A Stochastic Programming Approach to Analyze Design and Management of Flexibility in Infrastructure Systems Operating Under Long-Term Uncertainty

ZHANG SIZHE

(Master of Science, Huazhong University of Science and Technology)

(M.Sc., National University of Singapore)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INDUSTRIAL AND SYSTEMS ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2016

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Zhang Sizhe

Zhang Sizhe

1 January 2016

ACKNOWLEDGEMENTS

I would like to thank gratefully Prof. Michel-Alexandre Cardin for his mentorship, intellectual and diligent support, and for the enriching opportunities he gave me during my memorable journey at NUS since January 2012. Following your guidance, and working with you every day has been my most exciting and rewarding experience, and also my great honor and pleasure. I have been so fortunate to receive your advice on how to think carefully about research, how to well organize timetable, and how to keep an open mind in my life. The opportunities of studying in Delft University of Technology and the Open University you supported have been unforgotten pages in my life. Your encouragement and support has helped me building my confidence over the last four years. I hope I will continue contributing the legacy you has set forward.

I would like to thank my doctoral committee members Prof. Ang Beng Wah and Prof. Ng Szu Hui for their intellectual support, advice, and time. I would like to thank gratefully you for raising considerable questions during the committee meeting. It has been my honor and pleasure – to answer those questions thoroughly – to enrich my doctoral research. Your rigorous attitude mode of thinking will always be treasure in perseverance in my life.

I also would like to thank gratefully Prof. Paulien Herder and Prof.

William Nuttall for their intellectual support, generosity, and time. During my time in Delft, Prof. Herder warmly welcomed me to her team and colleagues, and provided me valuable comments on my research. Doubtless your passion and diligence for work is influential and encouraging, which I will always keep in my mind. Thank you Bill for welcoming me to your groups in Milton Keynes and Cambridge. This has been one of the most important phases in my doctoral work. Your expertise in energy sector has been greatly helpful for me to access this field and build up my own work. Your contribution to the nuclear power project has been part of this thesis and will definitely being honored.

I would like to thank my friends and classmates within our research group who made this time the greatest: Ranjbar Bourani Mehdi, Deng Yinghan, and Xie Qihui. I also would like to thank these research staffs who helped me a lot in research: Sun Chang, Aakil Mohammad Caunhye, Jiang Yixin, Hu Junfei, Yue Ka-Ho Howard, Mark Philip DeLessio, and Ng Kwok Kei Simon. In addition, I would like to thank my friends in the department who gave me inspiring suggestions over the last four years: Chao An-Kuo, Jing Lei, Yang Linchang, Weng Renrong, Zhang Linmiao, He Peijun, Yang Zhiyue, Zhou Chunhui, Liu Yue, Xu Xin, Zhang Chen, Qiu Zhipeng, Wang Xin, Wang Rui, Jiang Weiwei, and Liang Hongdei.

I am most grateful to the Department of Industrial and Systems Engineering, Faculty of Engineering, and National University of Singapore that generously provided funding over the last four years to support my research. I am also thankful to NUS Libraries that provided abundant resources – books and journals – to help me organizing this dissertation. Sincere gratitude is of course given to all other faculty members and staffs in the department.

Last but not least, I would like to thank gratefully my family members who always support me no matter how difficult it is. I would like to thank sincerely my parents who made this moment the greatest and happiest. I also would like to thank honestly my wife and daughter who made me the most fortunate man in the world. Without your support, I could not go this far and finish this dissertation. Remember, your happiness is my eternal motivation.

TABLE OF CONTENTS

SUMMARY

The design and management of infrastructure system under long-term uncertainty is a challenge to system planners and/or managers. Such infrastructures are critical to the well-functioning of modern cities and society, providing for example emergency and medical services, power, transportation, etc. Yet, they require billion if not hundreds of million dollar investments, and architecture/design decisions will last for several decades of planned operations. Typical approaches to system architecture and design do not account well for long term uncertainty, leaving the system to perform potentially in a sub-optimal manner some years later after launch. To improve the expected long-term lifecycle performance, a novel approach incorporating the concept of flexibility is proposed in this study. Similar to the concept of a real option, flexibility in engineering design provides "the right, but not the obligation, to change a system easily to adapt to the realization of uncertainty drivers." The proposed approach explored in this study is a multi-stage stochastic programming model based on a Sample Average Approximation scheme. The flexibility strategies are modeled using the novel concept of managerial decision rules and captured by non-anticipative constraints in the model. Such approach differs from standard methods in real options analysis used to analyze flexibility in irreversible investment projects and based on

dynamic programming techniques. It provides more freedom to emulate the actual decision making process to exercise flexibility at optimal times, parameterizing the design and decision making process, and making it suitable for stochastic programming. The proposed approach to analyze flexibility in infrastructure systems is explored through two engineering applications. The first application is about the design and management of Emergency Medical Service systems. The problem consists of deploying emergency resources over time and space during the system's life cycle considering flexibility, which is an extended version of existing capacity and resource allocation problems. The second application is about nuclear power plants. It aims to study deployment of nuclear plants capacity in phases to satisfy electricity demand and to deal with uncertainty related to the social acceptance of the technology. To make the approach applicable and validate the solutions found using the default Branch & Bound algorithm in the optimization software, a hybrid algorithm is introduced for solving the stochastic programming problem. The hybrid algorithm finds the same form of solutions as the Branch and Bound algorithm; similar expected lifecycle cost, and outperforms other alternatives in terms of the quality of solution and the time to best solution.

NOMENCLATURE

LIST OF ACRONYMS

ADSR (Accelerator-Driven Subcritical Reactor)

aOFAT (adaptive One-Factor-At-A-Time)

ALS (Advanced Life Support)

BFS (Basic Feasible Solution)

BLS (Basic Life Support)

B&B (Branch and Bound)

B-S (Black–Scholes)

CCS (Carbon Capture, Transport, and Storage)

CPA (Change Propagation Analysis)

DBD (Decision-Based Design)

DOE (Design of Experiments)

DSM (Design Structure Matrix)

DTM (Design Theory and Methodology)

DCF (Discounted Case Flow)

DP (Dynamic Programming)

EA (Evolutionary Algorithm)

EMS (Emergency Medical Service)

EoS (Economies of Scale)

ESM (Engineering System Matrix)

GDP (Gross Domestic Product)

GDT (General Design Theory)

GA (Genetic Algorithm)

GBM (Geometric Brownian Motion)

HC (Hill-Climbing)

IAEA (International Atomic Energy Agency)

INES (International Nuclear Events Scale)

KPI (Key Performance Indicator)

LCOE (Levelized Cost of Electricity)

LP (Linear Programming)

MALP (Maximal Availability Location Problem)

MCLP (Maximal Covering Location Problem)

MEXCLP (Maximal Expected Covering Location Problem)

MDP (Markov Decision Processes)

MIP (Mixed Integer Programming)

MATE (Multi-Attribute Tradespace Exploration)

MSCLP (Multi-stage Set Covering Location Problem)

NPV (Net Present Value)

NPP (Nuclear Power Plant)

OFAT (One-Factor-At-A-Time)

O&M (Operations and Management)

QoS (Quality of Solution)

RAM (Random Access Memory)

ROA (Real Options Analysis)

SAA (Sample Average Approximation)

SCLP (Set Covering Location Problem)

SA (Simulated Annealing)

sDSM (sensitivity Design Structure Matrix)

SMRs (Small and Medium Size Reactors)

STD (Standard Deviation)

TS (Tabu Search)

TBS (Time-to-Best-Solution)

LIST OF SYMBOLS

d_t = demand of uncertainty drivers at period t

I_t = international nuclear event scale at period t

μ = expected growth rate of Geometric Brownian Motion process

σ = standard deviation of Geometric Brownian Motion process

p = probability of a random event or a sample scenario

r = discount rate factor

R = coverage radius or fleet size of an emergency (fire) station

U = capacity of a facility in an infrastructure system

CoV = managerial requirement on incident coverage rate

b = economic of scale factor

ε = a very small tolerance

M = an arbitrary large integer

LIST OF FIGURES

LIST OF TABLES

Chapter 1 INTRODUCTION

"Human nature is knowledge." – Aristotle (c. 384 – 322 BC)

"The beginning is the most important part of the work." – Plato (c. 428 – 348 BC)

This thesis is concerned with the design and management of flexible infrastructure systems in urban contexts under uncertainty. The thesis presents a novel analytical methodology to design flexible infrastructure systems in uncertain environments and to evaluate design alternatives objectively and quantitatively based on the anticipated long term life cycle performance. The proposed methodology aims to complement existing methods for design and large-scale system project evaluation by analytically measuring the anticipated long term life cycle performance of the system in consideration of strategic-level flexibility. The strategic-level flexibility strategies are exercised by managerial decision rules and captured by the non-anticipative constraints in the mathematical models, which contrasts to standard Real Option Analysis (ROA) methods based on dynamic programming (DP), such as the Black-Scholes model (Black & Scholes, 1973), and binominal lattice-based analysis. Managerial decision rules are akin to "IF-THEN-ELSE" statements similar to triggering mechanisms that provide guidance on when and how it is best to exercise flexibility strategies based on the realization of uncertainty sources.

The idea of flexibility, also known as real options, is the right – but not the obligation – to undertake initiatives for mitigating unexpected risks or for seizing new opportunities to improve life cycle performance. For example, a parking garage sited near the Blue Water Shopping Mall in the United Kingdom was studied by considering this concept (de Neufville *et al.*, 2006). Typical design approaches favor deploying capacity all at once at the beginning of the life cycle. In the garage example, a typical approach would have been to design the garage to satisfy demand up to a certain (optimal) capacity. Such design would either lose demand and opportunity for profit when demand was higher than installed capacity, or waste money on high construction costs and unused capacity when the demand was lower than expected. In contrast, a flexible parking garage would deploy smaller capacity initially, thus reducing initial capital costs, but build in flexibility in the form of stronger pillars and infrastructure, for a potential vertical capacity expansion, *only if and when needed*. If the demand was not satisfied by current capacity for two consecutive years, managers would expand capacity vertically by adding one more floor. If the demand was flat or even lower down, then this capacity expansion may not be needed. This flexibility strategy on the one hand would help save money by deploying less capacity initially (i.e., mitigating downside risks), and on the other hand provide contingencies to capture more demand if needed and generate more profits (i.e., capitalizing on upside opportunities). Other real world examples of flexibility in engineering systems design include the Health Care Services Corporation building in Chicago, Tufts Dental School in downtown Boston (Guma *et al.*, 2009), and 26 de Abril Bridge in Lisbon. These examples mainly relate to the design of old-fashioned, capital-intensive infrastructure, which are typically long-lived and operate in significantly uncertain environments. The anticipated long term life cycle performance is an observational performance indicator that is expected under particular conditions over the system life cycle. This performance can be measured either by financial metrics, such as Net Present Value (NPV) and Levelized Cost of Electricity (LCOE), or by non-financial metrics, like the service level in supply chain management or incident coverage rate (or fleet size) in an Emergency Medical Service (EMS) system.

The design and management of infrastructure systems is challenged by the fast growth of urban population. Indeed system needs, demands, and regulations will inevitably change over a long system life cycle (Eckert *et al.*, 2009). The anticipated long term life cycle performance of a system will be influenced by such changes and may even fail to meet the target. In planning and architecture of this type of systems – such as emergency services, power plants, electricity distribution grids, and water supply networks – it is important to consider the ability to adapt and upgrade the infrastructure system to efficiently support essential public services and urban development under uncertainty, so as to ensure continued quality services. Plainly optimizing the planning and operations in a rigid manner may leave aside alternative solutions that could extract additional value from uncertainty, position better the system to mitigate risks and capitalize on upside opportunities. Another approach for dealing with uncertainty is robust design. The robust design allows a system to perform well in a range of possible scenarios without the need to change the configuration. The robust system is not changeable based on the realization of uncertainty sources, and thus may not be able to take advantage of upside opportunities as well as reducing downside risks. The analysis of flexibility also poses a challenge to existing engineering systems design and project evaluation approaches. The form of the solutions from typical real options analysis used to quantify the value of flexibility based on dynamic programming and ordinary differential equations (Dixit & Pindyck, 1994; Trigeorgis, 1996) may be difficult to understand and implement in operations when managers lack training in economics and/or the advanced backward induction processes required as part of these methods. This is because decision-makers need to position themselves in the decision tree, project the future uncertainty drivers, and roll back to the present time to determine the best course of action. This requires advanced mathematical training and a deep understanding of dynamic programming techniques not only to find the actual solution, but to use it in operations. The performance of a flexible system may be worse than that of an inflexible system if the flexibility is incorrectly implemented. Such issues will affect the appreciation of flexibility in engineering systems design.

