y_\omega = h_\omega - T_\omega x^p \quad (4.28)$$

$$y_\omega \geq 0 \quad (4.29)$$

Let π_ω^p be the dual vector corresponding to the optimal basis of the problem in (4.27) - (4.29).

If $(\pi_\omega^p)^T(h_\omega - T_\omega x^p) > \Phi_\omega^p$, then add the following optimality cut (4.30) to the master problem, set $l_\omega = l_\omega + 1$ and go to the **Step 1**.

$$(\pi_\omega^p)^T(h_\omega - T_\omega x^p) \leq \Phi_\omega^p \quad (4.30)$$

If $(\pi_\omega^p)^T(h_\omega - T_\omega x^p) \leq \Phi_\omega^p$ for all $\omega \in \Omega$, then stop, x^p is an optimal solution.

Our algorithm is based on the multi-cut L-shaped method suggested by Birge and Louveaux [28] for risk-neutral two-stage stochastic problems. The reader is referred to [28] for the details.

Chapter 5

Computational Study

This chapter consists of three sections. In Section 5.1, we present the scenario generation algorithm and then, in Section 5.2, we explain the data we use in this study. In Section 5.3, we provide the numerical results and their interpretations.

5.1 Scenario Generation

In the stochastic programming problems, the number of scenarios is important in order to model the uncertainty. Using an appropriate number of scenarios will enable one to achieve the true optimal solution. In our computational study, we assume that the amount of water inflow to the reservoirs is uncertain, and therefore, we use a scenario generation algorithm for the streamflow values. The method chosen will be explained in this section.

A nearest neighbor resampling algorithm that is developed by Lall and Sharma [30] to generate streamflows uses historical streamflow values of time period $t - 1$ and then resamples a flow value for the next time period t . For a lag-1 time dependence, this method first finds the k nearest neighbors of a given flow value at time period $t - 1$ and then resamples one of these k neighbors by using a kernel function that gives more weight to the nearest neighbor and less weight

to the farthest neighbor. The simulated flow value for the next time period t is found using the time index of the resampled neighbor. The flow values that are simulated using this method, however, are limited with the historical data.

As an extension to the nearest neighbor resampling algorithm, Prairie et al. [4] develop a new method that can simulate flow values that are not observed in the historical data. This method is called lag-1 modified k -NN. It uses local polynomial fitting for each time period t given the previous time period $t - 1$ to estimate the mean flow and simulates a flow value for the time period t by adding a resampled residual to the estimated mean flow. Kernel function in (5.2) is used to resample the residuals and it is the same with the one that was used by Lall and Sharma [30]. We use this lag-1 modified k -NN Algorithm in our study and the steps of the algorithm are as follows:

1. Fit a local polynomial to the flow values x_t of the time period t depending on the flow values x_{t-1} of previous time period $t - 1$ as in (5.1).

$$x_t = f(x_{t-1}) + e_t \quad (5.1)$$

Here, e_t are the residuals and will be saved while f is the fitted polynomial.

2. Find the estimated mean flow value, $(\hat{x}_t)$, that is obtained from (5.1) omitting e_t .
3. Find the k nearest neighbors of x_{t-1} and choose one of them by using the kernel function in (5.2). Simulate the flow value (x_t^*) for time period t by adding the residual of chosen neighbor (e_i) to the estimated mean flow $\hat{x}_t$. Obtained simulated value, then becomes $x_t^* = \hat{x}_t + e_i$.

$$K(i) = \frac{\frac{1}{i}}{\sum_{i=1}^k \frac{1}{i}} \quad (5.2)$$

5.2 Data

In our model, we use the same cost parameters as in [3] but we present our results for a case study from Turkey. Table 5.1 represents the adapted parameters from [3].

Parameter Name	Value
C_i^S	3 \$/m ³
C_i^{PG}	500 \$/kW
C_j^M	200 \$/m ²
C_{ij}^{TR}	55 \$/kW
μ_j	0.25 \$/kWh
β	0.88
γ	0.12
l_{ij}	0.05
d_{hydro}	60 years
d_{solar}	30 years
d_{tr}	40 years
i	0.05

Table 5.1: Parameters adapted from [3]

The length of each time period used in our study is one hour. We set the length of the planning horizons for solar energy to 30 years, for hydropower to 60 years and for transmission lines to 40 years and use the annualization formula ($d = i/(1-(i+1)^{-m})$), where m is the the length of the planning horizon and i is the discount rate. For demand, we consider the electrical distribution data of three cities, Antalya, Isparta, and Burdur, which are located in the Mediterranean Region of Turkey. Demand data are obtained from the records of the Energy Market Regulatory Authority of Turkey [31]. The solar radiation data of the region are obtained via a web-based data center provided by Homer Energy [32]. Figures 5.1 and 5.2 show the hourly solar radiation and demand data, respectively. In this computational study, we assume that the uncertainty is amount of water inflow to the reservoirs, while solar radiation and demand data are taken as deterministic.

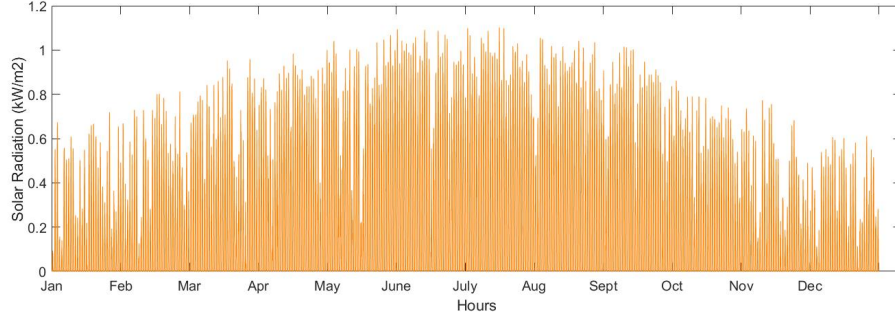


Figure 5.1: Hourly Solar Radiation Data

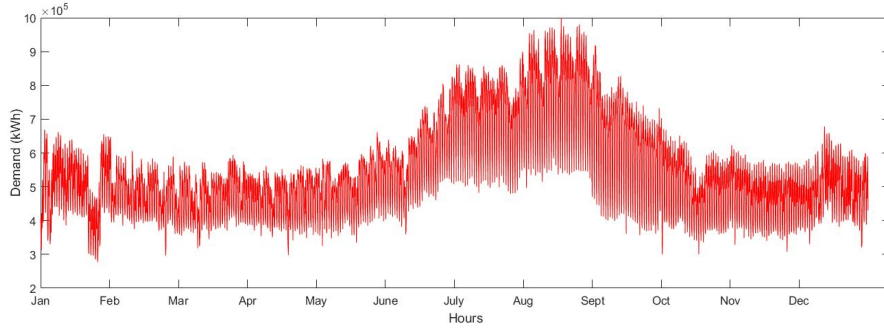


Figure 5.2: Hourly Demand Data

For streamflow data, we use Manavgat River's streamflow values for the years 1997 - 2015. The daily streamflows of the river are obtained from the General Directorate of State Hydraulic Works [33]. The streamflow data recorded consist of one reading for a day in the form of m^3/s . In order to obtain flow data for each hour, we assume that the value read in a day is constant and the same amount of water accumulates for 3600 seconds. Using this assumption, we obtain the hourly flow values for each of the 19 years.

In this study, we used lag-1 modified k -NN method of Prairie et al. [4], which is explained in Section 5.1, to generate scenarios for daily streamflow data based on the historical data of Manavgat River for the years of 1997 - 2015. The historical data for the daily flows of 19 years are represented in Figure 5.3 and the generated 50 scenarios of flow values are represented in Figure 5.4.

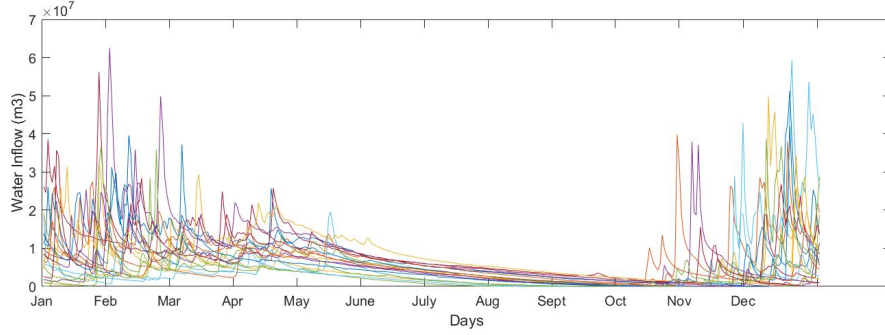


Figure 5.3: Daily Historical Streamflows of Manavgat River for Years 1997 - 2015

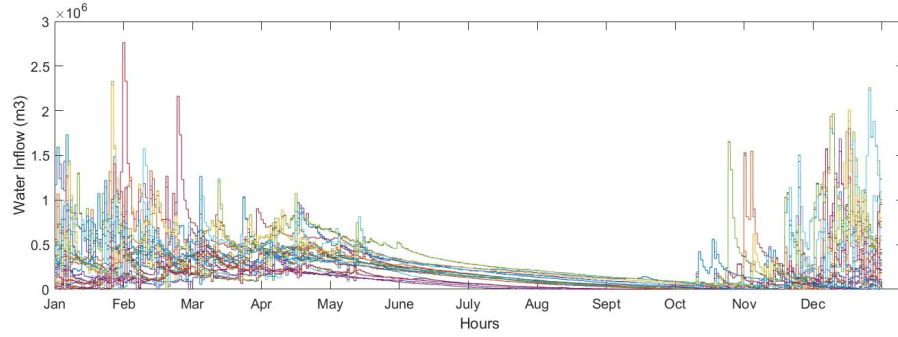


Figure 5.4: 50 Scenarios for Hourly Streamflows Generated Using Modified k -NN

For local polynomial fitting step of scenario generation algorithm, we use a toolbox developed by Jekabsons for Matlab [34]. In Lall and Sharma [30], in order to determine the appropriate number of neighbors, the Generalized Cross Validation (GCV) technique is used. This technique is based on leaving one value from the data at each time and fitting a model to the remaining data to forecast this leftout value. The reader may refer to Craven and Wahba [35] for details. The formula of GCV used in [30] is as follows:

$$GCV = \frac{\sum_{i=1}^n e_i^2/n}{(1 - 1/\sum_{j=1}^k 1/j)^2}. \quad (5.3)$$

In this formula, n is the sample size, while e_i corresponds to the error of estimate. In other words, it is the difference between real value of a data point and its estimate that is obtained by leaving it out and estimating its value using a fitted model to the $n - 1$ data points. They consider appropriate number of nearest neighbors as $\sqrt{n}$ since there is no significant difference in Generalized Cross Validation (GCV) values when n is between 50 and 200. However, in our case,

we have a sample size of 19, and therefore, in order to determine the appropriate number of nearest neighbors (k) we used GCV formula that is used by Lall and Sharma in [30]. A similar approach is used by Singhrattana et al. [36] in order to generate streamflows with modified k -NN algorithm by using a sample size of 21.

Similarly, to determine appropriate kernel function type and the order for local polynomial fitting, we use the same GCV formula. Local polynomial fitting is a non-parametric regression method and instead of fitting a global function to the data, it considers fitting low-degree polynomials to the localized subsets of the data [34]. Prairie et al. [4] use a similar approach to Loader [37] and Rajagopalan and Lall [38] to determine the localized subsets. They fit a local polynomial to the k -NN of each data point. Toolbox by Jekabsons [34], enable us to use this approach and provides a different types of kernel functions to be used as a k -NN weight function for local polynomial fitting. Therefore, we use GCV formula to choose the appropriate kernel function, number of neighbors and the polynomial order.

The appropriate kernel function type, number of neighbors and the polynomial order that end up with the minimum GCV value are “Triweight”, $k = 7$ and $p = 2$ respectively. The GCV values for different number of neighbors with a kernel type “Triweight” and an order $p = 2$ is presented in Table 5.2. The kernel function is used for density estimation for non-parametric analysis. For details reader may refer [34]. Let $D(u) : \mathbb{R} \rightarrow \mathbb{R}$ be any function that satisfies $\int_{-\infty}^{\infty} D(u)du = 1$. Then, the function $D(u)$ is called a kernel function. The formula of kernel function type “Triweight” is as follows:

$$D(u) = \begin{cases} (1 - u^2)^3 & \text{if } |u| < 1, \\ 0 & \text{o.w.} \end{cases} \quad (5.4)$$

5.3 Numerical Results

In this section, we give the numerical results and their interpretations. Our computational experiments are performed on a dual 2.4 GHz Intel Xeon E5-2630 v3

Number of Neighbors (k)	GCV
7	285,240.91
8	324,816.00
9	330,116.75
10	343,032.11
11	346,664.47
12	352,400.06
13	353,836.48
14	350,836.34
15	348,113.84
16	346,736.15
17	347,826.81
18	350,492.25

Table 5.2: GCV Values for Different Number of Neighbors (k)

CPU server with 64GB RAM. We solve the extensive formulation of our model using CPLEX with version 12.6. The proposed scenario-wise decomposition algorithm is coded in MATLAB R2016a and the optimization problems that are formed in Matlab are solved using CPLEX 12.6.3 in parallel mode using up to 32 threads. Then the performances of these two solution methods are compared based on CPU times. In order to observe the changes that are originated from the decision makers' different attitudes towards risk, we consider several instances with different α and λ values. Therefore, we are able to see how results change as the decision makers become more risk-averse. For the number of scenarios, we consider a range starting from 500 scenarios to 1000 scenarios. In this study, when we use the actual demand data, the amount of energy supplied through hydropower is relatively small compared to demand, which leads to high diesel usage. Therefore, in this section, we present numerical results that are obtained using one fourth of the actual demand data in order to observe clearly how uncertainty included in streamflow affects the decisions. We first present the numerical results for 500 scenarios with our interpretations and then we provide the results for larger number of scenarios at the end of this section. Numerical results that are obtained using the real demand data can be seen in Appendix C.

In order to evaluate the performance of the proposed decomposition algorithm, we solve the extensive formulation with CPLEX and compare the CPU times required to solve the model. Comparison of the CPU times for two solution methods for a model consisting 500 scenarios is given in Table 5.3. In this table, the CPU times that are required to solve the problem for 10 different instances are illustrated.