This thesis addresses the challenges of the design and management of infrastructure systems and the analysis of flexibility by explicitly considering uncertainty, strategic-level flexibility, and managerial decision rules. The objective is to help infrastructure systems obtain a better life cycle performance under uncertainty by indicating stochastically optimal flexible design alternatives, while providing solutions that are easy to in practice. To do so, a novel analytical methodology is proposed to rigorously design and evaluate flexible infrastructure systems. Section 1.1 specifically explains the motivations underlying this work. The intended audience and limits of this thesis are described in Section 1.2. Section 1.3 provides an overall summary of the structure of this thesis.

1.1 Motivations

1.1.1 Current Considerations in Design and Management of Infrastructure Systems

The design and management of infrastructure systems that mainly focus on distributing resources to particular customers in order to satisfy specific requirements (e.g., service level, demand) will be the major concerns of this thesis. Two examples of infrastructure systems (an EMS system and a nuclear power plants (NPP) system) will be discussed in the case studies of the thesis. The EMS system is dedicated to providing out-of-hospital medical care to patients and transporting them to the hospital if needed. The nuclear power system, which consists of multiple nuclear power plants, focuses on generating electricity and transmitting it to consumers through the power grid to fulfill a contract. These are conceptually good examples of infrastructure system as both EMS systems and NPP systems are critical to urban society and provide specialized, necessary services. Out-of-hospital treatment is absolutely needed for saving people's lives, while electricity is a daily necessity in modern society. How to deploy and maintain the capacity of those systems is a major concern for decision makers, especially considering a long term horizon, which is inevitably subject to demographic, market conditions, regulatory, and technological changes. More importantly, these requirements are highly uncertain and may evolve over time, which makes the deployment even more difficult. Design for flexibility – especially deploying the capacity of those infrastructures over time and space – gives us an opportunity to change the system to adapt to the future scenarios and thus improve the system's life cycle performance.

Approaches for the design and management of infrastructure systems are similar to those for typical engineering systems. Thus, the consideration of uncertainty in typical engineering systems design should be included in the design of infrastructure systems. Uncertainty is defined as anything (known or unknown) affecting the performance of an engineering system (Cardin, 2011). The importance of uncertainty in engineering systems is well demonstrated by the case of the Iridium cell-phone system. The Iridium satellite constellation was launched by Motorola® in 1998 to provide cell-phone communication over Earth's surface and consisted of 66 satellites in orbit plus spare satellites in case of failure (Fossa *et al.*, 1998). The demand forecast was over optimistic and the company then decided to launch all satellites once instead of deploying a part of the capacity. Soon the development of land based cellular network occupied the market. The Iridium system thus lost a lot of demand and eventually failed. The system cost $4 billion to create but sold for only about $20 million, making the company bankrupt (Hesseldahl, 2001). Overestimating the demand for Iridium technology, and underestimating the level of competition, was the main cause of this failure (de Weck *et al.*, 2004). This case clearly illustrates how uncertainty can affect the performance of engineering systems.

In the design and management of infrastructure systems, it is often the case that considerations of uncertainty are based on the short term. Moreover, such considerations are simplified by weighting one or several future scenarios. These current approaches make sense, but are incomplete in terms of reflecting the future. The life cycle or lifetime of an infrastructure system is relatively long (e.g., 10+ years), and the growth of the urban population during the life cycle has direct and significant impacts on the demand of the system.

Considering short-term uncertainty in the design procedure is of little help for strategic-level decision making and management. The system may be successful at the beginning of the life cycle, but become sub-optimal a few periods later. This ultimately affects the overall life cycle performance because the design does not fully and explicitly account for the future.

Due to bias in the uncertainty, designers and managers prefer robust systems that can resist uncertainty and maintain a certain level of performance over the life cycle. Uncertainty, however, does not only bring risks to engineering systems, but also creates opportunities. The upside of demand uncertainty could bring additional profit to a system designed for an average demand scenario, but only if the system has the ability to adapt to the change. Robust systems are incapable of taking advantage of upside opportunities, even though they are successful in dealing with downside risks. On the other hand, flexibility in engineering systems design provides "the ability but not the obligation, to easily change the system in face of uncertainty" (Trigeorgis, 1996). Design for flexibility enables systems to pro-actively deal with uncertainty and change themselves accordingly over a range of uncertainty scenarios in order to improve the expected system life cycle performance. Such applications demonstrate the positive impact of flexibility in engineering systems design, although design for flexibility has not spread widely yet.

1.1.2 Current Real Option Analysis Methods

The other motivation of this thesis is to find the stochastically optimal timing to exercise flexibility. In real life flexibility sometimes may not be exercised correctly, and/or never exercised in some cases. This is due to several reasons such as the lack of necessary knowledge for implementation, the change of leadership, or missing of document materials tracking such developments in infrastructures.

Several ROA methods exist for evaluating flexibility in engineering design (or real options). A close-form method, like Black-Scholes (B-S) from financial options analysis, is applicable when constant cost is assumed in the system (Black & Scholes, 1973). The finite difference method for option pricing is sometimes used if the option can be modeled using a partial differential equation (Brennan & Schwartz, 1985). As the most commonly used method, binominal lattices allow for flexibility analysis, where relevant and differing rules may be encoded at each node (Copeland & Antikarov, 2001). Monte Carlo simulations, in contrast to the above methods, provide opportunities to tackle high-dimension problems and deal with a large number of uncertainty scenarios (de Neufville & Scholtes, 2011).

For ROA in a current engineering context, lattice-based methods and Monte Carlo simulations are used more by decision makers than the B-S model. This is because the B-S model is not well suited for real options due to some important differences between financial options and real options (de

Neufville, 2010) – more details are provided on these in Sections 2.1.4 and 2.4.2. However, the solution emerging from a lattice-based ROA approach may be difficult to analyze and implement in practice. On the one hand, lattice-based methods can only consider and analyze one decision rule at a time (i.e., they often use the same expected value maximizing rule inspired from Bellman's formulation). On the other hand, without the necessary mathematical knowledge, practitioners and managers may find it difficult to apply the folding back procedures to determine the optimal policy. These raise questions regarding both the evaluation and the analysis of embedded flexibility for more practical applications. For simulation-based approaches, even though they are able to take into account multiple uncertainty sources and decision rules at a time, they may be time-consuming to use if finding the stochastically optimal initial configuration and the best decision rules for a flexible design is the objective. More details on standard ROA methods is provided in Section 2.1.4, while Section 2.4.2 provides more details on the criticisms with each ROA method.

Using only available tools, it may be challenging to analyze different rules when the total number of possible rules is considerably large. Only a subset of all the possible rules can be evaluated in a finite time. If lucky, the optimal rule may be found through an exhaustive search. However, in most cases, sub-optimal solutions are obtained, treated, and applied as optimal solutions. The value of flexibility given by such evaluation procedures may not be the stochastically optimal one (and may be far from it). This value could even be negative if the rule is sub-optimal. Thus, flexibility in engineering systems design may not be favorable to decision makers due to this incomplete evaluation. This biases practitioners' and managers' decision making regarding what type of design should be chosen (i.e., robust or flexible design).

The value of flexibility may also be affected by incorrect exercise in practice. Solutions from lattice-based models may not be intuitive to use without the required mathematical knowledge of dynamic programming, and thus are difficult to correctly implement. Indeed, determining when to exercise flexibility strategies in lattice-based ROA approaches requires decision makers to determine their position in the lattice, and apply a backward induction process to find the optimal policy. If the process is conceptually difficult to use, it has less chances of being used in practice. Finding the position in the lattice may be challenging, as it requires fitting the evolution of historical data to the closest and nearest stage and state. Another issue is that design for flexibility may require a larger budget for infrastructure (e.g., strong pillars for capacity expansion in the case of parking garage). Thus, the flexibility may become a burden to practitioners and managers if it is incorrectly or never exercised. The long term life cycle performance of the system will ultimately be affected by such inappropriate decision making.

The issues discussed above in the evaluation and analysis of flexibility have not been well addressed in the literature in the area of ROA; more specifically, no existing method can deliver an acceptable outcome as a simple, intuitive, and optimal policy for exercising flexibility strategies in the face of long term uncertainty so that the value of a flexible design solution could be fully appreciated by decision makers. Thus, this thesis aims to address this particular issue more analytically by modeling a design problem using mathematical programming, with the goal of optimizing the anticipated long term life cycle performance. The output of the proposed model is supposed to offer decision makers a guidance or policy to exercise flexibility strategies.

1.2 Intended Audience and Application

This thesis targets system designers, managers, and practitioners who are working in engineering systems design, especially in the field of urban systems. The design procedure for flexibility allows infrastructure systems to change themselves according to the different scenarios that may arise. Flexible designs have the ability to pro-actively deal with uncertainty, while robust designs focus on having a high performance under a range of scenarios without the need to change the system configuration. The design procedure was specifically crafted for the design of infrastructure systems, but can be applied to normal engineering/complex systems as well.

The proposed design approach aims to identify the stochastically optimal flexible design configuration, including the initial configuration and follow-up managerial decision rules. The approach aims to find the optimal design configuration among all possible flexible designs under the flexibility paradigm, and to improve the expected life cycle performance a step further. This differs slightly from the idea of "optimal design" in the traditional design paradigm. The ROA method considered in the approach is different from the existing methods often used in the evaluation of flexibility. In the proposed approach, the design problems are formulated using mathematical programming. Strategic-level flexibility is analyzed via managerial decision rules and captured by non-anticipative constraints. This is not the first time that managerial decision rules have been applied for the analysis of flexibility: past work considered Monte Carlo simulations as tools to evaluate decision rules under uncertainty (de Neufville & Scholtes, 2011). Such studies typically proposed a decision rule and analyzed the system flexibility in light of such decision rule, without necessarily finding the one is optimal in a stochastic sense. For example, in the parking garage case mentioned above, a threshold value was given for triggering the exercise of flexibility. That is, if the demand cannot be satisfied for N consecutive years, the decision rule will be exercised. The flexibility strategy may be exercised too often when this value is small (e.g., $N = 1$), thus increasing costs with too many unnecessary expansions. Also the strategy may seldom be exercised if this threshold is too high (e.g., $N = 3$ or more), so that the flexibility cannot help capitalizing on upside opportunities. It is clear that an inappropriate threshold may affect the value of flexibility and its corresponding exercise in practice. Mathematical programming, by contrast, allows for the evaluation of a variety of decision rules and the determination of the stochastically optimal one or the so called "sweet spot" for that threshold. It should be noted that the true optimal design configuration is not always available because of the complexity of the design problem. The procedure suggests a feasible approach for systematically evaluating and implementing flexibility in a finite amount of time, with the goal of getting as close as possible to the optimal design.

This thesis also targets researchers who are interested in studying and promoting flexibility in engineering systems design. The proposed methodology hopes to open their minds about analytically modeling design problems and collaborating with experts in different areas (e.g., optimization, game theory, system dynamics, graph theory, etc.). For specific design problems – such as power distribution grids and urban transit – particular expertise can be a catalyst.

The design approach proposed in this thesis can also be used more widely in engineering systems design and applied to systems that are not in urban contexts. In particular, the methodology is suited for the design and management of resource allocation systems (of which the EMS system is an example). Moreover, tactical- and/or operational-level flexibility can also be considered in the design procedure.

1.3 Thesis Structure

The thesis is organized as follows. Chapter 2 surveys the literature to give the audience an overview of the existing work regarding flexibility and to determine the contributions of previous research in the area of engineering systems design. This chapter also investigates the current work in terms of EMS systems and nuclear power systems and discusses the shortcomings of the existing work. Chapter 3 introduces the research questions, hypothesis, and research approaches of this study. Chapter 4 presents the methodology for the design and management of flexible infrastructure systems from a generic standpoint, including the design and evaluation procedures. Chapter 5 and 6 describe the applications used to verify the methodology in two urban contexts. Specifically, Chapter 5 describes an application of design and management of EMS systems for flexibility, while Chapter 6 describes an application of siting nuclear power plants for flexibility. A problem statement, uncertainty recognition, mathematical formulations, and numerical analysis are contained in each chapter. Chapter 7 summarizes the findings of the previous two chapters. Chapter 8 concludes the work and highlights research opportunities in related areas.

Chapter 2 LITERATURE REVIEW

"The only thing you absolutely have to know, is the location of the library." – Albert Einstein (1879 – 1955)

"Know the enemy and know yourself, and you can fight a hundred battles with no danger of defeat." – Sun Tzu (c. 545 – 480 BC, the Art of War)

This chapter aims to fulfill five objectives. The first objective is to position the thesis within the design for flexibility paradigm, and to identify potential contribution to current research work in this field, by analytically modeling design problems using mathematical programming. An overview of the existing state of flexibility in engineering systems design is provided in Section 2.1. The second objective is to identify potential contribution to existing designs of EMS systems by incorporating strategic-level flexibility, and to determine how expected life cycle performance can be improved by such flexible designs. Section 2.2 provides an overview of past and current designs in the context of EMS systems. The third objective is to identify potential contribution to nuclear power systems by considering a socio-technical factor in the design procedure. To this end an overview of nuclear power systems is provided in Section 2.3, including engineering design and social aspects. The EMS system and the nuclear power system are two representative cases that will be studied in the later chapters of this thesis.