		Extensive Formulation (CPLEX)	Scenario-wise Decomposition Algorithm
α	λ	CPU time (s)	CPU time (s)
0.7	0.3	137,696.77	20,519.79
0.7	0.5	129,266.52	21,734.31
0.7	0.7	131,700.02	19,736.97
0.8	0.3	135,288.21	21,296.40
0.8	0.5	135,363.54	26,418.78
0.8	0.7	131,514.29	28,325.60
0.9	0.3	131,114.26	27,559.41
0.9	0.5	133,299.39	26,820.16
0.9	0.7	174,300.88	23,133.45

Table 5.3: Comparison of CPU times for the algorithm and extensive formulation for 500 scenarios

The results in this table are obtained for different α and λ values that have effects on the decision makers' attitudes towards risk. For instance, as λ decreases or as α increases, the decision maker gets more risk-averse. The effect of these parameters on optimal solution and cost will be examined later in this section. In Table 5.3, the CPU times are recorded in seconds and it is seen that the CPU times of the scenario-wise decomposition algorithm for all instances are less than the CPU times of the extensive formulation. We compare the performance of the algorithm and the extensive formulation only for 500 scenarios since we receive a memory error from CPLEX when we use 600 scenarios. Using the scenario-wise decomposition algorithm, on the other hand, we are able to obtain solutions for 1000 scenarios.

In order to observe the changes in the first-stage decisions under different risk-aversion characteristics, we examine the optimal sizes for reservoirs, solar panels, transmission lines and power generators for different α and λ values in Figure 5.5.

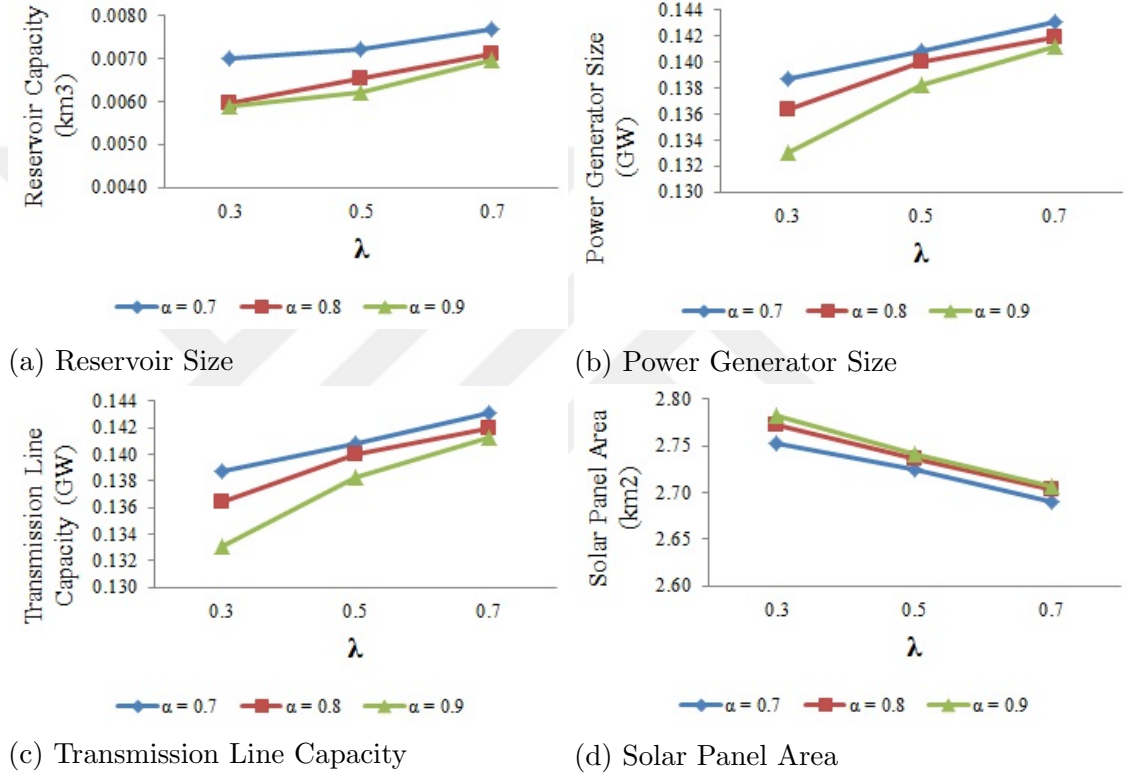


Figure 5.5: Optimal First Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenarios

It is seen from Figure 5.5a that the reservoir size decreases as λ decreases. In other words, as the decision maker gets more risk-averse, the investment made in the reservoirs decreases. Since the amount of water inflow to the reservoirs is uncertain, a risk-averse decision maker will tend to avoid investing in the reservoirs. Similarly, as α increases, the decision maker gets more risk-averse and the reservoir size decreases.

In Figures 5.5b and 5.5c, when the decision maker gets more risk-averse, the capacities of power generators at hydropower stations and transmission lines decrease similar to the reservoir size. In other words, as the decision maker gets

more risk-averse, then the investment made in the hydropower stations decreases. Due to the uncertainty in the amount of water inflow to the reservoirs, a risk-averse decision maker will avoid investing in hydropower.

In Figure 5.5d, when the decision maker gets more risk-averse, the investments made in solar power increase. Energy generated through the solar panels are related to the amount of solar radiation, which is deterministic (same realization is used in each scenario) in our model, and therefore, a risk-averse decision maker will tend to invest in a deterministic source of energy instead of uncertain ones.

The changes in optimal first-stage decisions with respect to different α and λ values for the number of scenarios larger than 500 are similar. See Appendix A.

In order to observe the changes in expected total cost and standard deviation of total cost as decision maker gets more risk-averse, we examine the percent deviations from the risk-neutral case. In Figure 5.6, we represent the percent deviation of expected total cost and the standard deviation of total cost obtained by the risk-averse model from the one obtained by risk-neutral model for the case of 500 scenarios. Figure 5.6a shows these percent deviations for $\alpha = 0.7$ and $\lambda \in \{0.3, 0.5, 0.7, 1\}$. It is seen from Figure 5.6a that when λ decreases from 1 to 0.3, the decision maker gets more risk-averse and in order to decrease standard deviation, she increases the expected total cost. When $\lambda = 0.3$, in order to decrease the standard deviation by 1.65 %, an increase of 0.12% is observed on the expected total cost. Note that the percent gain on standard deviation is much more than the percent increase of the expected total cost.

A similar behavior is obtained for the cases of $\alpha = 0.8$ and $\alpha = 0.9$ (see Figures 5.6b and 5.6c). Note that, as α increases, in other words, as the decision maker gets more risk-averse, the gain on the standard deviation increases while the amount of increment in total cost also increases.

In Figure 5.6d, the percent deviations of optimal expected total cost and the standard deviation of optimal total cost obtained by risk-averse model from the ones of risk-neutral model for $\lambda = 0.3$ and $\alpha \in \{0.7, 0.8, 0.9\}$ are presented. It is

seen that as α increases, the decision maker gets more risk-averse and expected total cost increases in order to reduce the standard deviation. When $\alpha = 0.7$, the amount of increase on the expected total cost is 0.12 %, while the amount of increase for $\alpha = 0.9$ is 0.33 %. Note that the percent gain on the standard deviation is much more than the percent increase of the expected total cost.

A similar behavior is observed for values of λ set to 0.5 and 0.7. See Figures 5.6e and 5.6f. As λ increases, the gain on the standard deviation decreases and the increase on the expected total cost also decreases.

The percent deviation of expected total cost and the standard deviation of total cost obtained by the risk-averse model from the one obtained by risk-neutral model for the the number of scenarios larger than 500 are similar. See Appendix B.

When we compare the numerical results that are obtained using 500 scenarios and 1000 scenarios together, it is seen that there is a significant change in the optimal expected total cost values. Table 5.4 presents the expected total cost values for the number of scenarios 500 and 1000.

α	λ	Exp Total Cost for 500 scens (\$)	Exp Total Cost for 1000 scens (\$)	% Deviation
0.7	0.3	137,846,191	138,381,928	0.388648
0.7	0.5	137,769,407	138,296,761	0.382780
0.7	0.7	137,713,413	138,242,384	0.384110
0.8	0.3	137,972,867	138,501,411	0.383078
0.8	0.5	137,808,872	138,340,612	0.385853
0.8	0.7	137,737,920	138,267,555	0.384524
0.9	0.3	138,132,920	138,665,944	0.385877
0.9	0.5	137,862,049	138,395,841	0.387193
0.9	0.7	137,751,508	138,283,308	0.386057

Table 5.4: Comparison of Total Expected Costs (\$) for 500 and 1000 scenarios

Similarly, Table 5.5 and Table 5.6 represent the expected total cost values for the number of scenarios 500 and 700, and for the number of scenarios 700 and 1000. It is seen that the percent deviation in Table 5.6 is less than the percent deviation in Table 5.5.

α	λ	Exp Total Cost for 500 scens (\$)	Exp Total Cost for 700 scens (\$)	% Deviation
0.7	0.3	137,846,191	138,357,491	0.370921
0.7	0.5	137,769,407	138,281,825	0.371939
0.7	0.7	137,713,413	138,227,870	0.373571
0.8	0.3	137,972,867	138,476,543	0.365054
0.8	0.5	137,808,872	138,322,720	0.372870
0.8	0.7	137,737,920	138,251,487	0.372858
0.9	0.3	138,132,920	138,643,700	0.369774
0.9	0.5	137,862,049	138,378,854	0.374871
0.9	0.7	137,751,508	138,266,434	0.373807

Table 5.5: Comparison of Total Expected Costs (\$) for 500 and 700 scenarios

α	λ	Exp Total Cost for 700 scens (\$)	Exp Total Cost for 1000 scens (\$)	% Deviation
0.7	0.3	137,846,191	138,381,928	0.017662
0.7	0.5	137,769,407	138,296,761	0.010802
0.7	0.7	137,713,413	138,242,384	0.010500
0.8	0.3	137,972,867	138,501,411	0.017958
0.8	0.5	137,808,872	138,340,612	0.012936
0.8	0.7	137,737,920	138,267,555	0.011622
0.9	0.3	138,132,920	138,665,944	0.016044
0.9	0.5	137,862,049	138,395,841	0.012276
0.9	0.7	137,751,508	138,283,308	0.012205

Table 5.6: Comparison of Total Expected Costs (\$) for 700 and 1000 scenarios

Tables 5.7 - 5.9 represent the numerical results consisting of objective function value, first-stage cost, second-stage cost, expected total cost, CVaR of mismatched demand, CVaR of total cost, cost of reservoirs, power generators, and solar panels and transmission cost for different α and λ values.

Values	$\lambda = 0.3$	$\lambda = 0.5$	$\lambda = 0.7$
Objective Function Value	157,447,452	151,832,734	146,193,785
First Stage Cost	41,047,697	40,782,161	40,477,253
Second Stage Cost	96,798,495	96,987,246	97,236,159
Expected Total Cost	137,846,191	137,769,407	137,713,413
CVaR of Mismatched Demand	124,800,295	125,113,900	125,504,066
CVaR of Total Cost	165,847,992	165,896,061	165,981,319
Cost of Reservoirs	21,043,890	21,615,315	23,095,884
Cost of Power Generators	69,349,574	70,400,095	71,538,732
Cost of Solar Panels	550,386,835	544,888,972	537,968,783
Transmission Cost	7,628,453	7,744,010	7,869,260

Table 5.7: Numerical Results for $\alpha = 0.7$ and $\lambda \in \{0.3, 0.5, 0.7\}$

Values	$\lambda = 0.3$	$\lambda = 0.5$	$\lambda = 0.7$
Objective Function Value	160,868,909	154,299,036	147,685,477
First Stage Cost	41,084,775	40,801,412	40,513,186
Second Stage Cost	96,888,092	97,007,460	97,224,734
Expected Total Cost	137,972,867	137,808,872	137,737,920
CVaR of Mismatched Demand	129,596,723	129,987,787	13,038,3257
CVaR of Total Cost	170,681,498	170,789,199	170,896,443
Cost of Reservoirs	17,902,034	19,563,480	21,363,580
Cost of Power Generators	68,185,689	70,005,389	70,962,232
Cost of Solar Panels	554,563,265	547,207,854	54,0450,119
Transmission Cost	7,500,426	7,700,593	7,805,845

Table 5.8: Numerical Results for $\alpha = 0.8$ and $\lambda \in \{0.3, 0.5, 0.7\}$

Values	$\lambda = 0.3$	$\lambda = 0.5$	$\lambda = 0.7$
Objective Function Value	163,752,724	156,385,930	148,949,323
First Stage Cost	41,093,082	40,766,129	40,527,731
Second Stage Cost	97,039,838	97,095,920	97,223,778
Expected Total Cost	138,132,920	137,862,049	137,751,508
CVaR of Mismatched Demand	133,639,558	134,143,682	134,549,827
CVaR of Total Cost	174,732,640	174,909,811	175,077,558
Cost of Reservoirs	17,697,677	18,633,503	20,883,618
Cost of Power Generators	66,533,700	69,113,674	70,608,863
Cost of Solar Panels	556,359,219	548,231,207	541,384,225
Transmission Cost	7,318,707	7,602,504	7,766,975

Table 5.9: Numerical Results for $\alpha = 0.9$ and $\lambda \in \{0.3, 0.5, 0.7\}$