First of all, both systems are good examples of infrastructure system. They also have more than one facility to maintain the regular operation, and the facilities are geographically located at different sites. Both systems offer their resource or service to customers to fulfill specific requirements, either from a single facility (e.g., an emergency station and/or an emergency vehicle), or a whole system (e.g., a power system). Last but not least, these two systems are obviously important, and there is still plenty of room to improve the system's performance. On the one hand, an EMS system is indispensable to the society as it provides out-of-hospital medical care to save patients' lives. On the other hand, the nuclear power system plays an important role in the power system, generating electricity without large CO_2 emissions. Any improvement in these systems could bring significant economic and/or social influence. Section 2.4 explains identified research gaps and opportunities in each field, to fulfill the fourth objective. The fifth objective is to highlight the contributions by indicating how this thesis intends to address these issues and extends to broader applications.

2.1 Flexibility in Engineering Systems Design

To best position the thesis in design for flexibility paradigm, this section summarizes the existing state of such research field. Since this thesis aims to propose a novel design approach for infrastructure systems, one may ask "how can one improve the long term performance of an infrastructure system?" The concept of flexibility will be considered as a possible design alternative to address this question. This idea of flexibility in engineering systems design, also known as real options, is to provide "the right but not the obligation, to easily change the system in the face of uncertainty" (Trigeorgis, 1996). Flexibility is a multi-disciplinary concept which is popular in many fields (Saleh *et al.*, 2009), such as decision theory, manufacturing systems, engineering design, etc. To people from different fields, flexibility indicates different meanings and even has different names (e.g., adaptability). In the realm of engineering systems design, Saleh *et al.* (2009) suggested there was a need to make a distinction between flexibility in the design process and flexibility of a design (or flexible systems), as they were considerably different and were implemented by different means. In a design process, flexibility is considered as the opposite of rigidity in specifying system requirements. On the other hand, flexibility of a design usually implies that such system has the ability to adapt to the change of circumstances, with certain characteristics. The thesis focuses on the latter concept, dedicating to improve system life cycle performance by designing flexible systems.

Tomiyama et al. (2006; 2009) developed a well-reputed classification of "Design Theory and Methodology" (DTM) for engineering design procedures based on General Design Theory (GDT) (Tomiyama & Yoshikawa, 1987; Yoshikawa, 1981; Yoshikawa & Uehara, 1985), where a variety of design methodologies were discussed, such as adaptable design, axiomatic design, Taguchi method (robust design) (Taguchi, 1987), Design Structure Matrix (DSM) (Steward, 1981), etc. de Neufville and Scholtes (2011) proposed a practical four-step process for developing flexibility in engineering systems design. In the light of de Neufville and Scholtes' work, Cardin (2014) provided a five-phase structure for enabling flexibility in engineering systems. The overview is based on the taxonomy prepared by Cardin (2014), as the five phases are discussed in Subsections 2.1.1 to 2.1.5 in sequence. Subsection 2.1.6 focuses on fulfilling the objective of situating this thesis as part of the flexibility in engineering systems design.

2.1.1 Phase I – Baseline Design

It is better to consider design for flexibility starting from an existing design configuration. Existing design configurations are usually optimized based on deterministic point forecasts, which can be considered as baseline designs or rigid designs. They are often not capable of changing themselves to adapt to uncertainties, as explained by Cardin (2014). Baseline designs may perform well in resisting downside risks attribute to robustness, but may fail in capturing upside opportunities. This phase aims to help designers and managers start from what they know and/or what they have on hand, and to help structuring necessary thought process regarding design for flexibility.

Design procedures for typical engineering systems (e.g., satellite systems, power plants or oil platform) are suited for baseline design. The concept of design procedure discussed here indicates the same idea as DTM does. Detailed procedures can be found in Tomiyama et al. (2006; 2009), Finger and Dixon (1989a, 1989b). Note that such designs are not necessarily to be "purely" rigid. Some designs for resource allocation problems considered phasing option (i.e., deploy resource in phases) to improve system life cycle performance. Relevant literature is discussed in Section 2.2. These designs, however, are not conceptually flexible designs, because the option is exercised via fixed rules. Systems cannot change themselves to adapt to different uncertainty scenarios by taking advantage of such option. Thus, those designs are treated as baseline designs as well in the thesis, particularly for infrastructure systems.

2.1.2 Phase II – Uncertainty Recognition

Uncertainty always exists, and really has impacts on system life cycle performance. Normally uncertainty is considered as a source of downsides as it can destroy value, such as unexpected performance of technology or painful environmental conditions (e.g., earthquake, tsunami). On the other hand, it can also bring opportunities for better performance, such as better market conditions or lower costs. As uncertainty is inevitable to engineering systems design, recognizing uncertainties is considered to be the prerequisite to design for flexibility.

The procedures in this phase help designer find out and model main uncertainty sources, based on formal and practical approaches suggested by de Weck and Eckert (2007), aiming at quantifying, characterizing and modeling major uncertainties for systems designs. Formal approaches are comprised of probability theory, statistics, and Bayesian theory (Bayes & Price, 1763). On the other hand, practical approaches consist of binominal lattice, decision trees, and scenario planning. The binominal lattice model allows for modeling both uncertainty and outcome at discrete time steps (Cox *et al.*, 1979). Dynamic programming is used in the recursive calculation at each node of the lattice. Lattices-based models are better suited for continuous events (i.e., uncertainties with continuous probability). In contrast, decision tree models are suited for discrete events (i.e., uncertainties with discrete probability), and so be scenario planning (Helmer, 1967; Howard, 1966). The limits of these practical approaches are fairly clear. To deal with more than one uncertainty source, the development of quadranomial and even multinominal approaches are alternative approaches. They are, however, very difficult to use and analyze because of the curse of dimensionality, which is exacerbated with considerations of multiple uncertainty sources (Copeland & Antikarov, 2001; Kamrad & Ritchken, 1991).

To moderately fulfill the target in this phase, either formal or practical approaches are necessary. Statistical tools are used for regression analysis with respect to historical data. Probability theory and Bayesian theory are applied for determining probability distribution, as complements to reversion analysis.

The difficulties of applying such formal approaches are selecting proper data set (e.g., range, availability), mean trend profile (e.g., linear, polynomial), stochastic model (e.g. Geometric Brownian Motion (GBM) process, mean reversion process), etc. All formal approaches including scenario planning are applicable in industry (de Weck & Eckert, 2007; Halpern, 2003; Morgan & Henrion, 1992). Lattices-based and decision tree models, by contrast, are not widely used in industry (Cardin & de Neufville, 2009; Engel & Browning, 2008).

2.1.3 Phase III – Concept Generation

Conceiving flexible designs is the main target in this phase, once the major uncertainties are identified and modeled. The procedures in the phase are concerned with strategy generation and enabler identification. Such procedures aim to provide directions to determine where to focus the design effort, as it is impossible to address all possible flexible concepts. Procedures are categorized according to what they specifically focus on. The outcome of the procedures is a set of flexible design concepts that can proactively deal with uncertainty, compared to baseline designs obtained in phase I. Overall, the performance of flexible design concepts can arguably better that that of baseline designs.

Strategy generation

The purpose of strategy generation is fairly obvious. It aims to generate flexible strategies and design alternatives to flexibly encounter uncertainty. Useful guidelines are provided in industrial contexts for this strategy generation (Fricke & Schulz, 2005; Suh, 1990). Trigeorgis (1996) suggested canonical real option strategies to designers for generating flexible design concepts, such as option to expand, option to deploy in phases over time and space, option to defer investment, option to abandonment, etc. They can be further divided into flexibility "in" and "on" the system (Wang & de Neufville, 2005). In depth technical knowledge is required for technical flexibility in the system, while it is not necessarily required for managerial flexibility on the system. Besides, Cardin *et al.* (2013a) suggested an experimental methodology for generating flexible concepts via a short-lecture regarding the idea of uncertainty and flexibility, with a prompting mechanism. Designers are able to be aware of the impacts by uncertainty, and how flexibility could help address such issues.

Enabler identification

The objective of enabler identification is to identify the components of systems, which can be applied to enable the flexibility. Plenty of authors have contributed to the development and analysis to access flexibility in products (Keese *et al.*, 2009; Qureshi *et al.*, 2006; Rajan *et al.*, 2005). Many principles that are used in design for flexible products can be applied to design for flexible complex systems, even though these two entities are quite different.

Design Structure Matrix (DSM) introduced by Steward (1981) aims to show design tasks as a series of network interactions. It lists all constituent subsystems/activities, corresponding information exchange, interactions, and dependency patterns (Eppinger & Browning, 2012). The matrix can highlight a large number of system elements and their relationship in a compact way. As one of DSM-based procedures, change propagation analysis (CPA) focuses on changing potential areas for inserting flexibility. Application of applying CPA can be found in (Giffin *et al.*, 2009). Hu (2013) developed a DSM-based method relying on Bayesian network and risk propagation analysis, which is conceptually similar to CPA. The sensitive DSM (sDSM) looks for the most sensitive design variables to changes and requirements to embed flexibility (Kalligeros, 2006), while engineering system matrix (ESM) considers human factors and system drivers by extending original DSM (Strauss & Corbin, 1998).

2.1.4 Phase IV – Design Space Exploration

The procedures in the phase aim to help designers explore the design space to determine the most valuable flexible design concepts and decision rules for the system. This phase includes quantitative concepts evaluation and computationally efficient search. The outcome of this phase should be a set of recommended flexible design concepts with better life cycle performance than baseline design in Phase I. The flexible concepts are provided with clear life cycle performance quantifications, and recommendations in terms of managerial decision rules. This thesis aims at proposing a better quantitative approach for evaluating flexibility strategies for use in an engineering context, which is one of the main contributions.

Quantitative concepts evaluation

Real option analysis has been applied to quantitatively evaluate flexible concepts more recently (Copeland & Antikarov, 2001; Dixit & Pindyck, 1994; Myers, 1984), as it can handle uncertainty that Discounted Case Flow (DCF) analysis cannot. The techniques are directly derived from the valuation approaches developed to financial options theory (Black & Scholes, 1973). These include the Black-Scholes model, lattice-based, and simulation-based approaches. The B-S model is developed for financial options valuation, and is not well suited in an engineering context because of the difference between financial and real options.

Lattice-based approaches, in particular the binominal lattice based approach, is one of the most popular approaches to perform ROA in irreversible investment in real, large-scale projects. The binominal lattice model was firstly proposed by Cox *et al.* (1979). Figure 2.1 shows a graphical example of a binominal lattice model in consideration of three periods. S denotes the state at the initial period of a system's life cycle, which could be price (in financial theory), or other uncertainty drivers such as emergency incidents (in EMS systems) and electricity demand (in energy systems). Variables u and d represent the up and down factors scaling the response in previous period, with corresponding probabilities p and $1-p$, respectively, for some given probability p, $0 < p < 1$.

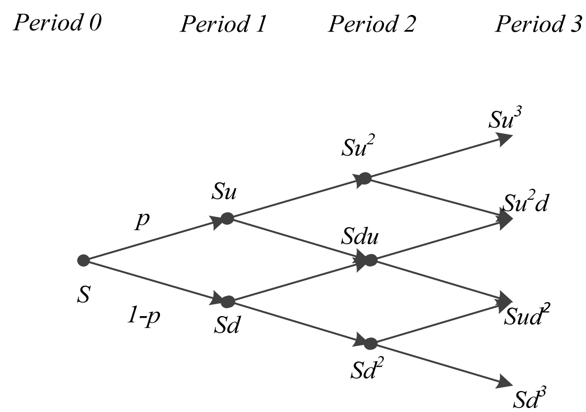


Figure 2.1 A three-period binominal lattice model.

In addition, μ and σ denote the periodically expected growth rate and standard deviation, respectively, which can be derived from historical data on a particular stochastic process. The model is derived on the implicit assumption of Geometric Brownian Motion (i.e., random walk). The values of the rates for moving up and down plus the probabilities can then be calculated as:

$$u = e^{\sigma}\sqrt{\Delta t} \tag{2.1}$$

$$d = e^{-\sigma}\sqrt{\Delta t} \tag{2.2}$$

$$p = \frac{1}{2} + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)\sqrt{\Delta t} \tag{2.3}$$

where Δt is the step length that could be a second, a day, or a year. The expected value of a n th node at period t with a discount rate r can be calculated by the following equation, based on Bellman's equations in dynamic programming:

$$E_{t,n} = \frac{1}{R}\left[pE_{t+\Delta t,n+1} + \left(1-p\right) E_{t+\Delta t,n-1} \right] \tag{2.4}$$

where $R = 1 + r$.