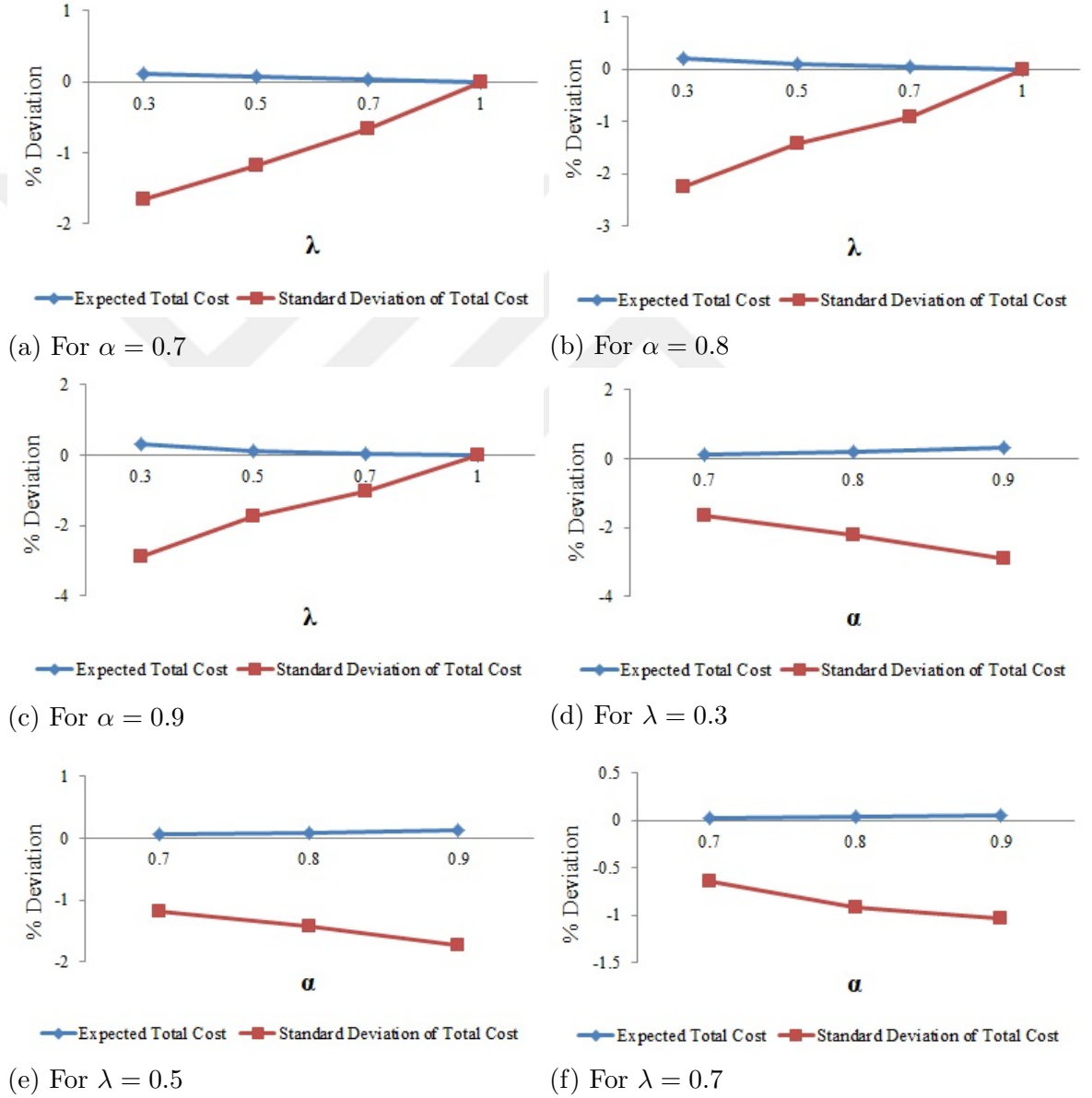


Figure 5.6: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenarios

Chapter 6

Conclusion

In this thesis, we consider a risk-averse approach for a hybrid energy system planning problem. The considered hybrid energy system consists of solar panels, conventional hydropower stations, transmission lines, and diesel generators. We model this energy system planning problem as a risk-averse two-stage stochastic program where the uncertainty is the amount of water inflow to the reservoirs. The objective is to minimize the risk adjusted cost while satisfying the demand for energy.

In classic two-stage stochastic optimization models, the aim is to optimize the total expected value of random outcomes. This approach is based on an implicit assumption that the decision makers are risk-neutral. However, in a cost minimization problem setting where there is need for large amount of investments, a risk-averse approach will be more realistic. In this context, we develop a risk-averse model by involving CVaR as the measure of risk in the objective function.

In order to solve our model, we propose a multi-cut scenario-wise decomposition algorithm that is based on the L-shaped method. This exact solution method is based on decomposing a single large sized problem into several small sized problems and therefore it decreases the solution time considerably. In stochastic optimization problems, the number of scenarios is crucial in order to model the

uncertainty involved in the problem. However, considering a large number of scenarios will increase the problem sizes. Our proposed decomposition algorithm, in this context, allows us to consider large number of scenarios by decreasing the solution time considerably.

We present a computational study for the Mediterranean Region in Turkey. In order to generate scenarios, we use a modified k -NN algorithm proposed by Prairie et al. [4]. We provide our numerical results for different number of scenarios ranging from 500 to 1000. We examine the changes in the decisions for different risk-aversion characteristics. Our results show that as the risk-aversion increases, the investments that are directly related to the source of the uncertainties are reduced. We also state that as the risk-aversion increases, in order to reduce the standard deviation of the total cost, the decision makers start to compromise on the expected total costs. We evaluate the performance of our solution method by solving the extensive formulation with CPLEX and comparing the CPU times required to solve the model. The results show that our proposed algorithm requires less CPU time compared to CPLEX.

For possible future research, we can consider involving the uncertainties in the solar radiation and the demand data. In this way, involvement of different types of uncertainty and their effects on the decisions taken under different risk-aversion characteristics can be examined. With the involvement of additional uncertainties, different types of scenario generation techniques can also be used.

Bibliography

- [1] “Key world energy statistics.” <https://www.iea.org/publications/freepublications/publication/key-world-energy-statistics.html>. Accessed: August, 2017.
- [2] “International energy agency.” <http://www.iea.org/publications/freepublications/publication/CO2-emissions-from-fuel-combustion-highlights-2016.html>. Accessed: August, 2017.
- [3] A. S. Kocaman, C. Abad, T. J. Troy, W. T. Huh, and V. Modi, “A stochastic model for a macroscale hybrid renewable energy system,” *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 688 – 703, 2016.
- [4] J. R. Prairie, B. Rajagopalan, T. J. Fulp, E. A. Zagona, and M. ASCE, “Modified k -NN model for stochastic streamflow simulation,” *Journal of Hydrologic Engineering*, vol. 11, no. 4, pp. 371 – 378, 2006.
- [5] R. V. Slyke and R. Wets, “L-shaped linear programs with applications to optimal control and stochastic programming,” *SIAM Journal on Applied Mathematics*, vol. 17, no. 4, pp. 638 – 663, 1969.
- [6] S. Ashok, “Optimised model for community-based hybrid energy system,” *Renewable Energy*, vol. 32, no. 7, pp. 1155 – 1164, 2007.
- [7] O. Ekren, B. Y. Ekren, and B. Ozerdem, “Break-even analysis and size optimization of a pv/wind hybrid energy conversion system with battery storage a case study,” *Applied Energy*, vol. 87, no. 7, pp. 1043 – 1054, 2009.

- [8] O. Ekren and B. Y. Ekren, "Size optimization of a pv/wind hybrid energy conversion system with battery storage using simulated annealing," *Applied Energy*, vol. 87, no. 2, pp. 592 – 598, 2010.
- [9] T. Senjyu, D. Hayashi, A. Yona, N. Urasaki, and T. Funabashi, "Optimal configuration of power generating systems in isolated island with renewable energy," *Renewable Energy*, vol. 32, no. 11, pp. 1917 – 1933, 2007.
- [10] L. Kuznia, B. Zeng, G. Centeno, and Z. Miao, "Stochastic optimization for power system configuration with renewable energy in remote areas," *Annals of Operations Research*, vol. 210, no. 1, pp. 411 – 432, 2013.
- [11] B. Sun, P. Krokhmal, and Y. Chen, "Risk-averse capacity planning for renewable energy production," *Energy Systems*, 2017.
- [12] Y. Merzifonluoglu and E. Uzgoren, "Photovoltaic power plant design considering multiple uncertainties and risk," *Annals of Operations Research*, 2017.
- [13] G. B. Dantzig, "Linear programming under uncertainty," *Management Science*, vol. 1, no. 3-4, pp. 197 – 206, 1955.
- [14] E. Beale, "On minimizing a convex function subject to linear inequalities," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 17, no. 2, pp. 173 – 184, 1955.
- [15] S. Ahmed, "Mean-risk objectives in stochastic programming," *Technical report, Georgia Institute of Technology*, 2004.
- [16] J. Linderoth, A. Shapiro, and S. Wright, "The empirical behavior of sampling methods for stochastic programming," *Annals of Operations Research*, vol. 142, no. 1, pp. 215 – 241, 2006.
- [17] S. Ahmed, "Convexity and decomposition of mean-risk stochastic programs," *Mathematical Programming*, vol. 106, no. 3, pp. 433 – 446, 2006.
- [18] T. K. Kristoffersen, "Deviation measures in linear two-stage stochastic programming," *Mathematical Methods of Operations Research*, vol. 62, no. 2, pp. 255 – 274, 2005.

- [19] R. Schultz and S. Tiedemann, “Conditional value-at-risk in stochastic programs with mixed-integer recourse,” *Mathematical Programming*, vol. 105, no. 2-3, pp. 365 – 386, 2006.
- [20] C. I. Fabian, “Handling CVaR objectives and constraints in two-stage stochastic models,” *European Journal of Operational Research*, vol. 191, no. 3, pp. 888 – 911, 2008.
- [21] N. Miller and A. Ruszczyński, “Risk-averse two-stage stochastic linear programming: Modeling and decomposition,” *Operations Research*, vol. 59, no. 1, pp. 125 – 132, 2011.
- [22] N. Noyan, “Risk-averse two-stage stochastic programming with an application to disaster management,” *Computers & Operations Research*, vol. 39, no. 3, pp. 541 – 559, 2012.
- [23] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, “Coherent measures of risk,” *Computers & Operations Research*, vol. 9, no. 3, pp. 203 – 228, 1999.
- [24] R. T. Rockafellar and S. Uryasev, “Conditional value-at-risk for general loss distributions,” *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443 – 1471, 2002.
- [25] P. Jorion, *Value at risk the new benchmark for measuring derivatives risk*. Irwin Professional Publishing, 1996.
- [26] M. Pritsker, “Evaluating value at risk methodologies: Accuracy versus computational time,” *Journal of Financial Services Research*, vol. 12, no. 2-3, pp. 201 – 242, 1997.
- [27] W. Ogryczak and A. Ruszczyński, “Dual stochastic dominance and related mean-risk models,” *Tutorials in Operations Research*, vol. 13, no. 1, pp. 60 – 78, 2002.
- [28] J. R. Birge and F. V. Louveaux, “A multicut algorithm for two-stage stochastic linear programs,” *European Journal of Operational Research*, vol. 34, no. 3, pp. 384 – 392, 1988.

- [29] G. Laporte and F. V. Louveaux, “The integer L-shaped method for stochastic integer programs with complete recourse,” *Operations Research Letters*, vol. 13, no. 3, pp. 133 – 142, 1993.
- [30] U. Lall and A. Sharma, “A nearest neighbor bootstrap for resampling hydrologic time series,” *Water Resources Research*, vol. 32, no. 3, pp. 679 – 693, 1996.
- [31] “Republic of Turkey Energy Market Regulatory.” <http://www.epdk.org.tr>. Accessed: August, 2017.
- [32] “Homer energy.” <http://www.homerenergy.com/>. Accessed: August, 2017.
- [33] “General directorate of state hydraulic works of turkey.” <http://www.dsi.gov.tr/>. Accessed: August, 2017.
- [34] G. Jekabsons, “Locally weighted polynomials toolbox for matlab/octave, 2016.” <http://www.cs.rtu.lv/jekabsons/>. Accessed: August, 2017.
- [35] P. Craven and G. Wahba, “Smoothing noisy data with spline functions,” *Numerische Mathematik*, vol. 31, no. 4, pp. 377 – 403, 1979.
- [36] N. Singhrattana, B. Rajagopalan, M. Clark, and K. K. Kumar, “Seasonal forecasting of thailand summer monsoon rainfall,” *International Journal of Climatology*, vol. 25, no. 5, pp. 649 – 664, 2005.
- [37] C. Loader, *Local Regression and Likelihood*. Springer, New York, NY, 1999.
- [38] B. Rajagopalan and U. Lall, “Locally weighted polynomial estimation of spatial precipitation,” *Journal of Geographic Information and Decision Analysis*, vol. 2, no. 3, pp. 48 – 57, 1998.

Appendix A

**Numerical Results for Optimal
First-Stage Decisions for the
number of Scenarios 600, 700,
800, 900 and 1000**

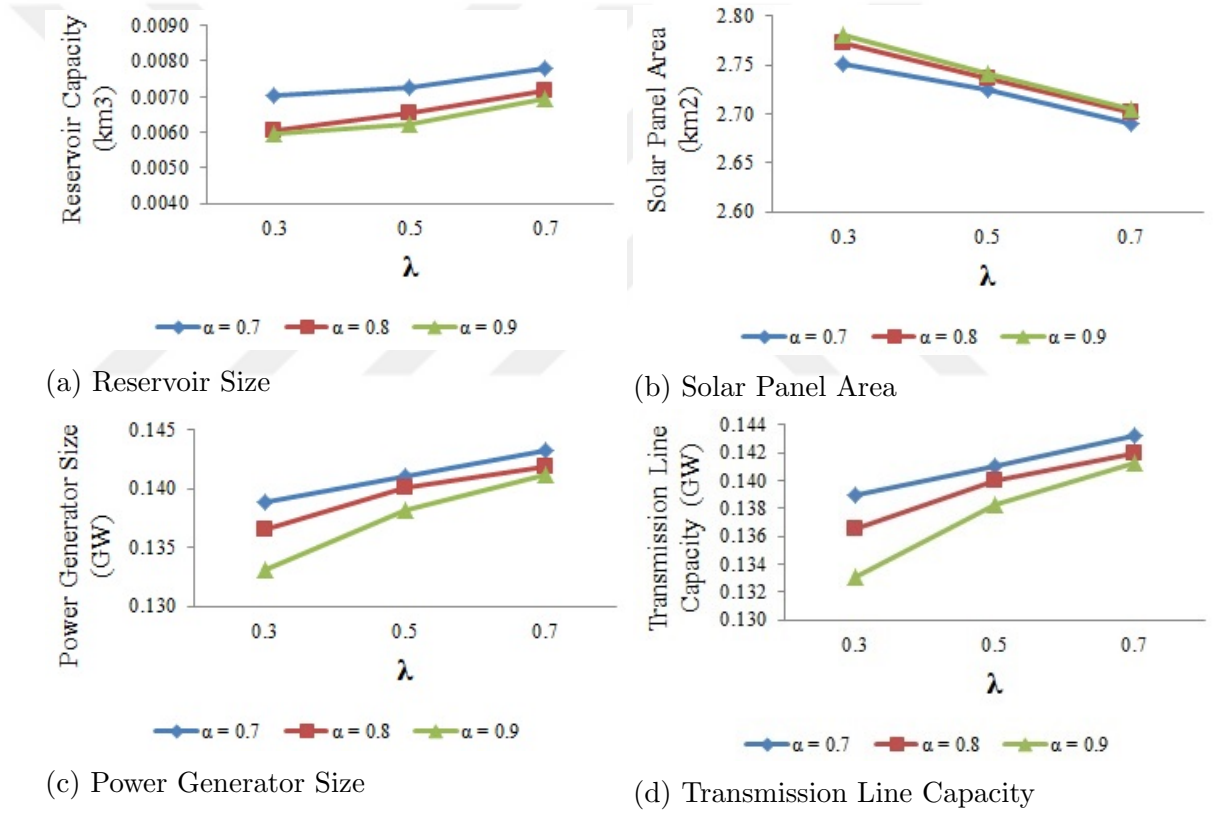


Figure A.1: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case

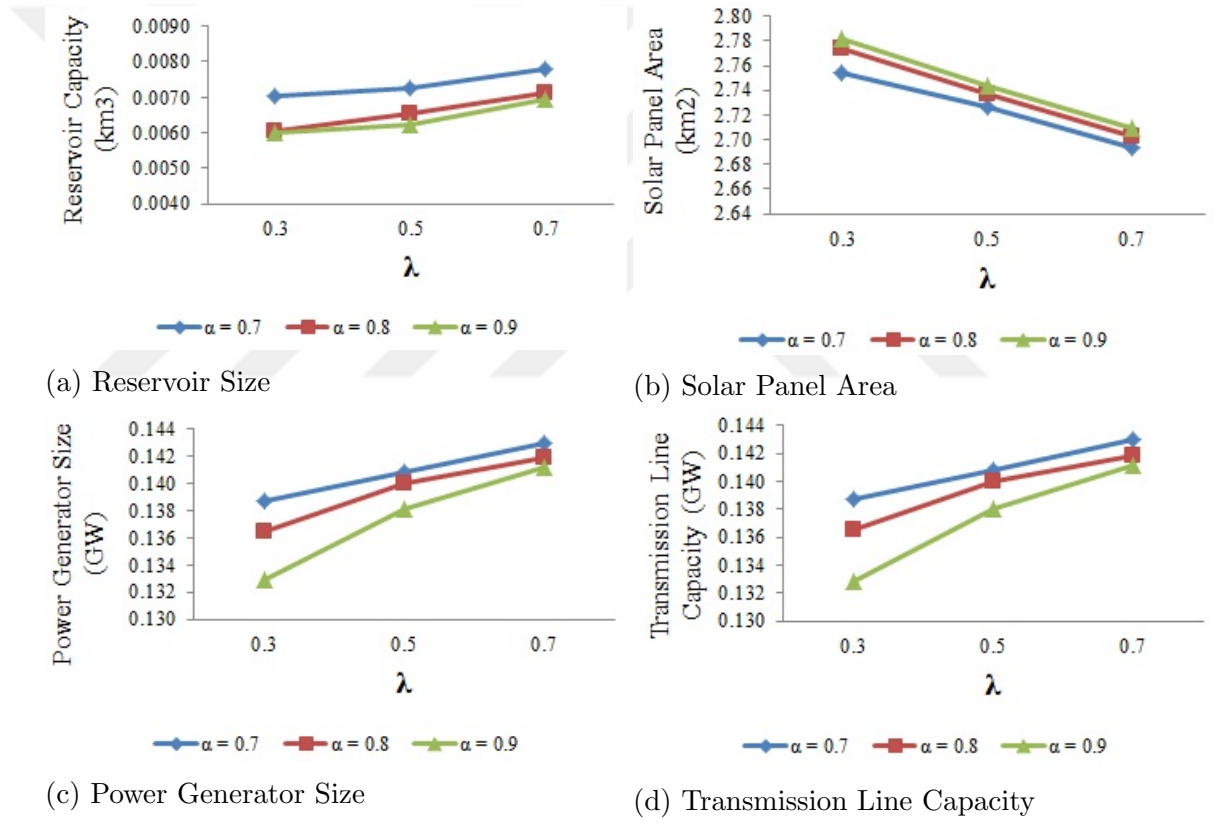


Figure A.2: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case

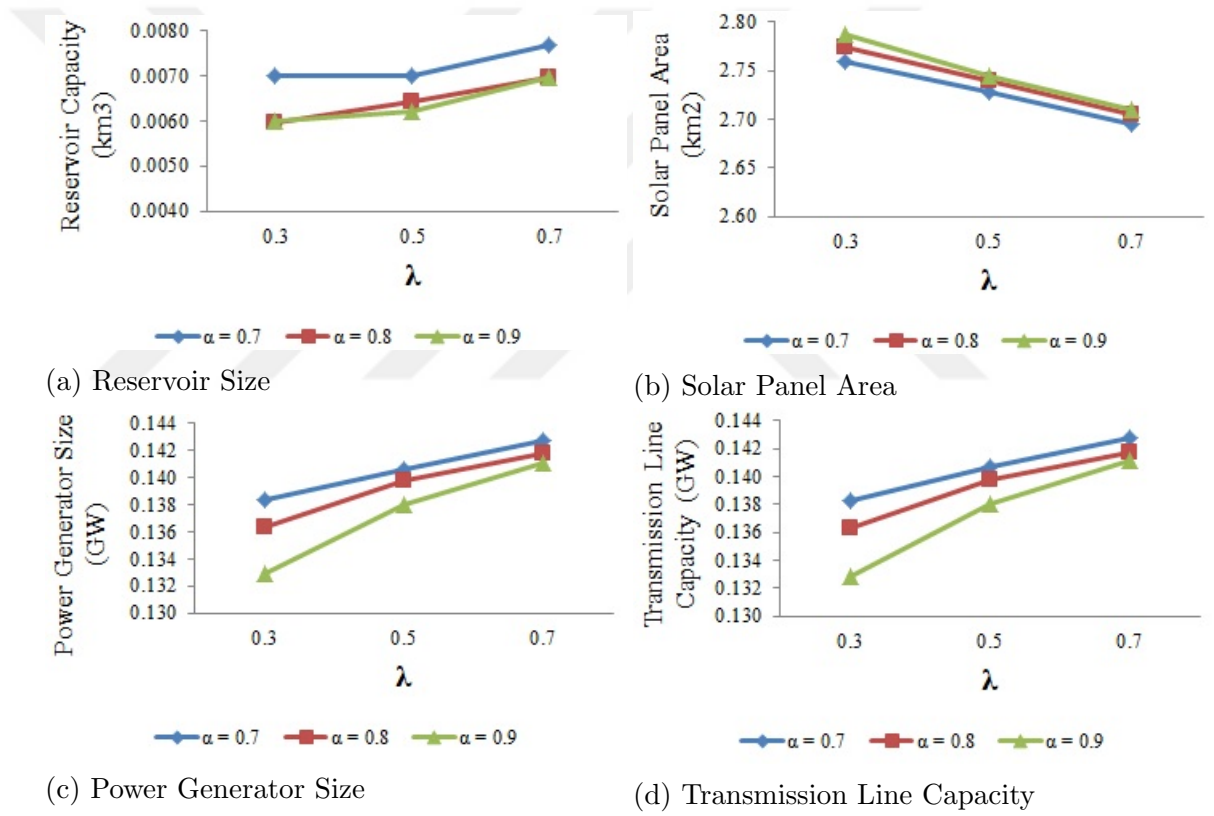


Figure A.3: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case

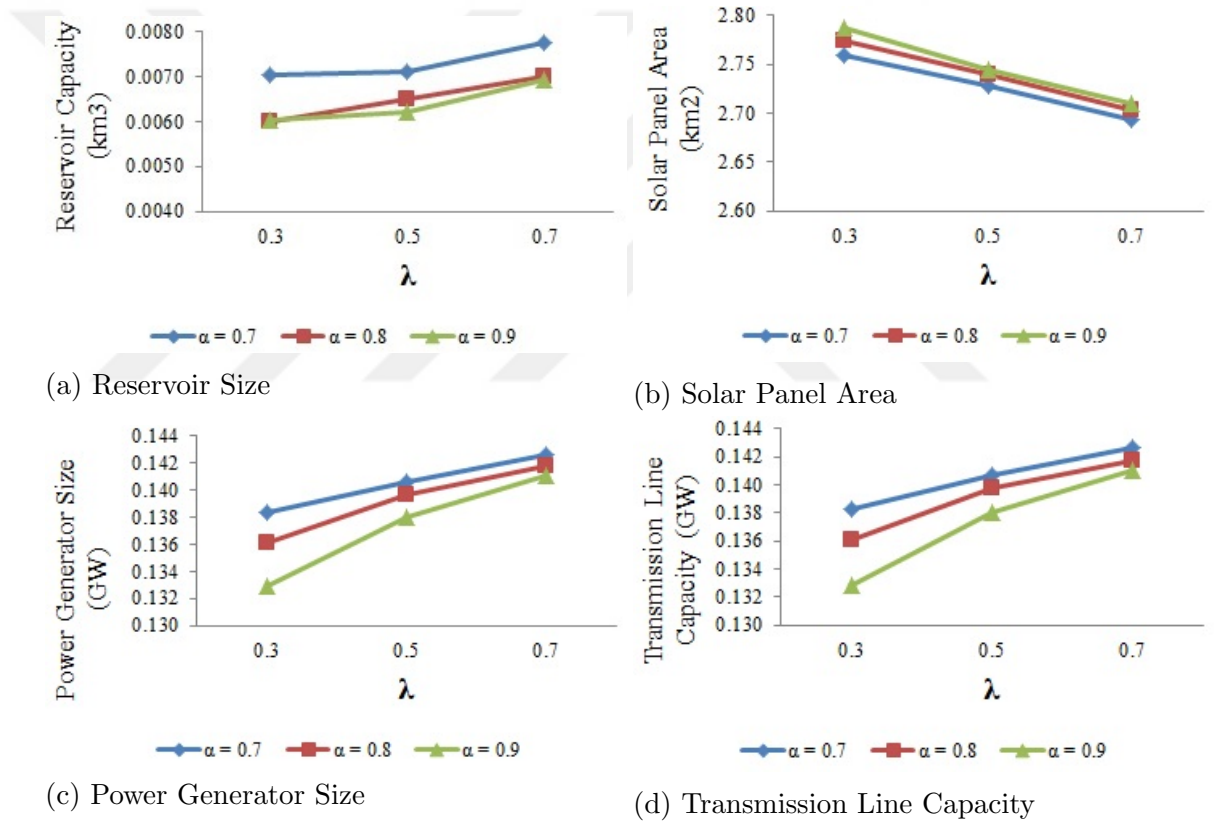
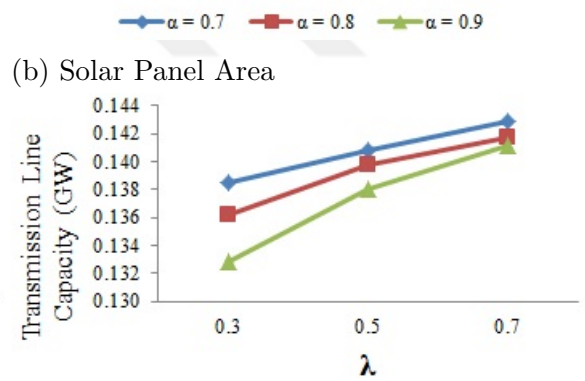
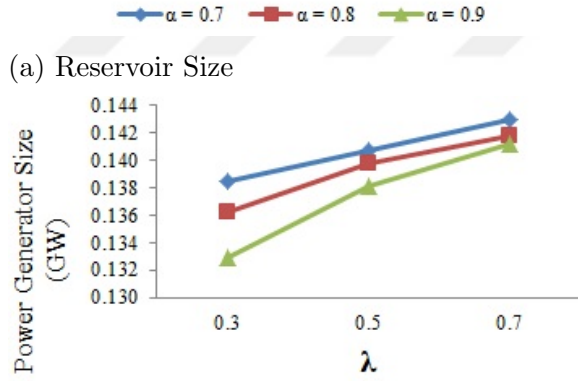
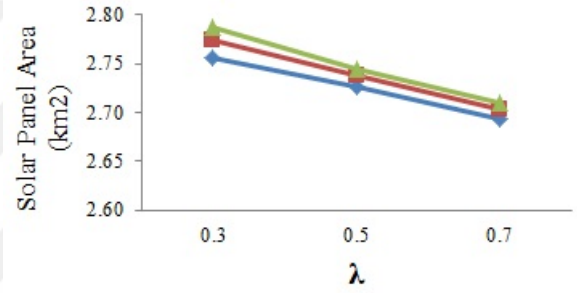
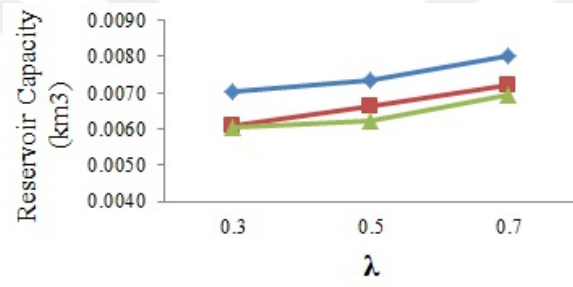


Figure A.4: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case



(a) Reservoir Size

(b) Solar Panel Area

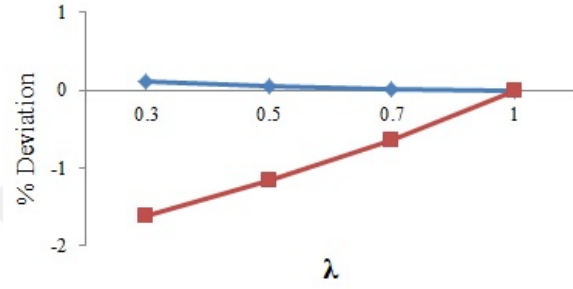
(c) Power Generator Size

(d) Transmission Line Capacity

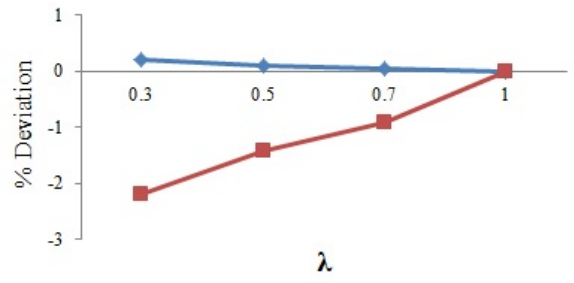
Figure A.5: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case

Appendix B

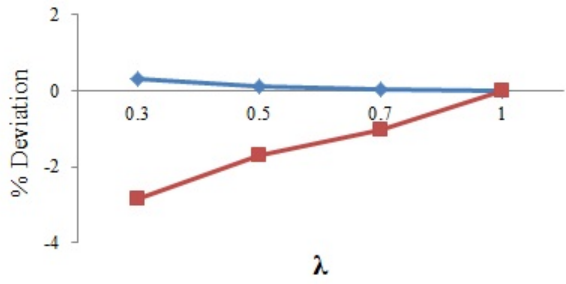
Numerical Results for the
Percent Deviations from Risk
Neutral Case for the number of
scenarios 600, 700, 800, 900 and
1000