In decision analysis, the best design decision is determined at each stage by a folding back process, developed based on dynamic programming. An important different with lattice analysis is that paths do not recombine. The process starts from optimizing the system life cycle performance at the final stage, and goes backward until the initial stage where the expected life cycle performance is calculated. Babajide *et al.* (2009) applied such technique for evaluating a flexible oil platform design under uncertainty regarding reservoir compartmentalization and future oil price. Cardin *et al.* (2012) delivered another case study regarding Accelerator-Driven Subcritical Reactor (ADSR) technology, followed by the similar approach. As an innovative nuclear technology, ADSR is suggested to generate electricity by coupling a LINear Accelerator with a standard nuclear reactor core. Their case study demonstrated that design for flexibility could be valuable and worthwhile in nuclear technology.

In addition to techniques based on folding back principle, Monte-Carlo simulation is suggested as an effective alternative to evaluate flexible concepts. de Neufville *et al.* (2006) applied the approach to quantify a flexible design for a parking garage, compared to a rigid *n*-floor design. Designs are analyzed and compared under different scenarios of demand generated by GBM, instead of deterministic demand projections. The flexible design allowed for capacity expansion if the demand was not satisfied in two consecutive years. A simple logical "IF-THEN-ELSE" statement is a representation of managerial decision rules, which can be used to guide managers to response to an observation. The decision rule for implementing flexibility is not unique. Even for the same statement, the value of parameters can vary in a certain range (e.g., "two" in the last statement could be "one" or "three"), depending on expert recommendations.

As discussed in Subsection 2.1.2, decision analysis and binominal lattice do not account for multiple uncertainty sources. In particular, decision analysis is better suited for discrete events, while binominal lattice can handle most continuous uncertainty sources. Monte-Carlo simulation is capable of dealing with multiple uncertainty sources, no matter what type the uncertainty is. It provides more freedom for modeling decision rules, design variables, and parameters. However, exploring and evaluating many decision rules and design variables is difficult and time-consuming.

Computational efficient search

As the design space can be extremely large, designers need to efficiently search the whole or part of design space to determine the most valuable flexible design configuration in finite time. The procedures - including decision-based design, multi-attribute tradespace exploration, screen methods, and design category - help designers systematically and efficiently fulfill this target.

The basis of the decision-based design (DBD) was developed by Simon (1977), and the mechanism of evaluation and selection of designs was proposed by Hazelrigg (1998). This method is eventually applied for flexibility based on the framework of Olewnik *et al.* (2003) and Olewnik and Lewis (2006), which is an extension of Hazelrigg (1998). DBD is suited for the case that there are more than one metrics of interest excluding economics. However, developing models and searching representative scenarios can be time-consuming. The framework of multi-attribute tradespace exploration (MATE) was suggested by Ross (2006), exploring the design space to determine configuration based on decision-makers' utility attributes. Since the attributes normally are not single, a Pareto set is used for comparing candidate configurations. Hu and Poh (2010) proposed a Pareto set-based model to efficiently search the design space. Similar to DBD, MATE has difficulties in computational time regarding modeling and searching, and calculating expected utilities in design context. Screed methods were firstly introduced by Jacoby and Loucks (1972), relying on techniques such as optimization algorithms and design of experiments (DOE) to speed up the search procedure. de Neufville and Scholtes (2011) suggested three types of screen methods: top-down, bottom-up, and simulator. Top-down screen methods use the similar means as in system dynamics to represent the relationship between the parts of systems. Bottom-up screen methods use simplified model to describe the detailed one. Statistical tools and/or fundamental principles are used in simulators. Heuristic algorithms and simulation-optimization algorithms are applied for design space exploration. Relevant applications can be found in (Buurman *et al.*, 2009; Deng *et al.*, 2013; Hassan & de Neufville, 2006; Wang, 2005). To efficiently model the problem and search the design space, the model fidelity could be reduced. The shortage of screen methods is such loss in model resolution. Cardin (2007) proposed an approach called operating plan procedure, which explores the design space by selecting a small group of representative scenarios of uncertainty. The selected scenarios keep high fidelity so the model represents the real problem in a certain level of degree. This approach was applied in mining (Cardin *et al.*, 2008) and infrastructures (Cardin & de Neufville, 2013). Both screen methods and category methods cannot guarantee the global optimality of the solutions, due to limited time and computational resource. Moreover, DBD, screen methods, and category methods are not widely applied in industry (Cardin, 2014).

2.1.5 Phase V – Process Management

The procedures in this phase aim to address social and collaborative issues under which flexibility design and configuration is generated. The whole decision-making process involves many stakeholders with different backgrounds in the hierarchical structure. Although designers can embed flexibility into the system, managers may exercise flexibility in suboptimal timing or never exercise it due to many reasons such as the unawareness of the concept, budgetary constraints, and operating conditions. This ultimately affects the system life cycle performance as the embedded flexibility becomes a burden instead of an advantage. To address such issues, this phase studies how to create a comfortable condition for flexible concept generation, how to reduce barriers for exercising flexibility, and how to weaken the affects by agency problems and/or information asymmetries.

As an important methodology for studying agency behavior, game theory has received much attention from outside world since it was introduced by von Neumann and Morgenstern (1944) and Nash (1950). Game theory is considered as "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers" (Myerson, 1997). It is useful in addressing agency problems, capturing the dynamics of information asymmetries, and thus facilitating flexibility in design. This methodology has been explored recently and there exists opportunities for research in a further step (Dias & Teixeira, 2003). More specifically, it is demonstrated that the integration of real option and game theory – also known as option games – could be a valuable tool for dealing with multi-stakeholder decision-making problems with respect to flexibility in design, in the context of public infrastructures (Smit & Trigeorgis, 2009), enterprises (Ferreira *et al.*, 2009; Smit, 2001).

As an experimental methodology, serious gaming (also referred as “serious games” or “simulation gaming”) provides a research platform for studying the process management of flexibility in design. Serious games are designed for specific purpose rather than pure entertainment. The “serious” adjective is generally prepended to refer to products used by industries like defense, education, scientific exploration, heath care, emergency management, city planning, engineering, religion, and politics (GAIA, 2015). This technique can be helpful in engineering contexts for understanding the dynamics of decision-making, even though it is usually used in business school (Faria *et al.*, 2008). For instance, Cardin *et al.* (2013b) investigated how decision-making processes regarding siting fire stations affect the system performance of an EMS system in the Singapore context, firstly considering flexibility in design in serious gaming.

2.1.6 Thesis Positioning within Flexibility in Engineering Design

This thesis fits best within the phases of design space exploration and process management. The reason is that the thesis presents a general methodology to support the design for flexibility via mathematical programming tools in urban contexts. This methodology provides an opportunity to analytically and systematically investigate possible decision rules, and to determine the stochastically optimal initial configurations. As a means for exercising flexibility, managerial decision rules are simple, easy, and fairly intuitive to managers. These rules can help designers promote their flexible design concepts, and help managers take advantage of embedded flexibility to improve life cycle performance of the system. Even though mathematical models may have the issue on model resolution loss, the advantages on modeling and computational efficiency can be attractive in both academia and industry.

2.2 Design and Management of EMS Systems

Emergency services (or rescue services) are defined by the Wikipedia as the organizations that ensure public safety and health by addressing different types of emergencies. Emergency services are comprised of police, fire rescue services, and emergency medical services. Police departments focus on tackling crimes, and ensuring safety for both health and assets. Fire departments provide fire trucks and firefighters to deal with fire emergencies. Emergency medical services dedicate on supplying ambulances and professionals for out-of-hospital care. This section focuses on providing an overview of the design and management of EMS systems.

More specifically, the design and management of EMS systems is mainly concerned with the decision-making processes about siting ambulance/fire stations, and allocating/reallocating ambulances. These issues have been of interest to researchers over recent decades (Başar *et al.*, 2012). Savas (1969) suggested a standard four-step approach for systematically analyzing EMS systems. Owen and Daskin (1998), Brotcorne *et al.* (2003), Goldberg (2004), and Farahani *et al.* (2012) reviewed much of the body of research applying operations research methodology to solve resource allocation problems in the context of EMS systems. Past work majorly falls into two categories in terms of objectives: maximal covering location problems (MCLP), and set covering location problems (SCLP). Typical decisions in EMS systems can be categorized into strategic-, tactical- and operational-levels. Strategic-level decision-making processes care about where to site stations among a set of candidate sites, and how many vehicles/ambulances are needed for the system. Tactical-level decision-making processes focus on allocating and reallocating ambulances every tactical period (e.g., 3 to 6 months). Operational-level decision-making processes are concerned with daily incident response issues, i.e., sending available ambulances to respond to specific incidents/accidents. As the concept of flexibility was reviewed in Section 2.1, one may now ask "can the concept of flexibility be used for the design and management of EMS systems?" This question is best addressed before reviewing the literature regarding EMS systems.

This section provides an overview of design and management of EMS systems based on two types of location problems. More specifically, subsection 2.2.1 provides an overview of MCLP, while subsection 2.2.2 provides an overview of SCLP. Subsection 2.2.4 indicates where to situate this thesis best within the area regarding design and management of EMS systems.

2.2.1 Maximal Covering Location Problem

MCLP aims to maximize preferred goals subject to limited resources (e.g., facilities, budget, etc.). Goals can be any system target of interest, such as demand coverage, incident coverage rate, system service level, etc. MCLP is normally applied for the situation that one or a group of companies operate the EMS system, because the restrictions of resources are always highlighted during the design and management.

Short-term (single period) problems

MCLP was firstly proposed by Church and ReVelle (1974), referred to the following mathematical formulation of the design problem:

$$\begin{aligned} \text{maximize } \; & z = \sum_{i \in I} a_i y_i \\ \text{s.t. } \; & \sum_{j \in N_i} x_j \ge y_i \quad \text{for all } i \in I \\ & \sum_{j \in J} x_j = P \\ & x_j = (0,1) \quad \text{for all } j \in J \\ & y_i = (0,1) \quad \text{for all } i \in I \end{aligned} \tag{2.5}$$

where

$I =$ denotes the set of demand nodes;

$J =$ denotes the set of facility sites;

$x_i = 1$ if a facility is allocated to site j, and 0 otherwise;

$y_i = 1$ if demand node i is covered by facilities, and 0 otherwise;

$a_i =$ population to be served at demand node i;

$p =$ the number of facilities to be located.

In Problem (2.5), a demand node is a geographic region where emergency incidents arise. As in Toregas *et al.* (1971), it was assumed that the sites in the set N_i could respond to the incidents occurred in node i within predetermined time. Its target was to maximize the covered population in a single time period, subject to limited number of facilities. No uncertainty was considered in this model. Real world applications of Problem (2.5) can be found in Eaton *et al.* (1985) and Schilling *et al.* (1979).

Daskin (1983) extended MCLP into the Maximal Expected Covering Location Problem (MEXCLP), by considering facilities being unavailable when a call entered the service system. The model aimed at maximizing the expected number of covered demand covered by multiple times instead of the population. The objective function was written as the formulation:

$$\max \ \sum_{j=1}^{M}(1-p)\,p^{j-1}\sum_{k=1}^{N}h_k y_{jk}. \tag{2.6}$$

Differed from Church and ReVelle (1974), it was assumed that there could be more than one facilities at a node. The coefficient term $(1-p)p^{j-1}$ was the weight associated with the objective of maximizing demand covered by at least j times. Facilities were assumed to be identical and independent, i.e., the availability of each facility was independent with busyness probability p, and independent from candidate nodes. MEXCLP showed the benefits of considering multiple demand coverage for improving system performance; however, it does not consider multiple time periods.

Batta *et al.* (1989) revisited MEXCLP and relaxed three underlying assumptions of MEXCLP: facilities operated independently, facilities had the same busyness probability, and the busyness probability was independent from locations. By embedding the hypercube queuing model developed by Larson (1974), the adjusted MEXCLP used correction factors to relax the independency of facilities (or servers) availability. The formulation of correction factors is as follows:

$$Q(M,p,j)=\frac{\left[\sum_{k=j}^{M-1}\left\{\left(\frac{(M-j-1)!(M-k)}{(k-j)!}\right)\left(\frac{M^k}{M!}\right)p^{k-j}\right\}\right]}{\left[\left((1-p)\sum_{i=0}^{M-1}\left(\frac{M^i}{i!}\right)p^i\right)+\left(\frac{M^M p^M}{M!}\right)\right]}. \qquad (2.7)$$

The idea behind approximation (2.7) was that the probability of dispatching jth closest facility was proportional to the probability of $j-1$th closest facilities being unavailable and the probability of the target facility being available. As can be seen, MEXCLP was the special case when the approximation was equal to 1. In addition, Rajagopalan *et al.* (2007) replaced the original formulation of MEXCLP by non-linear terms, with the assumption of calls being non-uniformly distributed. The non-linear formulation used for replacing the inner summation term of formulation (2.6) was $h_j(1-p^{y_j})$. This modeling trick reduced the size of the problem, however, turned the problem into a non-linear one. Rajagopalan *et al.* (2007) applied meta-heuristic algorithms to solve the non-linear MEXCLP, due to the limit of linear solvers.