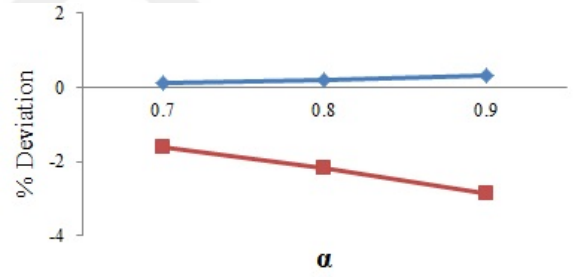
(a) For $\alpha = 0.7$



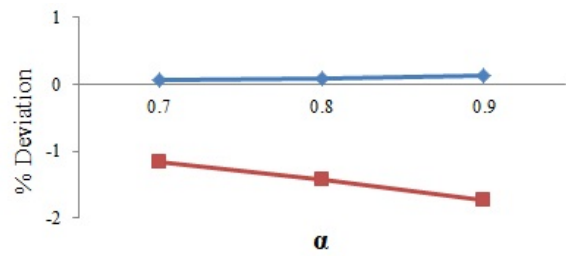
(b) For $\alpha = 0.8$



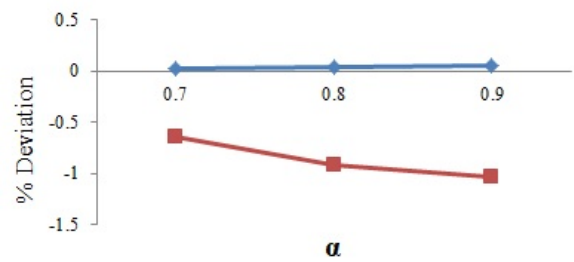
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

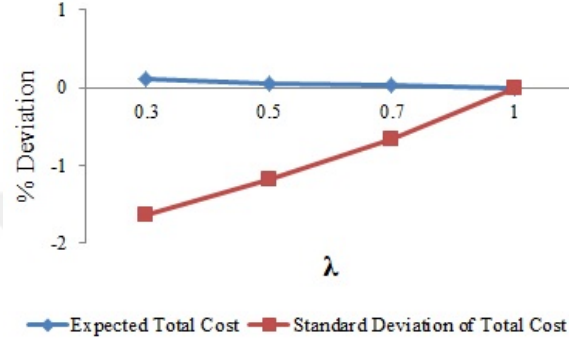


(e) For $\lambda = 0.5$

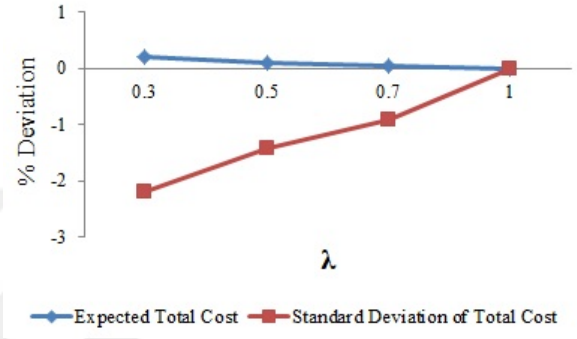


(f) For $\lambda = 0.7$

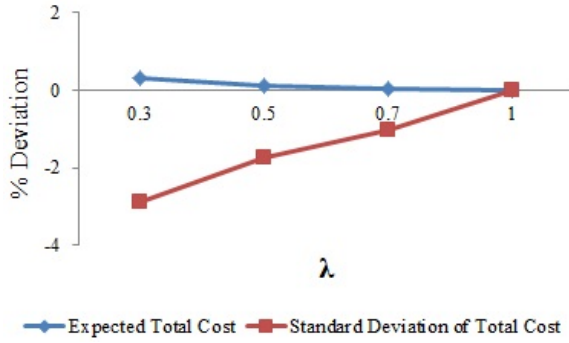
Figure B.1: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case



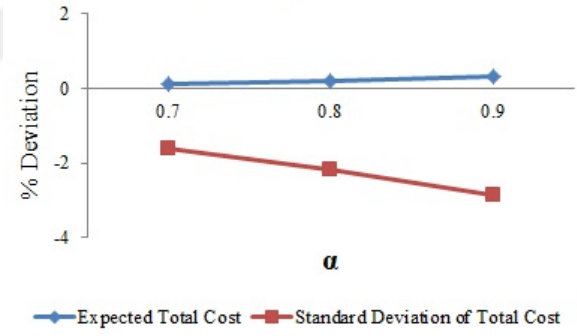
(a) For $\alpha = 0.7$



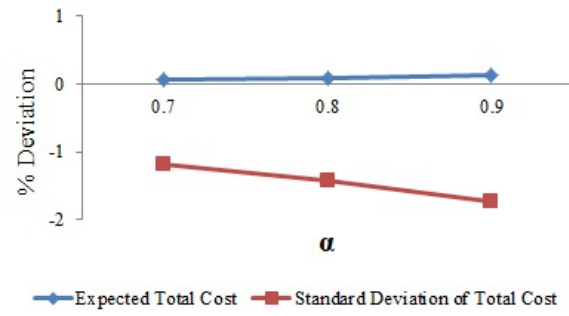
(b) For $\alpha = 0.8$



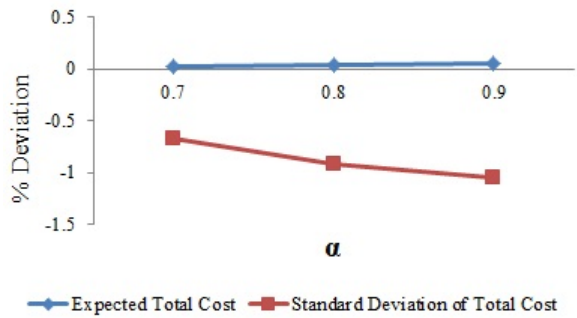
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

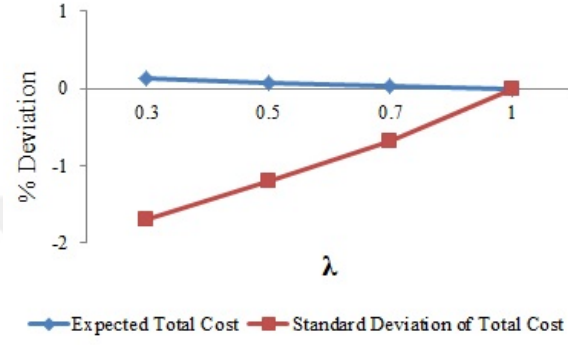


(e) For $\lambda = 0.5$

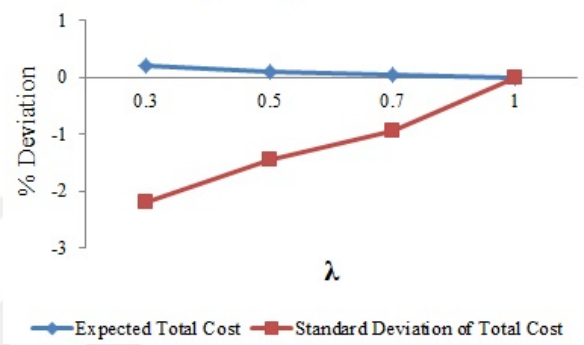


(f) For $\lambda = 0.7$

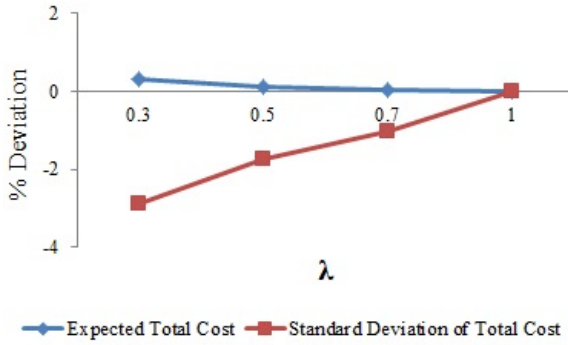
Figure B.2: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case



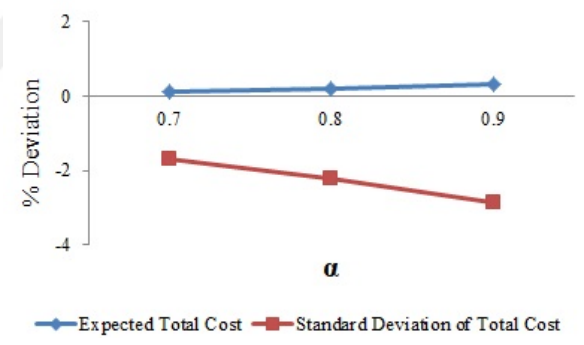
(a) For $\alpha = 0.7$



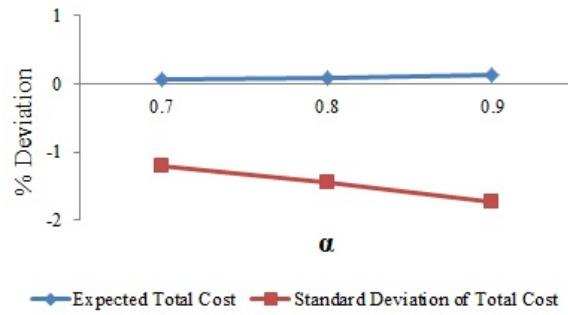
(b) For $\alpha = 0.8$



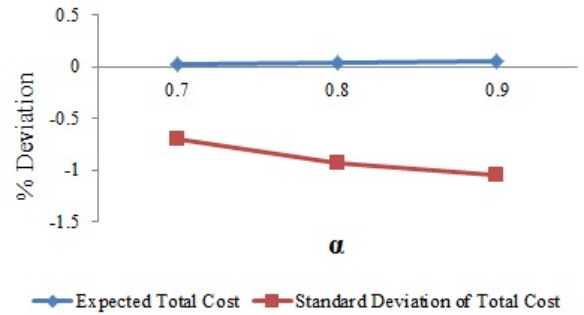
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

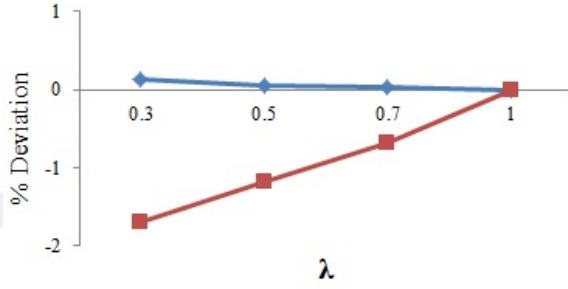


(e) For $\lambda = 0.5$

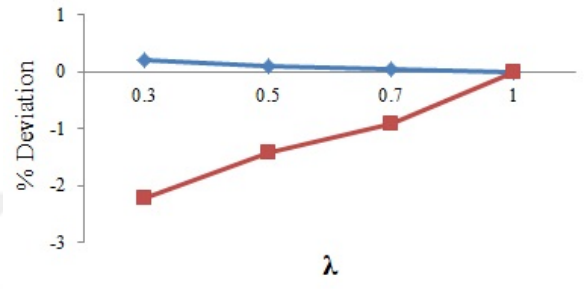


(f) For $\lambda = 0.7$

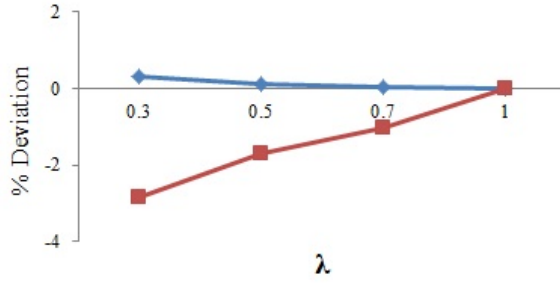
Figure B.3: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case



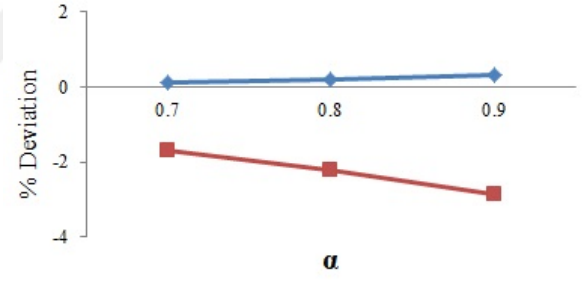
(a) For $\alpha = 0.7$



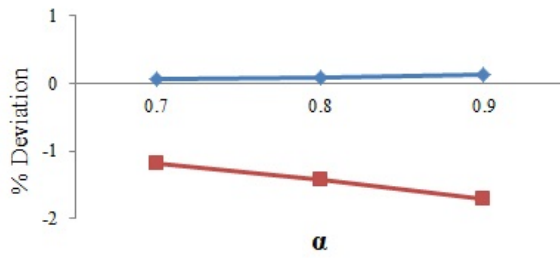
(b) For $\alpha = 0.8$



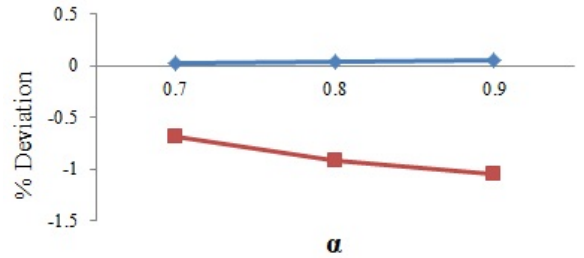
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