ReVelle and Hogan (1989a) extended the basic MCLP into a Maximal Availability Location Problem (MALP), by taking the system reliability into consideration. The concept of busy fraction proposed by Daskin (1983) was applied here as well. The reliability was represented by a chance constraint originated by Charnes and Cooper (1959), in which server availability was either uniform system-wide or not. Marianov and ReVelle (1996) proposed a queuing model to represent the availability neighborhood of a demand node. This neighborhood was modeled as an M/G/s-loss system, and the probability of at least one server being available in the neighborhood was shown as $1-\left(\frac{1}{p_{s-1}+su_i/\lambda_i}\right)p_{s-1}$. The term of λ_i is the Poisson intensity at node i, and p_s is the probability of system being in the state s. The queueing MALP model had a better approximation of system availability, and thus had a better performance. It was assumed that travel times were normally distributed in this model. Sorensen and Church (2010) introduced a hybrid model called Local Reliability Maximal Expected Covering Location Problem, which combined the objective of MEXCLP and the local busyness estimates of MALP. This hybrid model allowed for calculating a series of local $q_{i,k}$ values instead of a single s_i value for each node. The parameters $q_{i,k}$ and s_i represent the reliability of service at node i for different number of servers. Berman and Krass (2002) discussed a generalized maximal covering location problem with the consideration of partial coverage of customers. That is, a customer at a node could be covered at different levels of coverage by facilities, depending on the distance from the facility to the customer. Berman *et al.* (2013) studied three basic types of location problems: expected covering problem, robust covering problem, and expected *p*-robust covering problem, by considering uncertain travel time. The objectives aimed at maximizing the covered demand nodes under different conditions.

Hogan and ReVelle (1986) suggested an idea of backup coverage as a decision criterion. In addition to maximizing the number of demand nodes covered once, the Backup Coverage Problem further aimed at maximizing the number of demand nodes covered twice (i.e., the backup coverage). The objective function could be formulated as:

$$\max \quad Z = w\sum_{i \in I} a_i y_i + (1-w)\sum_{i \in I} a_i u_i . \tag{2.8}$$

The weight parameter w in the formulation (2.8) represented the preference between the first coverage and the backup coverage. Based on this decision criterion, Gendreau *et al.* (1997) proposed a Tabu search algorithm originated by Glover (1990) to handle a design problem, with the goal of maximizing double coverage subject to limited units of servers. Araz *et al.* (2007) developed a maximal covering location model based on fuzzy logic. The objective was to maximize the overall satisfactory level of individual goals (e.g., coverage based on different requirements), which were formulated by the fuzzy membership functions to represent the linguistic terms of “approximately greater than or equal to ” and “approximately less than or equal to”. Silva and Serra (2008) proposed a covering model to handle emergency calls with different priorities, incorporating priority queuing theory. The objective of Priority Queuing Covering Location Problem is to maximize the population covered by all priorities, as shown in formulation (2.9).

$$\max \quad Z = \sum_k \sum_i \sum_j a_i X_{ij}^{[k]}. \tag{2.9}$$

The assignment variable x_{ij}^{k} denoted whether a demand node i was assigned to a site j for priority k's urgencies. The values of all x_{ij}^{k} with different priorities were not necessary to coincide, which could be considered as a form of operational flexibility implemented by the optimizer.

Pirkul and Schilling (1991) introduced a capacitated covering model by considering workload limit on the facilities. Demand nodes now could only be responded to when facilities did not exceed the workload, and the system thus needed more facilities in order to have the same level of covered population as MCLP. Marianov and ReVelle (1992a, 1992b) discussed fire protection siting problems with capacitated stations, based on the framework of MCLP and MEXCLP. The problems considered the cases that a demand node covered by more than one type of servers. McLay (2009) introduced a Maximal Expected Covering Location Problem with Two Types of Servers: Advanced Life Support (ALS) and Basic Life Support (BLS). BLS can only provide basic medical care service and it is powerless if a Priority I incident occurs. ALS, by contrary, provides advance medical care service that can handle such Priority I incidents. The design issue was that the number of ALS was limited, and the response time for ALS was strictly less than a specific time. The objective function was formulated as follows:

$$\max \quad \sum_{i=1}^{n}\sum_{k_a=1}^{K_i^A} h_i W_{k_a}^{A} y_{i,k_a}^{A} + \sum_{i=1}^{n}\sum_{k_b=1}^{K_i^B}\sum_{k_{a=0}}^{K_i^A} h_i W_{k_b,k_a}^{B} y_{i,k_b,k_a}^{B}. \tag{2.10}$$

The objective was to maximize the expected number of Priority I covered in an amount of time. The performance of the system depended on the definition of Priority I and public recognition of such definition. Besides, Serra (1996) considered a problem where emergency service was coherent hierarchical. That is, regions corresponding to a given hierarchical level belonged to the same one region in the next hierarchical level. The concept of coherent hierarchy applied here was helpful for defining optimal capacities of the facilities because the coherent hierarchy existed in real world, leading to an applicable model.

Long-term (multiple periods) problems

To investigate long term life cycle performance, some past work focused on system-wide strategic-level decision-making dynamics by developing multi-period models. Schilling (1980) presented a Multi-objective Dynamic Location model that merged T static MCLP models in such a manner that future service levels were considered collectively. The multi-objective was to maximize the coverage at each time period. The outcome of Multi-objective Dynamic Location was different from the simple summation of T static MCLP models because the decisions were inherited over the life cycle. In other words, station was decided to be constructed at site j at time t would remain available at time $t = t + 1 \dots T$. Gunawardane (1982) introduced dynamic extensions to MCLP with deterministic demand incidents in a given planning horizon, referred to the following formulations:

$$
\begin{aligned}
\min \quad & z = \sum_{i \in I} \sum_{t=1}^{T} a_{it} y_{it}, \\
\text{s.t.} \quad & \sum_{j \in N_i} x_{jt} + y_{it} \geq 1, \\
& \sum_{j \in J} x_{jt} \leq p_t, \\
& x_{jt} \leq x_{j,t-1} \; j \in J_1, \\
& x_{jt} \geq x_{j,t-1} \; j \in J_2.
\end{aligned}
\tag{2.11}
$$

where

$J_1 =$ set of facilities 'open' at the beginning;

$J_2 = J - J_1$;

$a_{it} =$ population of demand center i at time t;

$p_t =$ the limit on number of facilities in period t;

$y_{it} = 1$ if demand center i is *not* served in period t, and 0 otherwise;

$x_{jt} = 1$ if j is 'open' in t.

Problem (2.11) aimed to minimize the uncovered population over life cycle. The goal was the same as maximizing the covered population, because the total population was fixed here. Differing from Schilling (1980), facilities in Problem (2.11) were allowed to be closed. However, a facility could not be opened again if it was closed in a previous period.

Schmid and Doerner (2010) developed a multi-period version of the

Double Standard Model considering time-varying coverage areas. The model maximized the demand covered by at least two vehicles in the long term and allowed vehicles to be reallocated throughout the planning horizon simultaneously. Başar *et al.* (2011) suggested a Multi-period Backup Double Covering Model with two response regulations, with the goal of maximizing the double covered population over life cycle in Istanbul. The assumption made by Schilling (1980) was applied here as well. In general, for such long term planning, the number of facilities that was allowed for operating at time t was limited. Work discussed above was deterministic, i.e., no uncertainty was considered.

2.2.2 Set Covering Location Problem

SCLP, in contrast to MCLP, aims at minimizing targets subject to specific requirements such as system coverage, service level, etc. The targets of interest can be economic (e.g., total costs) and non-economic (e.g., travel distances). SCLP is basically applied for the case that the government or public organizations operate the EMS system. To achieve a high level of coverage (e.g., 90% to 95%), the service supplier would like to fulfill this goal with minimal costs.

Short-term (single period) problems

Toregas *et al.* (1971) developed the first deterministic SCLP for the design of EMS systems, referred to the following formulations:

$$
\begin{aligned}
\min \quad & z=\sum_{j=1}^{j=n} x_j \\
\text{s.t.} \quad & \sum_{j \in N_i} x_j \geq 1, \\
& x_j=(0,\ 1).
\end{aligned} \tag{2.12}
$$

where

n = denotes the number of demand points;

$x_j = 1$ if a facility is established at point j, and 0 otherwise;

$N_i = \{j | d_{ji} \leq s\}$.

Problem (2.12) firstly assumed that an urgency at node j could only be responded by a facility on time if the distance between the node j and the facility was less than or equal to s. This assumption was widely accepted and applied in later work; see, for example, Church and ReVelle (1974), Daskin (1983), Ball and Lin (1993), etc. The deterministic SCLP aimed at minimizing the total number of facilities for operating the system.

Chapman and White (1974) developed an early probabilistic version of SCLP by taking system-wide busy fraction into consideration. Aly and White (1978) considered the traveling time as the uncertainty. Instead of defining a response zone, the demand point could be responded on time anywhere if the traveling time was no greater than a specific upper bound. The chance constraint $Pr(t_{ij} \leq t_i) \leq \gamma_i$ represented the requirements on the system service level, i.e., the probability of response time for region i being less than or equal to the upper bound should be no more than required service level γ_i. Revelle and Hogan (1989b) suggested an enhanced probabilistic model by considering a detailed average busy fraction of vehicles. This average busy fraction is shown as:

$$q_i = \frac{\bar{t} \cdot \sum_{k \in M_i} f_k}{24 \sum_{j \in N_i} x_j}. \tag{2.13}$$

where

$\bar{t}$ = the average duration of a call (hours);

f_k = frequency of calls at demand node k (calls/day);

M_i = the set of demand nodes within S of node i.

This idea of busy fraction was interpreted as the reliability of the system, and higher busy fraction indicated lower system reliability. In addition to Revelle and Hogan (1989b), Ball and Lin (1993) introduced a reliability constraint on the number of vehicles to guarantee that emergency incidents could be responded to with a given probability p. The demand here was generated according to a new better than used distribution. That is, the probability of a stochastic system will survive addition t time when it has survived for s time, is no greater than the probability of a new system will survive t time. In contrast to previous work, this reliability model aimed at minimizing the total costs for operating the system. To clarify, the definition of reliability here is fairly different from the definition in quality engineering contexts. Reliability describes “the ability of a system or component to function under stated conditions for a specific period of time” (Institute of Electrical and Electronics Engineers, 1990). The concept of reliability in EMS systems is better interpreted as availability.

Beraldi *et al.* (2004) proposed a stochastic formulation of SCLP where the probabilistic constraint was formulated as $P\left(\sum_{j\in N_i} x_{ij} \geq \xi_i\right) \geq p$, with the goal of minimizing the costs. The random variable ξ_i denoted the service request at node i, and x_{ij} denoted the number of vehicles located at site j for responding to request from node i. It was assumed that requests could be fully responded if the number of vehicles was no less than the number of requests, regardless the density of the requests. Beraldi and Bruni (2009) subsequently extended the work of Beraldi *et al.* (2004) to a new stochastic programming paradigm incorporating representative scenarios, based on the same assumption on the request response. The availability or reliability of the system was defined as the proportion of representative scenarios being covered. In other words, if the system availability was required to be 0.95, service requests in at least 95% of the scenarios needed to be fulfilled.