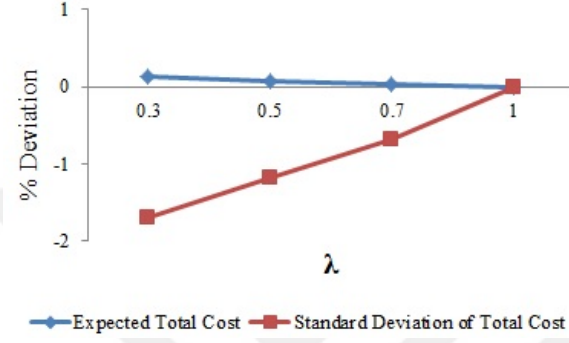


(e) For $\lambda = 0.5$

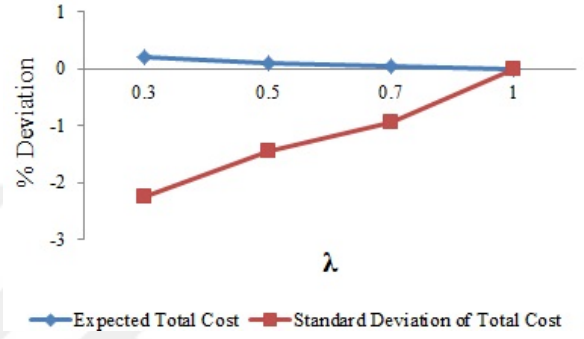


(f) For $\lambda = 0.7$

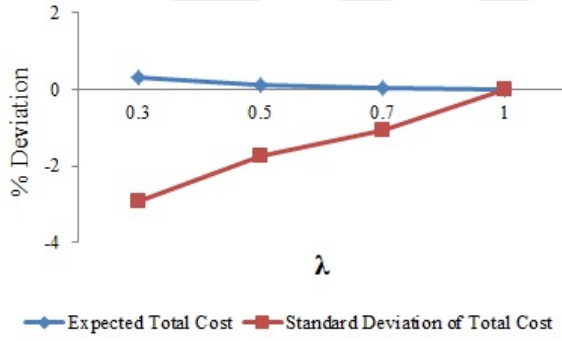
Figure B.4: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case



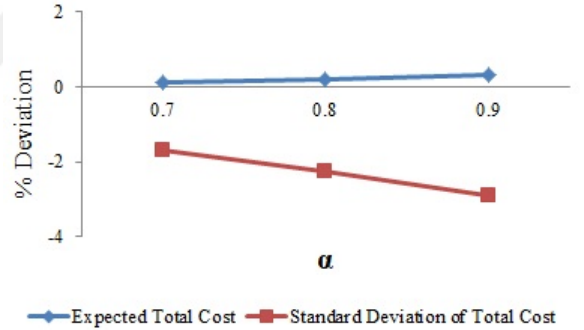
(a) For $\alpha = 0.7$



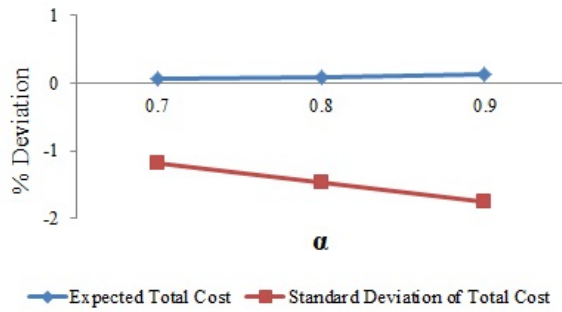
(b) For $\alpha = 0.8$



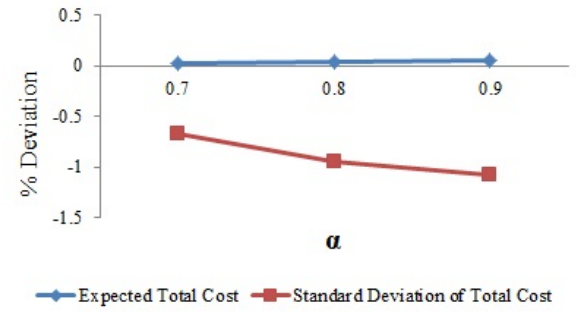
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$



(e) For $\lambda = 0.5$



(f) For $\lambda = 0.7$

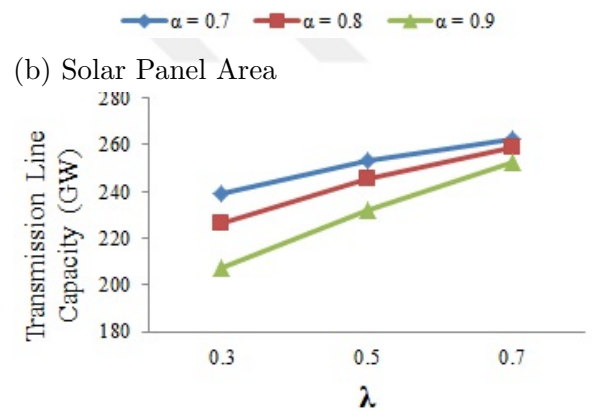
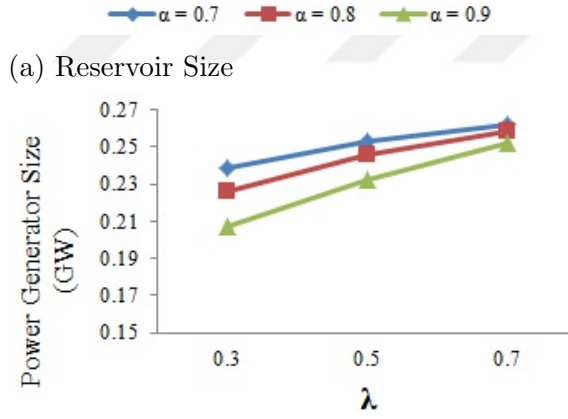
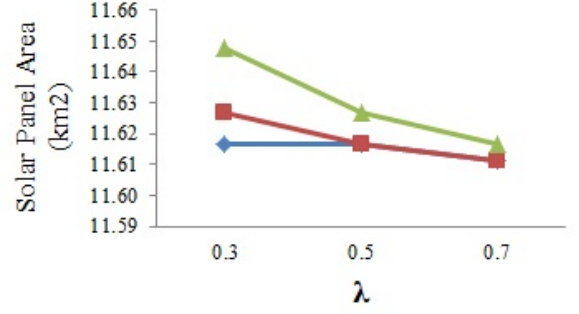
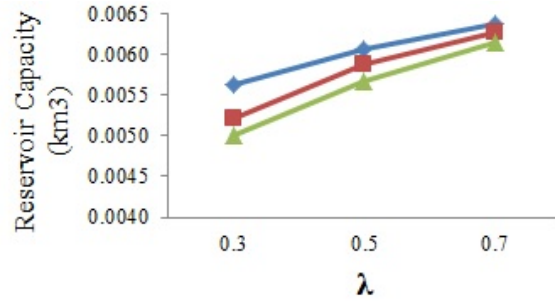
Figure B.5: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case

Appendix C

Numerical Results with Real Demand Data

		Extensive Formula- tion (CPLEX)	Scenraio-wise Decom- position Algorithm
α	λ	CPU time (s)	CPU time (s)
0.7	0.3	NA	13,223.15
0.7	0.5	NA	14,484.33
0.7	0.7	NA	12,291.05
0.8	0.3	168151.31	14,766.80
0.8	0.5	100,407.49	17,291.46
0.8	0.7	NA	18,182.12
0.9	0.3	141,225.94	14,371.83
0.9	0.5	132,831.74	15,916.61
0.9	0.7	NA	18,768.85

Table C.1: Comparison of CPU times for the algorithm and Extensive Formulation for 500 scenarios



(a) Reservoir Size

(b) Solar Panel Area

Figure C.1: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenario Case

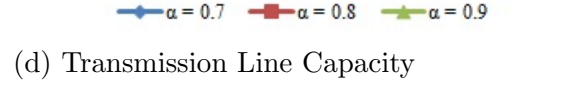
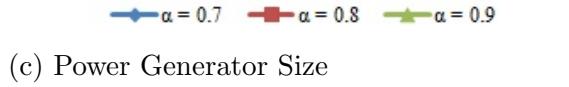
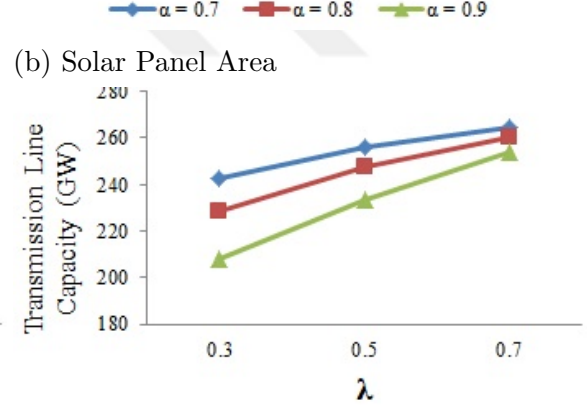
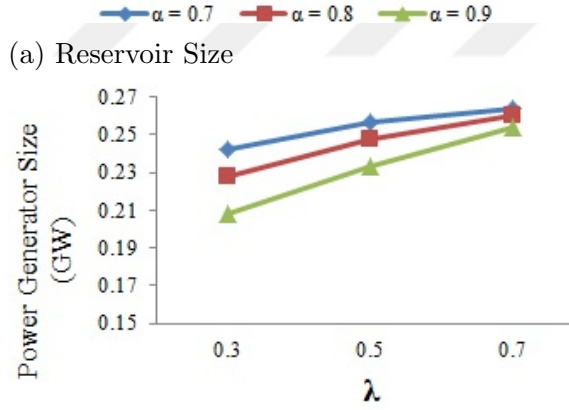
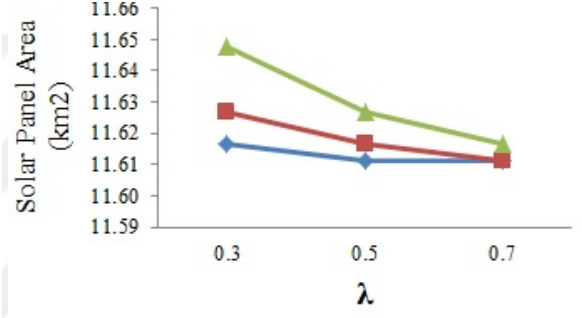
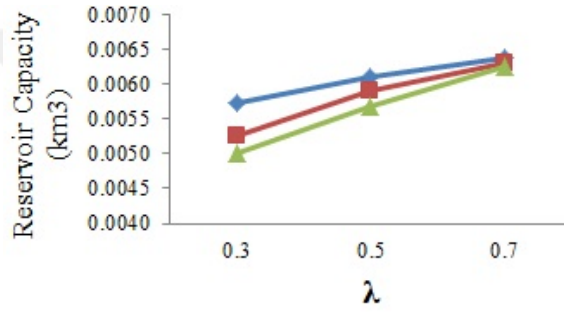
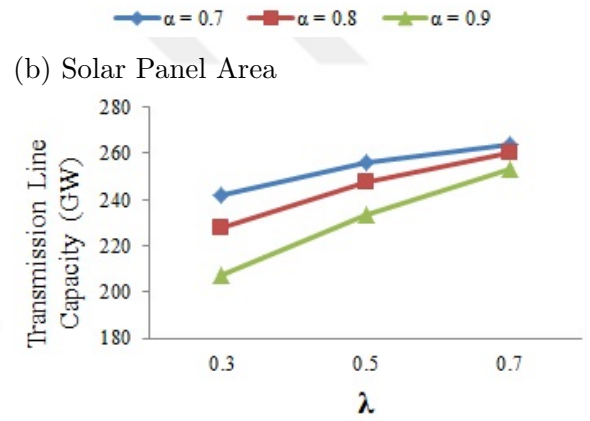
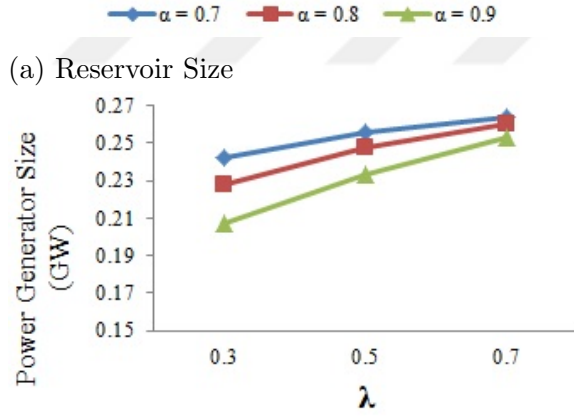
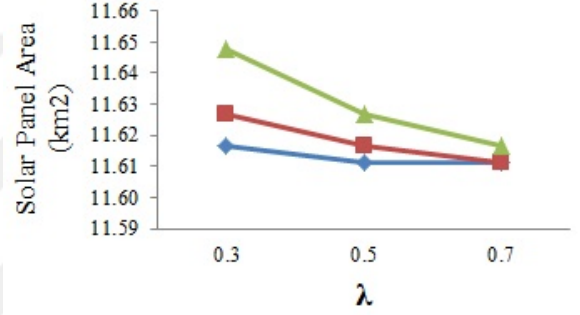
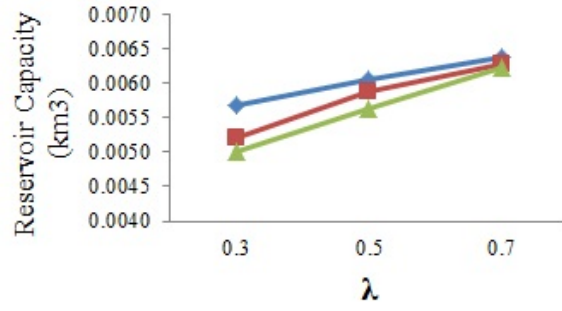


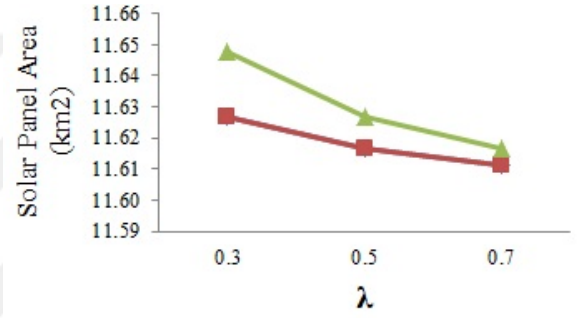
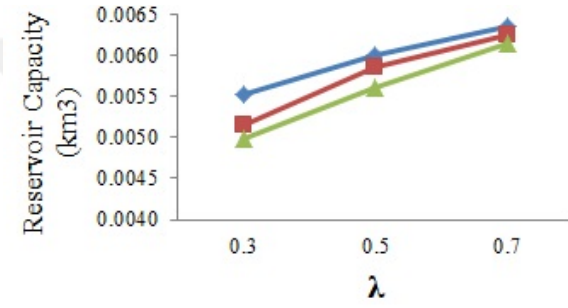
Figure C.2: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case