Queuing theory was applied for SCLP to study the stochastic processes, as for MCLP. Marianov and ReVelle (1994) considered the service system as a *M/M/s*-loss queuing system. The probability of all servers being busy could be derived from the following equation:

$$p_s = \frac{\frac{1}{s!}\rho_i^s}{1+\rho_i+\frac{1}{2!}\rho_i^2+\ldots+\frac{1}{s!}\rho_i^s} \leq 1-\alpha. \tag{2.14}$$

The left hand side of formulation (2.14) was actually the Erlang B formula, which was applied mostly in telecommunication systems. The parameter

$1-\alpha$ is the requirement on system reliability. Marianov and Serra (2002) used the formulation (2.15) to express the cumulative distribution function of the waiting time in a manner of *M/M/m* queue system.

$$P\left(w_j \geq \tau\right) = e^{-\mu\tau}\left[1+\frac{p_0\rho^m}{m!\left(1-\rho/m\right)}\left(\frac{1-e^{-\mu\tau(m-1-\rho)}}{m-1-\rho}\right)\right], \qquad (2.15)$$

where τ is the bound for the waiting time that cannot be exceeded by α percent of times. This probability was independent from p_0, but only on m, μ, ρ, and τ. In addition, Baron *et al.* (2009) suggested a location problem with stochastic demand and congestion. It studied the system stability condition and the lower bounds on the system availability, based on partially accessible queuing system. The model used a server availability constraint to guarantee the service quality of the mobile servers. Noyan (2010) considered a similar problem as Baron *et al.* (2009) in a single-stage stochastic model, integrating a chance constraint and a stochastic dominance constraint. Besides, Pirkul and Schilling (1988) suggested a set covering model incorporating the concept of backup coverage and capacitated workload, which were applied in MCLP as discussed in subsection 2.2.1. Marianov and Serra (2001) proposed a model for a hierarchical EMS system, with the goal of minimizing the cost of locating high-level and low-level response centers. Jia *et al.* (2007a) discussed covering problems with respect to large scale emergencies. The number of demand nodes and facilities could be hundreds or even thousands. Typical solution approaches could not efficiently address such large-scale problems.

Jia *et al.* (2007b) then provided three heuristic algorithms for finding optimal solutions. The heuristics were also applied to sequential decision-making processes (i.e., long-term problems).

Long-term (multiple periods) problems

Similar to MCLP, some work was developed based on SCLP from a long-term perspective. Gunawardane (1982) extended a deterministic single-period SCLP to a multi-period SCLP, with the assumptions of stations being open cannot be closed, and vice versa. Daskin *et al.* (1992) looked into a dynamic location problem in order to find a planning horizon τ^* and a corresponding initial design configurations X^* for at least one optimal decision policy. The initial configuration was suited for planning horizons that were less than or equal to τ^*. The model was developed under the same assumption as in Gunawardane (1982). Zarandi *et al.* (2013) proposed a large-scale dynamic location model by relaxing this assumption. The proposed model used for dealing with large-scale systems where the number of facilities and demand nodes could be thousands.

Farahani *et al.* (2009) studied a multi-period location problem by considering multiple reallocation opportunities for one facility in a discrete planning horizon. The weight associated with a demand point was a function of traveling time. Different types of distances were taken into account in the analysis. Ghaderi and Jabalameli (2013) proposed a budget-constrained model to study the multi-period SCLP, with the objective of minimizing the traveling and operational costs over the time horizon. The objective costs were not discounted over time though. The purpose of the budget constraints was to investigate the influences by real-world investment policies on the facility location and network design. For instance, whether or not unspent budget would be used in the next period was one consideration. Ghaderi and Jabalameli (2013) proposed a hybrid Simulated Annealing (SA) heuristic for solving the problem because of the complexity of the problem. Binary variables were mainly solved by Branch and Bound (B&B) algorithm, and the budget-constrained variables were systematically solved by heuristics.

2.2.3 Miscellaneous Problems

Apart from MCLP and SCLP, there are much work focusing on the design and management of EMS systems. These studies, however, cannot be simply categorized into MCLP or SCLP in terms of the objective functions. Such work is discussed in this subsection.

Compared to single objective, multi-objective problems aim to find out solutions that are fit for different requirements, and some requirements may contradict. For instance, minimizing the number of facilities and maximizing the service level or covered demand nodes contradict in general. The solution for this combined objective is somewhere between the solutions for separate objectives. Alsalloum and Rand (2006) suggested a mathematical model with two goals that is introduced above. The objective function was expressed as a goal constraint in a goal-programming framework:

$$\sum_{i=1}^{n}\sum_{j=1}^{m} a_i P_{ij} Y_{ij} + d_0^- = 1. \tag{2.16}$$

Chanta *et al.* (2011) proposed a bi-objective model for the design of EMS systems in rural area, where the second objective was not unique. This second objective could be one of the three options: minimizing the maximum traveling distance, minimizing the uncovered rural zones, and minimizing the uncovered zones. Regardless of the second objective, the main goal of the model was to maximize the expected number of covered requests.

Badri *et al.* (1998) discussed a multi-objective location problem with respect to fire stations. The model incorporated several objectives, such as, minimizing the total costs, maximizing the service of area required most, minimizing the traveling distance, minimizing the traveling time, etc. Some objectives could be represented in a similar way, like minimizing the traveling distance and traveling time.

As part of decision-making processes, ambulance deployment and redeployment has received much attention recently. Gendreau *et al.* (2001) discussed a redeployment problem by considering several penalty cases. These cases were considered as constraints that could not be violated. Rajagopalan *et al.* (2008) proposed a dynamic redeployment model, with the goal of minimizing the number of servers (e.g., ambulances, fire trucks) deployed. Yue *et al.* (2012) introduced a simulation-based approach based on greedy algorithm for the dynamic ambulance allocation and redeployment problem. Sung and Lee (2012) proposed a modeling framework for EMS system design and examined different dispatch and redeployment policies, such as locally nearest one, shortest travel time, dynamic assignment, etc.

The ambulance redeployment is also known as the dispatch policy for ambulances. McLay and Mayorga (2011) examined the impacts introduced by the dispatch policy on a maximal covering location problem. McLay and Mayorga (2013b) proposed a model based on Markov Decision Processes (MDP) for optimally dispatching ambulances to requests, with classification errors in patient priorities. The classification error is one that the patient true condition is misclassified. This error always exists because patients may not have the ability or professional knowledge to classify their conditions correctly. McLay and Mayorga (2013a) suggested a MDP model for service-to-customer systems by considering four types of equity inside the model - two of which were related to customers and two of which were related to servers. As a critical factor when deciding how to allocate public source, it is a challenging to balance both efficiency and equity. The proposed model considered these equity as compulsory constraints, and reflected them using MDP. The details could be found in McLay and Mayorga (2013a).

Moreover, Rekik *et al.* (2013) developed a decision support system for EMS systems, embedding strategic-, tactical-, and operational-levels of decision-making processes. The decision support system was based on a hierarchical structure, where strategic-level decisions were on the top and operational-level decisions were at the bottom. The objectives of decision-making processes at each level were optimized according to their unique requirements. To the strategic-level, for instance, the goal is to minimize the number of stations. The decision support system was built for reflecting the real-world EMS system, and it thus had a high complexity.

2.2.4 Thesis Position within Design and Management of EMS Systems

The proposed approach intends to offer decision makers an opportunity to design the EMS system incorporating the concept of flexibility. It focuses on the strategic and tactical levels decision making processes and uses decision rules to analyze and exercise the flexibility. Operational level decision making processes however are not the concern. This thesis thus fits best within the area of long-term (multiple periods) design problems regarding EMS systems.

2.3 Siting Problems of Nuclear Power Plant

According to Wikipedia, a nuclear power plant (NPP) is one type of thermal power stations where the heat source is a nuclear power reactor. The nuclear reactor is a fundamental device to release nuclear energy, and thereby generate electricity. Nuclear energy or nuclear power is known as a green energy for the generation of electricity with low carbon emission, along with other sustainable energy source (e.g., solar, wind). Nuclear power has the ability to stably generate a large amount of electricity, which is the main advantage of nuclear power versus other sustainable energy source. Solar energy and wind power, by contrast, could only steadily generate electricity under specific environmental conditions. Hydropower may have the ability to stably generate electricity, but the installation of hydropower plants and dams seriously breaks ecological equilibrium alongside the water source (Walsh, 2007). In addition, the capacity of a single nuclear reactor is 600 – 1400 MW, which is much greater than the capacity of a wind turbine or a solar panel. The average capacity factor of a nuclear power plant is as well typically higher than other power plants such as wind farms and hydropower plants.

As for now, nuclear power is one of the two major sustainable energy source for the generation of electricity (the other one is hydropower). The international atomic energy agency (IAEA) reports that there are 440 operational nuclear reactors in 30 countries in 2015 (Power Reactor Information System (IAEA), 2015). This number will be increasing in the near future due to the growth in electricity demand, even though some debate about the use of nuclear power remains ever since this energy source has been used for electricity generation. The debate is focusing on safety related issues of using nuclear energy, severe nuclear accidents, as well as waste disposal and management. Although nuclear power is typically considered as safe, nuclear accidents with extremely low probability such as the ones in Chernobyl and Fukushima can be disasters to the world. Such accidents have profound implications for human health, environmental protection, industrial production, etc. A disaster anywhere around the world may impact policy making in any country. For instance, Germany and Switzerland planned to gradually shut down all NPPs in operation till 2025 in the wake of the Fukushima disaster, while China and India delayed their construction of new NPPs (Joskow & Parsons, 2012).

As part of the design and management of NPP, the siting problem has received much attention from society including academia and industry. Grimston *et al.* (2014) suggested six types of factors that could be considered in the siting of NPPs: Radiological and safety factors, Economic factors, Technical factors, Social factors, Environmental factors, and 'Political' factors. Again, one may ask a similar question as in Section 2.2 regarding the possibility of improving system performance in terms of total costs. For instance, one may ask: "could the concept of flexibility help improve the expected performance of a nuclear power plant system?"

This section provides an overview of economic assessment of siting problems for NPPs under uncertainty. The existing research that applying real option analysis as a means to deal with uncertainty in the siting problem is discussed. This section also provides an overview of social acceptance recognized in the context of nuclear energy systems.

2.3.1 Uncertainty in Investments on Nuclear Power Plants

Uncertainty is inevitable in the siting of NPPs, as in any other siting projects of large-scale industrial facilities (e.g., EMS systems). Kessides (2010) identified economic uncertainty and risks in fundamental elements regarding investments on nuclear power. More specifically, Kessides (2010) proposed a framework to identify underlying uncertainties and quantify their impacts on the costs of nuclear power. In general, the costs consist of four major components: construction costs, operations and management (O&M) costs, fuel costs, and back-end costs (Joskow, 2006). The costs of nuclear power are driven by high up-front construction costs (or capital costs), which approximately represent 60% of the total costs. Fuel costs accounts for 20% of the total costs, while O&M plus back-end costs account for the rest 20% (OECD/NEA, 2003).

Uncertainty in construction costs

Construction costs contain any costs incurred during the stages of "planning, preparation and construction of a new nuclear power plant" (Kessides, 2010). However, there is no internationally unified definition of construction (capital) costs of nuclear power plants, even though much work has been done with regard to standardize the costs of nuclear power over past four decades (DTI, 2007). There are several reasons that influence the estimation of construction costs. Firstly, not all the relevant data required for the estimation are available. Du and Parsons (2009) emphasized that collecting market data for both projects completed in recent years and under implementation from developed and developing countries (e.g., US, Japan, France, China, India) can be really helpful for such estimation. Although construction time and operating performance is open to the public (Rothwell, 1998), construction costs and other costs data are not available except for the commercial US nuclear fleet (Hultman *et al.*, 2007). The possible methods for such "blind" estimation could only rely on adjusted data announced by the government and/or the bidding data provided by vendors. The government however due to political and nationalistic reasons may present optimistic cost estimates. Vendors also have incentives to present biased contract price for winning the bid. All these make estimating the construction costs difficult.

Secondly, the construction of a nuclear power plant requires a large amount of on-site engineering. On-site engineering is a notorious barrier for economic assessment of large-scale projects, and it accounts for a significant proportion of the construction costs of nuclear plants (Thomas, 2005). Thirdly, a major part of construction costs heavily depends on the type of reactor chosen by system designers, and the selection of nuclear reactors ultimately influences the costs. There are four generations of nuclear reactors until now, and Generation IV is still under development (Grimston *et al.*, 2014). Generation I is the early prototype of power reactors, such as the Magnox power stations in UK and the Fermi 1 power stations in US. Generation II nuclear reactors that are most common reactors in use include pressurized water reactor, boiling water reactor, advanced gas-cooled reactor, and so on (Grimston, 2006). PWR and BWR are cooled by water during the operation, while advanced gas-cooled reactor is cooled by gas instead. A Generation III reactor is an improvement of any Generation II reactors done by improving fuel technology, thermal efficiency, passive safety, etc. The first Generation III reactor used for generating electricity is Kashiwazaki in 1996. The selection of reactors also influences the geological consideration and thus influence construction costs, as well as back-end costs within O&M costs.

Uncertainty in O&M costs

Operations and management costs discussed here contain typical costs associated with administration, management, support and upkeep of a power plant plus back-end costs. The costs of license and regulatory compliance is independent from the capacity of a NPP, while the costs of planned maintenance, insurance, contractor services, security, and corporate overhead are dependent on the scale of a NPP. The above costs included in the O&M costs can be considered as the known costs once the capacity of a NPP is determined.

On the other hand, back-end costs, the costs of human health and environmental protection, and the supply and maintenance costs of unplanned shut down of reactors are uncertain in the O&M costs. Back-end costs include the costs related to decommissioning and dismantling of nuclear facilities at the end (or even middle) of their operating time, and the long-term costs related to disposal and management of radioactive waste (Joskow, 2006). The costs related to decommissioning and dismantling could be understood as depreciation costs, depending on the scale and the age of nuclear power plants. This cost is estimated as 10-15% of the construction costs (Rose, 1985). The costs related to disposal and management of radioactive waste however is dependent on “the sequence and timing of various stages of the program” (World Nuclear Assocation, 2015). The lifetime of a typical Generation II reactor is 30 or 40 years, but some reactors are being life-extended to 50 or 60 years in United States. The actual operating time is longer than the anticipated life time in this case, and the costs for disposal and management of radioactive waste thus are smaller than the expected due to decreasing radioactivity.

In addition to the costs discussed above, there is another uncertain costs associated with the unplanned shut down of power generators. For ease of maintenance, a nuclear reactor is planned to shut down regularly and it does not influence the output of electricity for the power plant. The unplanned shut down can be caused by human mistakes or any issue regarding the nuclear reactor. This can affect the output of electricity, as the redundant device needs a while for warming up and generating electricity. Between this time gaps, the output of electricity of the power plant is definitely lower than the anticipated output. In some electricity market (e.g., UK), this short of the contracted sales can cause a lot of expense for buying electricity from the competitors (Steer *et al.*, 2011) Additional supply and maintenance costs are required accordingly as well.

Uncertainty in fuel costs

Rothwell (2009, 2010) suggested that fuel costs could be inferred from public information including reactor technology, nuclear fuel fabrication, the length of refueling outage, etc. Most nuclear fuels contain fissile elements such as uranium-235 (^{235}U), plutonium-239 (^{239}P), plutonium-241 (^{241}P), etc. The quantity of these fissile elements required by reactors are different, as well as the type of fissile elements. A typical Pressurized Water Reactor would have about 150-250 assemblies of 200 to 300 rods each, and the reactor would require 80-100 tons of uranium in all glass (Glasstone & Sesonske, 1994). A modern Boiling Water Reactor, by contrast, would have 800 assemblies of 74 to 100 rods each, requiring 140 tons of low-enriched uranium. The fuel costs heavily depend on the selection of nuclear reactors.

2.3.2 Real Options in Nuclear Power Systems

As discussed in Subsection 2.3.1, uncertainty exists in nuclear power systems and affects the investment assessment. To pro-actively deal with uncertainty, real options theory (or flexibility in design) is considered as a means for economic analysis and improvement in nuclear power systems design and management. Cavender (2011) provided an overview of the methods for economic analysis in valuing nuclear power plants under uncertainty, including the discussion and comparison between typical DCF and ROA. DCF is an economic method for valuing a project by using the ideas of the time value of money, but it does not handle uncertainty well, even though it is widely used in cash flow analysis. This is because in traditional DCF (e.g., standard NPV rule) approaches, an expected or most likely scenario is projected and used to model the cash flows. The approach presumes of management's passive commitment to a rigid deployment strategy, which is not an adequate representation of reality. For example, a capital project may be initiated immediately and operated continuously at the same scale at the beginning until the end of the life cycle, although in reality managers may seek to adjust capacity as needed over time, so as to adjust to changing needs. On the other hand, many real world applications showed that the ideas of flexibility could improve the expected life cycle performance by 10-30% on average (de Neufville & Scholtes, 2011). However, design for flexibility is not fully accepted by researchers in the context of nuclear technology due to several misconceptions introduced by Martinez-Cesena *et al.* (2013):

1. Real options theory is similar to a black box that is difficult to understand it really works without strong background in mathematics and/or finance;
2. Real options theory is one tool used to enlarge the value of projects, despite the projects being flexible or not;
3. Real options theory prefers risky projects (designs or systems) rather than safe ones;

4. Real options theory is only applicable for tradable assets;
5. Real options theory may be a good idea conceptually but do not work practically.

However, these misconceptions are not difficult to address. For instance, there is much research meant to provide accessible tools for applying real options under a wide range of circumstances, of which one approach is the use of managerial decision rules, as proposed in this thesis. In addition, real options theory only develops value within a flexible project or system. If the project is inflexible, real options theory cannot bring any additional value at all. Moreover, if the value of flexibility for a relatively safe project is not favorable or even negative, that indicates that design for flexibility is not necessarily needed in this particular case. As can be found in de Neufville and Scholtes (2011), there are some real cases applying real options theory that are not tradable.

In the broader literature, Louberge *et al.* (2002) investigated an optimal stopping problem with respect to geological disposal of nuclear waste using a real option model. The embedded flexibility is the option of switching from surface storage to deep geological disposal. The decision-making processes was represented and modeled by a GBM, where the objective is to minimize the expected net present value of waste management, under uncertainty of costs of future accidents, institutional control, and hazard management. Kiriyama and Suzuki (2004) proposed a real option model to analyze an optimal stopping problem under the uncertainty of CO_2 emission by considering the phasing flexibility. The model was a simplified diffusion model based on the work of Pindyck (2000) by replacing the original technical uncertainty with the CO_2 emission, and it was solved by DP for determining the optimal timing to construct new a power plant in order to reduce the rate of CO_2 emission in excess of the cap. Gollier *et al.* (2005) examined a phasing flexibility through a real option, where the capacity of nuclear power plants deployed at different times were not necessarily the same. The model aimed at determining the optimal timing for investing the first module and the sizable effect on the value of modularity. Siddiqui and Fleten (2010) examined the value of a staged commercialism program for a unconventional energy technology by taking the real option approach, motivated by concerns about CO_2 emissions. Jain *et al.* (2013) focused on the small and medium size reactors (SMRs) and investigated the economic impact of modular construction of such reactors. Locatelli *et al.* (2015) discussed a load-following problem with SMRs and demonstrated its viability based on the real option approach.

In addition to DP and DCF, simulation is considered as an important tool in analyzing the value of flexibility in nuclear power systems. Rothwell (2006) considered three uncertain sources in valuing the investment of new nuclear power plants using a real option approach: price volatility, capacity factor risk, and cost volatility. The capacity factor of a power plant is defined as a ratio of the output of a specific power plant, and it shows great variation during the life cycle of a power plant (Du & Parsons, 2012). These uncertain parameters were estimated based on available real data and sampled in the Monte-Carlo simulation for sensitivity analysis. The target system was a dual-unit advanced Boiling Water Reactor, where managers had the option to decide whether or not to invest the second reactor. Abdelhamid *et al.* (2009) also analyzed the deferral option in siting the first nuclear power plant in Tunisia using the similar approach. Zhu (2012) established a simulation-based model for the economic assessment of investing nuclear power plants in China. Several uncertain sources, such as technological and economic uncertainty, were taken into account in the analysis. Uncertainty factors were explicitly modeled using different mathematical models, and the proposed model was solved by Least Squares Monte-Carlo simulation. The results showed that Generation III reactors were not worth investing, unless the total investment costs of such reactor could reduce to 1.2 times to that of current domestic reactors and the electricity price increase by 30%. Besides, Cardin *et al.* (2012) investigated an innovative nuclear technology (i.e., ADSR) by considering the uncertainty of unplanned shut downs caused by nuclear reactor cores, and then solved the real option model via decision tree analysis.

2.3.3 Social Acceptance of Nuclear Power

Social acceptance or public acceptance can be understood as “essential for any activity that affects large sectors of a nation" (Golay, 2001). The applied objects of social acceptance can be human beings, abstract concepts, real technologies, etc. For nuclear power, the corresponding social acceptance is represented as the public support for maintaining existing NPPs and escalating more NPPs nationwide. Abrecht *et al.* (1977) discussed some ethical issues that can affect the public acceptance of nuclear power, including the public appraisal and risks of nuclear technology. More specifically, these issues contained nuclear waste disposal, catastrophic accidents, releases of radioactive substances, and the threat of nuclear weapons. This acceptance is obviously regional, because the factors that could influence it are different across the world, such as the national economic condition (e.g., Gross Domestic Product (GDP)), recognition of current nuclear technology, etc. Kidd (2013) introduced that some western countries, such as the USA and UK, were not obviously constrained by public acceptance issues, as other developed countries like Germany and Switzerland. The public acceptance nevertheless still plays a role in affecting the capital investment cost, which is currently the major problem for the escalation.

Among the above-mentioned factors, historical local and/or foreign nuclear events are ones that could significantly reduce the public support for nuclear power. The Chernobyl disaster in Ukraine was the worst NPP accident since nuclear power had been used as an alternative energy source, in terms of cost and casualties. This catastrophe ultimately cost 18 billion rubles and involved over 500,000 workers, while 31 people died during the accident itself (Gorbachev, 1996). Thirteen countries was contaminated by radioactive substances, and 985,000 premature deaths as a result of radioactive released was reflected between 1986, the year of the accident, and 2004. After the Chernobyl disaster, the escalation of NPPs had been slowed down significantly across the globe until 2005. In 2011, the other disaster at Fukushima Daiichi NPP in Japan received much attention from the outside, as it was considered as the second worst accident in terms of the level of radioactive materials. The accident occurred when the power plant was hit by a tsunami, and it resulted in a nuclear meltdown of three nuclear reactors (Wakatsuki, 2014). 300,000 people evacuated the contaminated zone and about 18,500 people died during the evacuation in this accident due to the earthquake and tsunami (Aliyu *et al.*, 2015). The estimates of the economic losses (including the total disaster and consequence) range from $250 billion to $650 billion US dollar (Lavelle, 2012). Influenced by the Fukushima accident, Germany immediately closed 8 oldest NPPs, and will shut down the remaining plants gradually till 2025 (Joskow & Parsons, 2012). Switzerland government made similar decisions at the same moment under this circumstance. Other countries like the USA, Sweden, China, and India delayed or did nothing with ongoing projects even though they were confronted with social acceptance as a key issue when they made decisions regarding nuclear power (Mishra, 2012). The South Korean government just continued to operate 20 existing NPPs and will establish more NPPs by 2015 (Song *et al.*, 2013).

For technology recognition, Golay (2001) questioned the proposition that creating demonstrably safer technology could gain the social acceptance of nuclear power. The original argument would be true if the safer technology is demonstrated with satisfied performance by trustable organizations or communities. The exact social acceptance gained by the corresponding technology, however, is uncertain and difficult to quantify. In addition, the social acceptance of nuclear power does not vary directly or inversely with the GDP. The United States had the highest GDP around the world in 2008, but India had the highest public acceptance index at that moment (IAEA, 2008). An inverse conclusion related to the USA and Germany can be drawn according to the same source. There might be some relationship between the social acceptance of nuclear power and the GDP. Such relationship, however, has not been confirmed yet.

International Nuclear Events Scale (INES)

As discussed above, current literature shows that the historical data of nuclear events could be an applicable tool for some countries to estimate the future social acceptance of nuclear power. For instance, the social acceptance of nuclear power in Germany is obviously affected by the recent nuclear disasters, while countries like the USA and UK are not. Despite other factors such as the GDP and technology recognition, one may consider to project the social acceptance by taking into account the historical data of recent nuclear disasters in a numerical way. That is, qualifying the safety significance of nuclear events (incidents or accidents) first before going to further steps, such as regression analysis and hypothesis test.

To enable prompt communication of safety significance to the public, the International Atomic Energy Agency introduced the International Nuclear Events Scale (INES) in 1990 (World Nuclear News, 2015). Similar to the Moment Magnitude Scale used to measure the earthquake, the INES is also designed to be logarithmic such that the severity of an event is approximately ten times greater for each increase in the level of the scale, which aims to "keep the public as well as nuclear authorities accurately informed on the occurrence and consequences of reported events" (International Atomic Energy Agency, 2015). There are eight levels in total on the INES scale. For seven nonzero levels, three of them are incident-levels and the rest are accident-levels. IAEA also provided at least one example for each level in order to facilitate the qualification of nuclear events.