(a) Reservoir Size

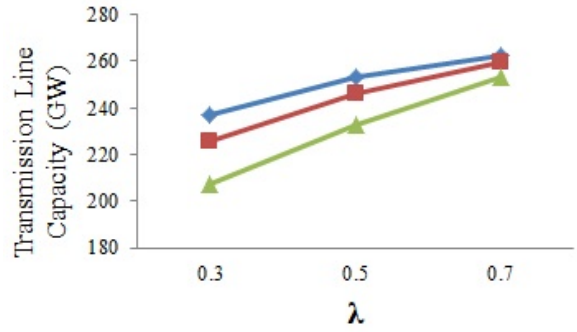
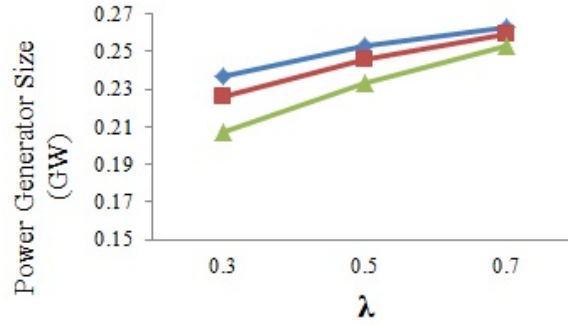
(b) Solar Panel Area

Figure C.3: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case



(a) Reservoir Size

(b) Solar Panel Area



(c) Power Generator Size

(d) Transmission Line Capacity

Figure C.4: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case

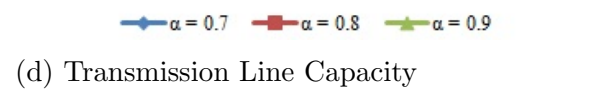
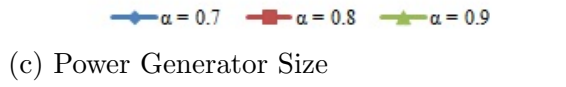
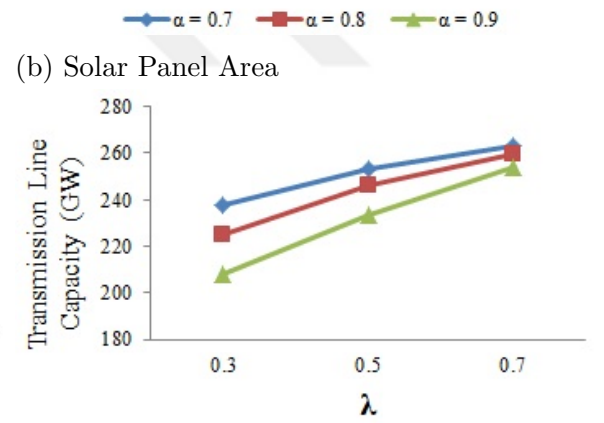
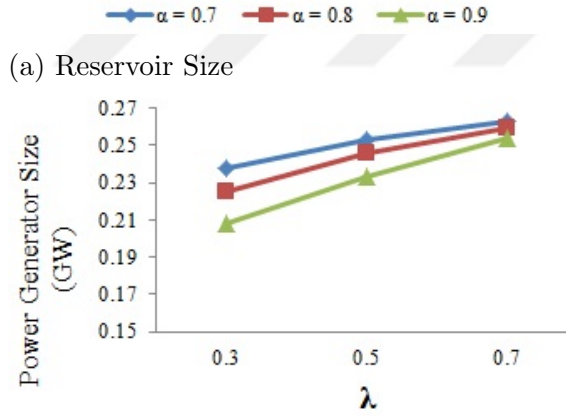
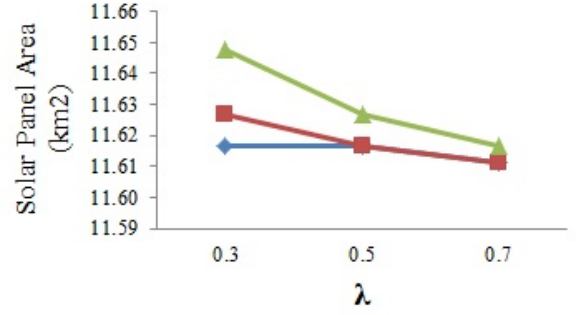
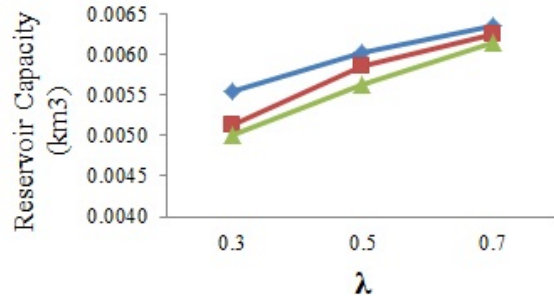


Figure C.5: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case

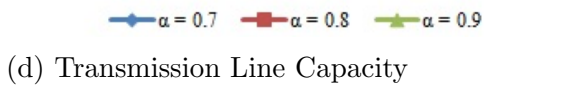
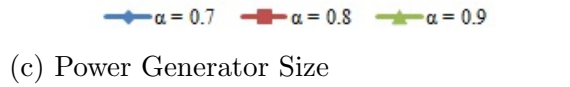
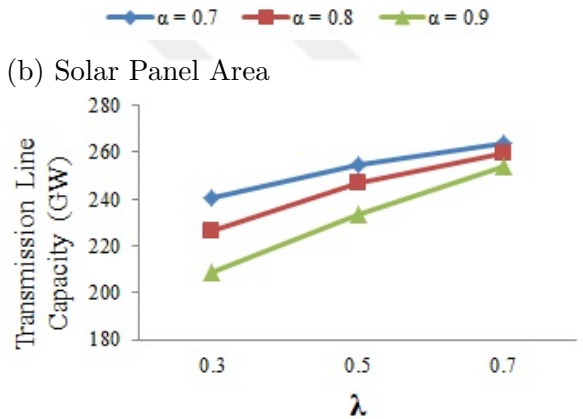
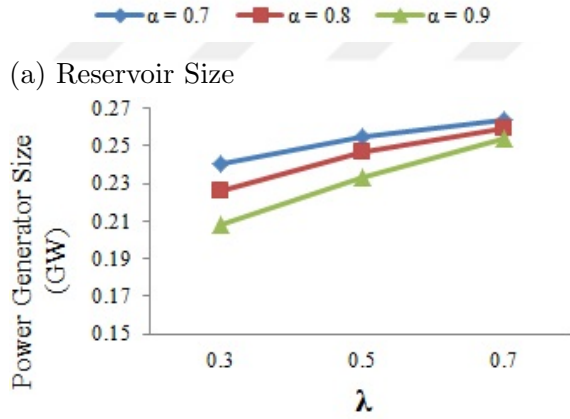
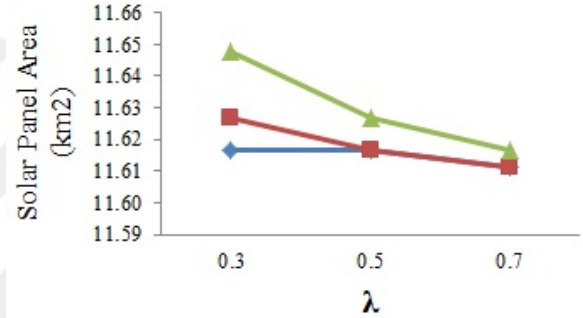
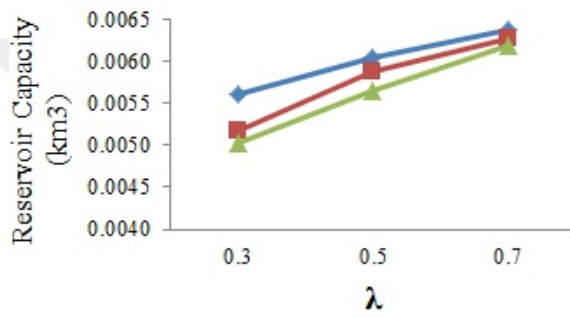
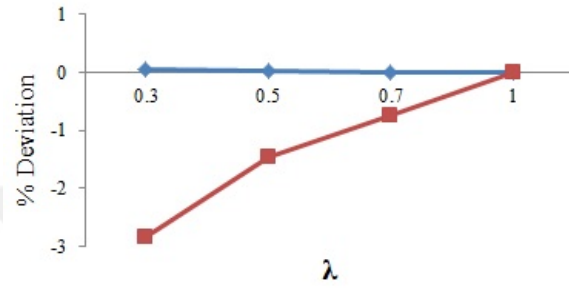
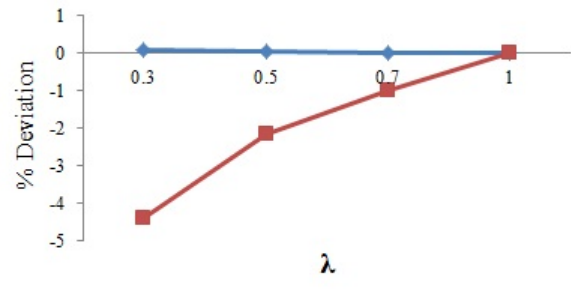


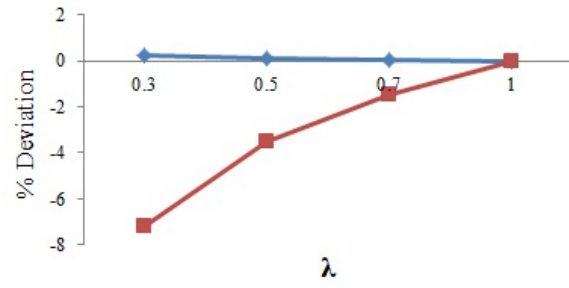
Figure C.6: Optimal First-Stage Decision Variables for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case



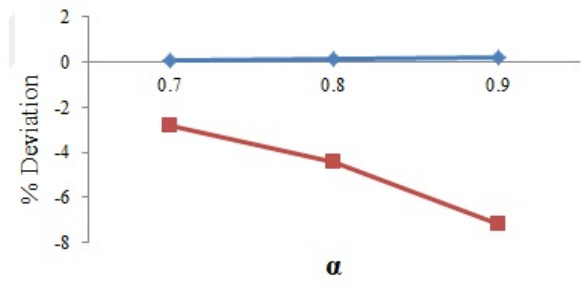
(a) For $\alpha = 0.7$



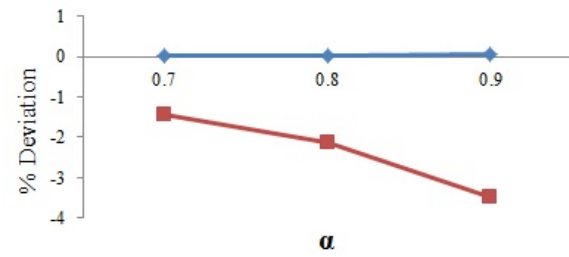
(b) For $\alpha = 0.8$



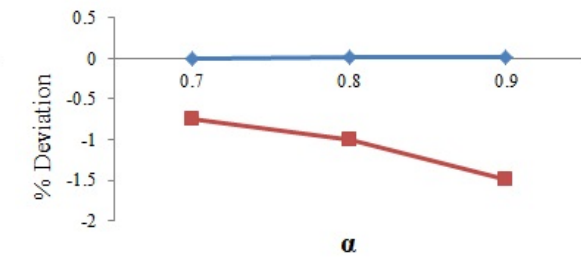
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

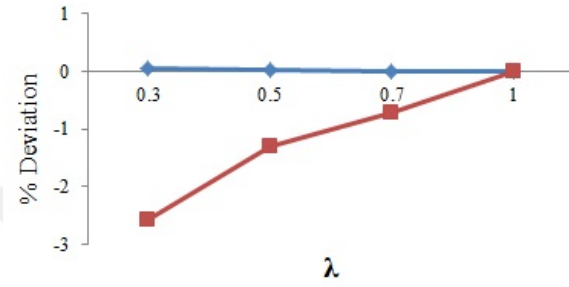


(e) For $\lambda = 0.5$

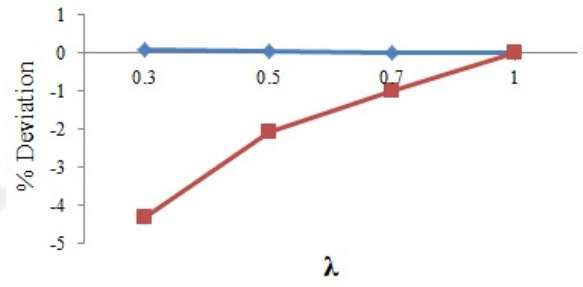


(f) For $\lambda = 0.7$

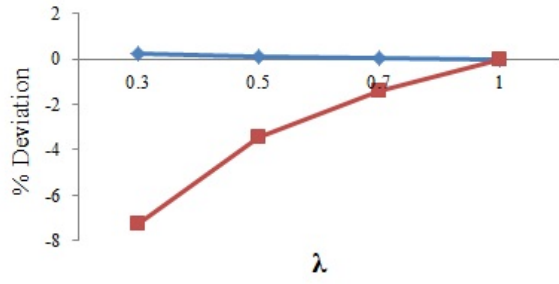
Figure C.7: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 500 Scenario Case



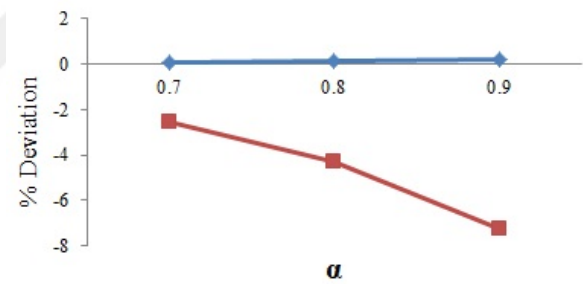
(a) For $\alpha = 0.7$



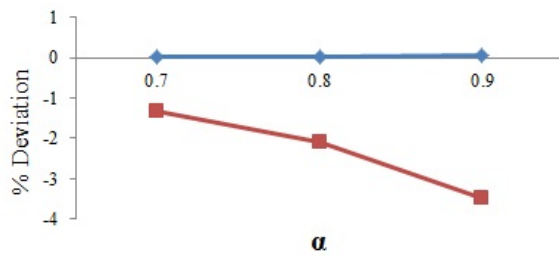
(b) For $\alpha = 0.8$



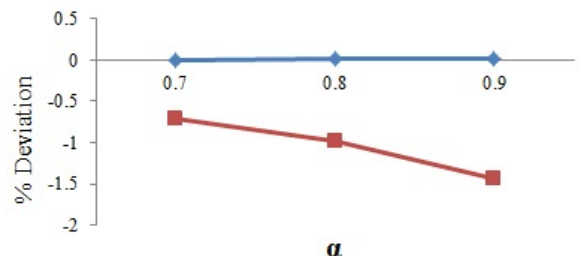
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

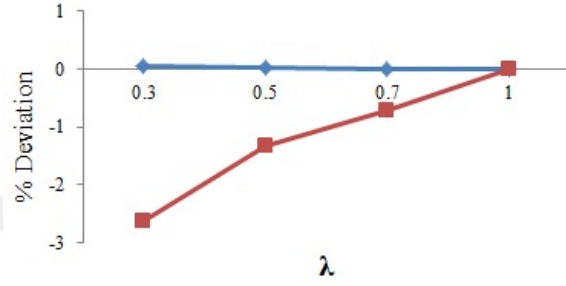


(e) For $\lambda = 0.5$

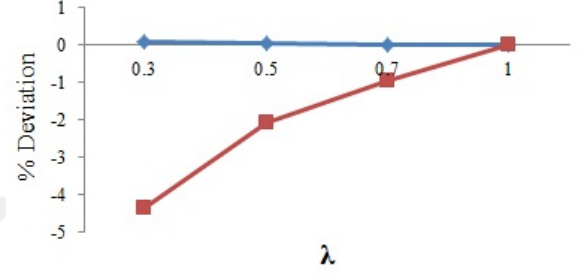


(f) For $\lambda = 0.7$

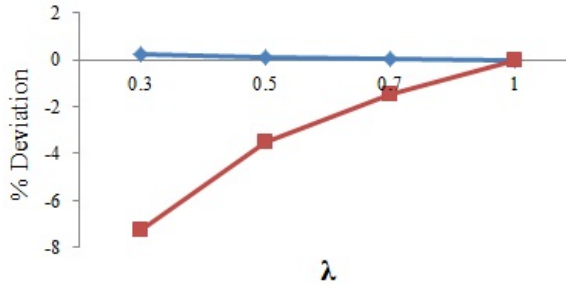
Figure C.8: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 600 Scenario Case



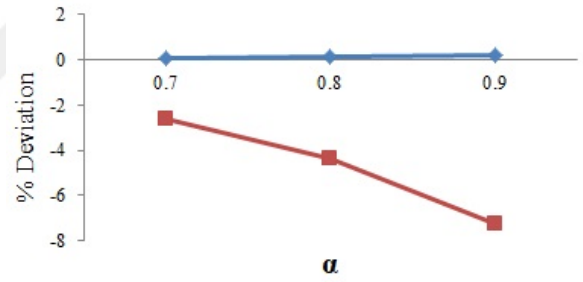
(a) For $\alpha = 0.7$



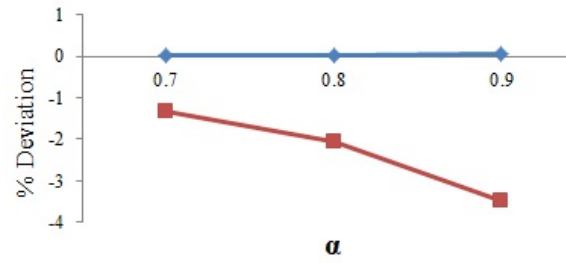
(b) For $\alpha = 0.8$



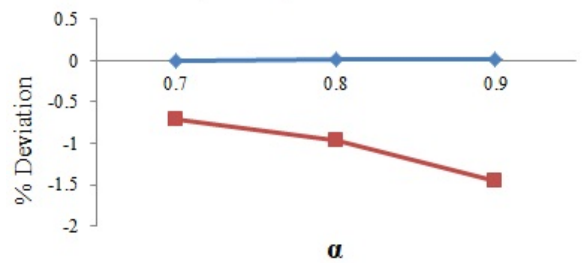
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

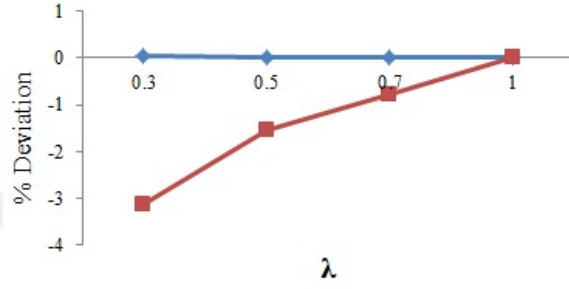


(e) For $\lambda = 0.5$

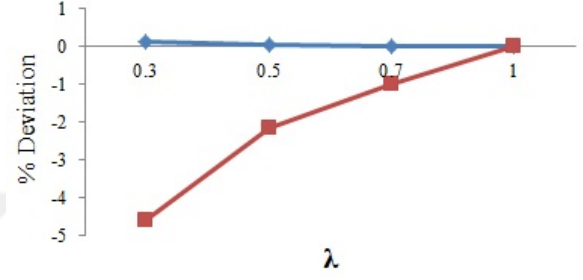


(f) For $\lambda = 0.7$

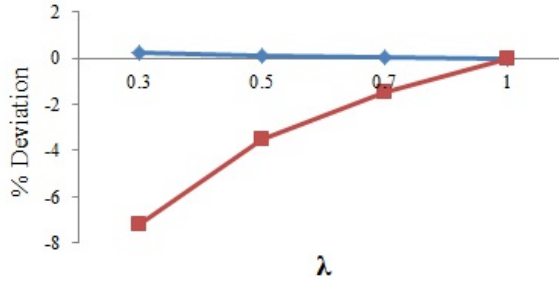
Figure C.9: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 700 Scenario Case



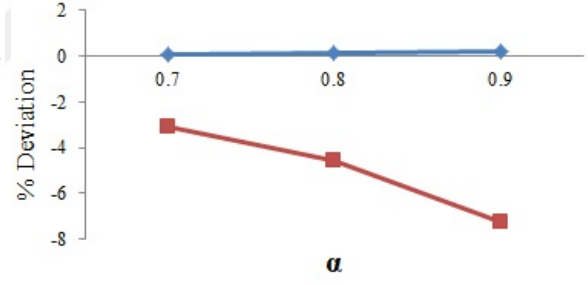
(a) For $\alpha = 0.7$



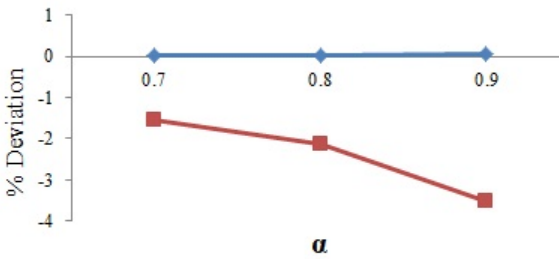
(b) For $\alpha = 0.8$



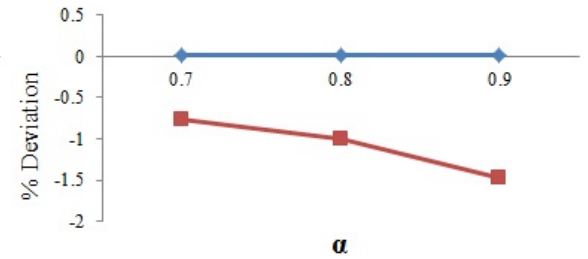
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

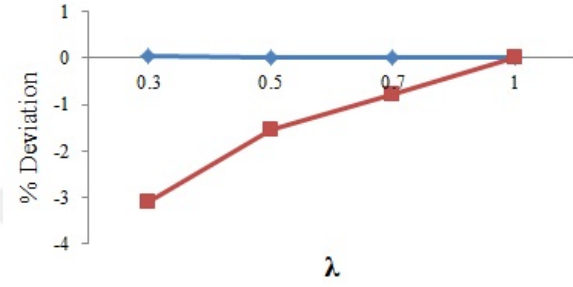


(e) For $\lambda = 0.5$

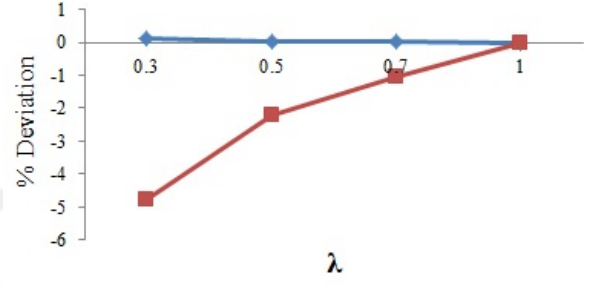


(f) For $\lambda = 0.7$

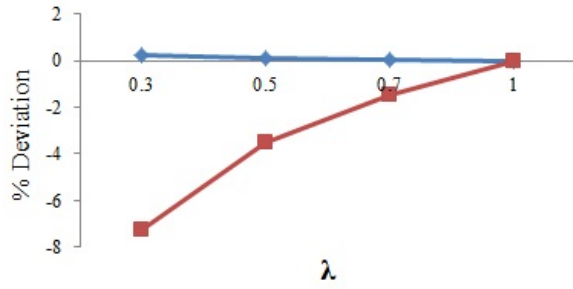
Figure C.10: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 800 Scenario Case



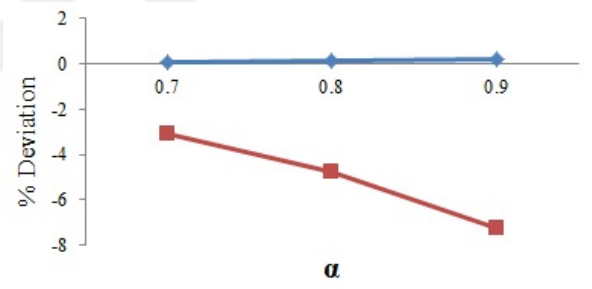
(a) For $\alpha = 0.7$



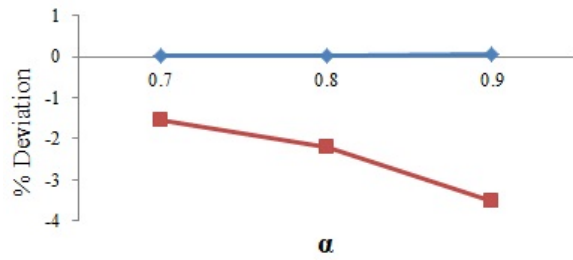
(b) For $\alpha = 0.8$



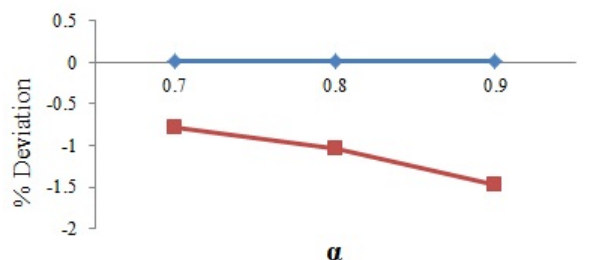
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$

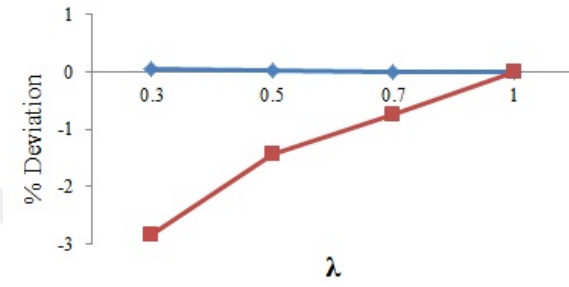


(e) For $\lambda = 0.5$

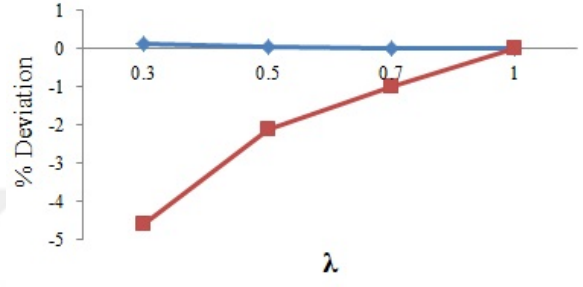


(f) For $\lambda = 0.7$

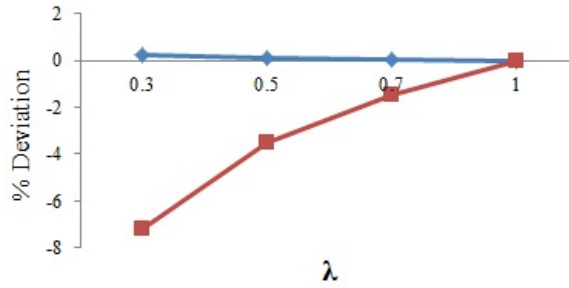
Figure C.11: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 900 Scenario Case



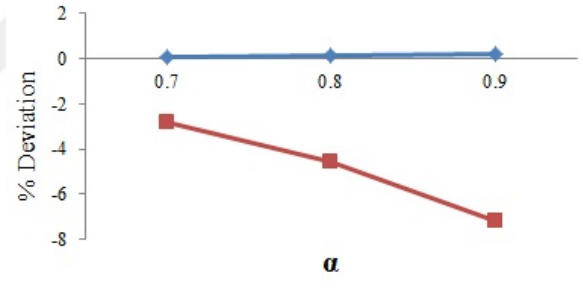
(a) For $\alpha = 0.7$



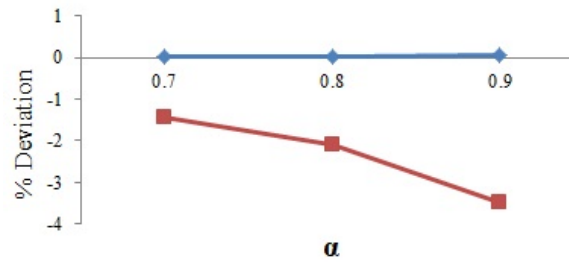
(b) For $\alpha = 0.8$



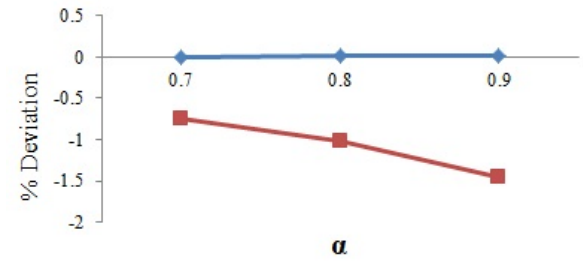
(c) For $\alpha = 0.9$



(d) For $\lambda = 0.3$



(e) For $\lambda = 0.5$



(f) For $\lambda = 0.7$

Figure C.12: % Deviation from Risk-Neutral Case for $\alpha \in \{0.7, 0.8, 0.9\}$ and $\lambda \in \{0.3, 0.5, 0.7\}$ for 1000 Scenario Case