As an important indicator suggested by IAEA, INES has been applied for studying the risk perception of a nuclear power plant in contexts of different countries. Huang *et al.* (2013) studied how the Fukushima accident impacts on the risk perception of residents near nuclear power plants in China. Four perception factors (i.e., knowledge, perceived risk, benefit and trust) were considered in a structural equation model, and three levels of nuclear events (i.e., level 1 to 3 of INES) were taken into account for assessing the median public acceptable frequencies. It showed that this accident in Japan had significant impacts on risk perception, and the most sensitive groups of residents were those not in public service, those with low income, and those living near the plant. Research participants were required to go through survey for gathering data where INES provided them better understanding on the classification of nuclear events. Similar study of He *et al.* (2013) showed that the public in China did not lose trust in government authorities (not include state-owned enterprises), even after the Fukushima accident. However, such trust was waning due to information asymmetries (i.e., lack of transparency) and information incompleteness (i.e., lack of information sources). A strategy that aimed at dealing with those issues would be developed in the near future by the Chinese government. Case studies in both papers were with respect to the nuclear power plant at Lianyunguang and Haiyang, respectively, which were relatively new candidates for establishing nuclear power plants. Besides, Webb *et al.* (2006) reviewed the nuclear events over 50 years at the Sellafield nuclear installation in England, and accessed past events referring to the levels of INES in a unified system of off-site impact rating. Each event was noted in a specific INES level with detailed description, helping people better understand the INES User's Manual. This study too demonstrated that the current INES rating scheme was probably applicable for various radiological events.

2.3.4 Thesis Position within Siting Nuclear Power Plants

The proposed approach is applied in a deployment problem with regards to nuclear power plants. To the best of our knowledge, no existing work considers the uncertain factor of social acceptance of nuclear power technology when evaluating the investment of siting nuclear power plants. This factor, however, is valuable for consideration as the social acceptance could affect the decision-making processes, such as the capacity deployed at a plant initially and the time for expanding its capacity over the life cycle. The social acceptance within the second application is simplified, being expressed in the INES rating scheme. It is assumed that decision-making processes of this siting problem will be significantly affected by such acceptance. That is, if the social acceptance is too low to tolerate, existing plants will be shut down immediately. It is further assumed that existing plants will not be allowed to extend their service life if the social acceptance falls into a medium level of severity, even though they are permitted to keep operating. Therefore, this thesis fits best within the long term siting problem of nuclear power plants incorporating strategic-level flexibility.

2.4 Research Opportunities

2.4.1 Related to Design and Management of Infrastructure Systems

There is an opportunity to enhance the current design and management of infrastructure system in terms of the long term (e.g., 10+ years) life cycle performance by incorporating the idea of strategic-level flexibility. The opportunity arises because this idea performs well in the design of typical infrastructure systems, and the strategic-level decision-making processes of infrastructure systems are similar to those of typical ones. The goal in this case is to find appropriate flexible design alternatives to improve expected long term life cycle performance of infrastructure systems, as demonstrated in many case studies of typical infrastructure systems (de Neufville & Scholtes, 2011).

Table 2.1 provides some examples of designs in the literature regarding an urban system that fits along the four features of interests. The first column on the left shows the example studies, and the other four list the features. As can be seen, very few of them cover all these features and some of them even are not infrastructure systems. Except for Cardin and Hu (2016), there is not much work that addresses all the engineering system features considered. This is where this thesis hopes to make a contribution. Also, the concept of flexibility is not well applied in the infrastructure system sector.

A typical infrastructure system consists of some infrastructures (e.g., stations, power plants) and/or multiple network links (e.g., roads, power grids). The system aims at distributing corresponding resources - such as emergency vehicles and electricity – to the customers in order to satisfy certain predetermined requirements. Uncertainty sources like customer demand and supply volatility significantly affect the expected system performance in the long term. Plenty of work demonstrates that design for robustness has considerable impacts on dealing with uncertainty in the downside; see, for example, Farahani *et al.* (2012). The typical designs, however, cannot change themselves according to the external situations, and thus lose the opportunities in the upside. The flexible design, in contrast to typical designs, has the ability to pro-actively deal with uncertainty, and adapt to changing future conditions.

Table 2.1 Example studies of different designs aiming at infrastructure systems.

Example studies	**Design Features**				
	System details	infrastructure system	Long term life cycle	Strategic-level flexibility	With uncertainty
Tseng and Graydon (2002)	Power plant	✓		✓	✓
de Neufville *et al.* (2006)	Parking garage	✓	✓	✓	✓
Guma *et al.* (2009)	Real estate	✓	✓	✓	✓
Urich and Rauch (2014)	Urban water infrastructure	✓	✓		✓
Zhou *et al.* (2015)	Electric power system	✓			✓
Cardin and Hu (2016)	Waste-to-energy plant	✓	✓	✓	✓
This thesis	EMS and nuclear systems	✓	✓	✓	✓

2.4.2 Related to Real Option Analysis

There is a clear opportunity to enhance and complement existing approaches used in real option analysis to help identify best flexible systems design concepts by means of performance quantification (i.e. economic or others). This can be done by developing a framework systematically for valuing the flexibility under uncertainty, by determining the stochastically optimal initial configuration and the best implementation plan. Standard ROA methods may be difficult to use for the evaluation of flexibility in infrastructure systems for the reasons discussed below.

Typical DCF valuation methods do not account well for inevitable change of uncertainty drivers during the lifetime of an engineering system. It is usually assumed that the deployment path or management strategy is determined at $t = 0$ over the long term life cycle of the system. However, the uncertain factors keep changing since the start of the project and managers/planners have to operate the system based on the best available condition, in order to maximize the profit or minimize the loss. This is not captured in typical DCF valuation methods, and it can affect investment decisions on large-scale systems significantly.

For the lattice-based model, there are some assumptions implicit in it, which may not be realistic for use in an engineering context. Firstly, it is assumed that path independence holds in the model. That is, the value of a node at period t with an up-down movement in previous two periods ($t - 2\Delta t,\ t - \Delta t$) is equivalent to that of a down-up movement. While this is fine in the context of financial options valuation, this assumption may not hold true in an engineering context because one may act differently in those two scenarios. For instance, decision makers may expand system capacity if demand rises, in which case the system may have extra capacity when demand drops, leading to a different performance and value for the system than if a down-up movement occurs. In this case, decision makers may not expand system if demand drops first, and the system may thus have no extra capacity when demand rises. This highlights some of the path dependencies that are inherent to the analysis of complex engineering systems, which is not captured well in a lattice model.

Secondly, considering more than one uncertainty source in a binominal lattice model can be challenging. Quadranominal and even multinominal lattice approaches have been developed to deal with multiple uncertainty sources, however at the expense of a higher curse of dimensionality, which increases computational complexity, and detracts from the main advantages of a recombining lattice (Copeland & Antikarov, 2001). Also, it is challenging when multiple flexibility strategies are considered in the design. The computational time may be consideraly large when more than one strategies are taken into account, because decision makers need to determine the appropriate positions to exercise flexibility strategies by fitting historical data based on a pre-determined recursive formula. It may already be challenging for decision-makers to determine the optimal strategy at any given time, it becomes even more difficult to do so when considering a multinomial lattice. In the case of a simple lattice, the decision-maker must determine its state and stage in the lattice by fitting historical data, project the up-down movement based on this state, and then apply a recursive backward induction process to determine the best strategy at that particular time (e.g., exercise option, do not exercise). This may be challenging if the decision-maker does not have prior training in this advanced mathematical technique. This problem is exacerbated if the decision-maker needs to find the optimal strategy in a multinomial lattice space. Also, due to the nature of the model (with $u > 0$, $d > 0$), the performance of the system will never be negative. However, the expected value (i.e., $E_{t,n}$) sometimes could be negative when the profit of a system is the aim.

Finally, it is normally challenging to analyze several systems simultaneously (i.e., infrastructures in a complex system) using lattice-based approaches. Although it is possible to describe each system with a single lattice model, it is unclear how the interactions between systems can be explicitly considered in those models. These interactions however can heavily affect decision makings with regards to implementing the flexibility strategies. For example, each fire station can be considered as an individual system in an EMS system. The emergency incidents covered by one station could be covered by other stations, too. Assigning one incident to different stations may produce different situations and ultimately affects the implementation of flexibility strategies. Unfortunately, lattice-based approaches may prove difficult to use to address this issue, and therefore may not suitable for the analysis of flexibility in infrastructure systems.

Table 2.2 summarizes engineering applications making use of different ROA methods for valuing the flexibility. The second to the fourth columns list the typical methods applied for evaluating the flexibility. The term “CCS” is the abbreviation of carbon capture, transport, and storage. The idea of decision rules is considered as an implementation of flexibility strategies in the second column on the right. Check marks indicate the approach used in a given study, or whether decision rules are used as an implementation method with the optimal configuration. It is shown that little work making use of decision rules, and even less with the optimal configuration (i.e., initial configuration plus best decision rules).

Table 2.2 Engineering applications of studies making use of ROA methods for flexibility evaluation.

Example studies	System details	Real Option Analysis Methods			Method	
		Decision analysis	Dynamic programming	Simulation	Decision rules	Optimal settings
Babajide *et al.* (2009)	Oil	✓				
Wang *et al.* (2014)	Biomass power	✓				
Ajak and Topal (2015)	Mining	✓				
Kelly (1998)	Mining		✓			
Khansa and Liginlal (2009)	Security process innovation		✓			
Eckhause and Herold (2014)	CCS system		✓			
Chow and Regan (2011)	Transportation network			✓		
Pringles *et al.* (2015)	Electricity grid			✓		
Melese *et al.* (2015)	CCS system			✓		
Cardin *et al.* (2015c)	On-shore LNG			✓	✓	
Cardin and Hu (2016)	Waste-to-energy plant			✓	✓	✓

2.4.3 Related to EMS Systems

There is an opportunity to analyze EMS systems incorporating the idea of flexibility dedicated to the issue of improving expected system performance in terms of key performance indicators (KPIs) such as total costs and/or incident coverage rate. This opportunity arises due to the limits of current design and management strategies in EMS systems, as captured in the literature – and practice to some extent. Most existing analyses typically consider full capacity deployment (i.e., emergency vehicles and stations) at once at the beginning of the system life cycle (i.e., $t = 0$) and consider short term demand fluctuations as the main uncertainty driver. Those solutions (or designs) do not account well for fluctuations in emergency calls or incidents in the long term (e.g., 5-10 years) (Başar *et al.*, 2012; Goldberg, 2004). As a result, an EMS system may function well for the first few years, but then start underperforming later if adjustments are not made to accommodate changing incident patterns in light of changing demographics, economic situations, regulatory environments, and technology.

A few studies account for designing an EMS system in the long term. An important issue in such work is that long term demand of emergency calls or incidents is considered as a deterministic parameter rather than an uncertainty driver. The projection of the long term incidents is obtained from the analysis regarding historical or artificial data. This so-called forecast can be biased and incorrect for some reasons (Morgan & Henrion, 1992). Thus, it is most likely that the future will not turn out as planned for the whole project life cycle, and thus the system may be in a less than ideal configuration quite rapidly. Even if the forecasts are correct, which is highly unlikely, the idea of "Flaw of Averages" (Savage, 2002) indicates that decision based on the "most likely" or "average" scenario is usually inappropriate and may lead to bad investment, unless the engineering system response is purely linear. Such flaw is the consequence of Jensen's inequality for non-linear systems:

$$E[f(x)] \neq f(E[x]) \tag{2.17}$$

In short, this inequality represents that the expectation of a system output (i.e., left hand side) is not the same as the system output evaluated based on an expected input x (i.e., right hand side). Since an EMS system almost necessarily does not have a linear performance response, the "Flaw of Averages" may have significant impacts on the expected performance associated with typical designs where expected emergency incidents are used as input parameters to the modeling. Besides, the EMS system might be called to adapt its configuration flexibly to achieve better life cycle performance and reduce costs over time. This may pose a challenge when the system has been designed and planned based on rigid deterministic projections of future needs, so that flexible adaptations may be more costly.

To address the issues stated above, this thesis proposes a novel design approach and applies it to a design problem within EMS systems. The proposed flexible design allows the system to deploy capacity over time and space such that the system is able to adapt to the realization of uncertainty. The flexible system considers long term uncertainty in the design process, and it is thus able to perform well in the long term. Table 2.3 summarizes an overview of current designs and studies for EMS systems with the considerations of different features of interest. The second and third columns on the left indicate the types of objective functions for the design. As can be seen, most work focuses on short-term planning or daily operation (i.e., location and relocation) under uncertainty, and little focus on long term planning. Also, the uncertainty driver like emergency calls is typically not considered in the long term. Furthermore, none of them considers strategic-level flexibility as a means to deal with long term uncertainty.

Table 2.3 Example studies of current designs for EMS systems.

	Design Features				
Example studies	Obj. of Min.	Obj. of Max.	Long term planning	With uncertainty	Strategic-level flexibility
Toregas *et al.* (1971)	✓				
Schilling *et al.* (1979)		✓			
Daskin (1983)		✓		✓	
Ball and Lin (1993)	✓			✓	
Beraldi and Bruni (2009)	✓			✓	
McLay (2009)		✓		✓	
Gendreau *et al.* (2001)		✓		✓	
Gunawardane (1982)	✓	✓	✓		
Başar *et al.* (2011)		✓	✓		
Ghaderi and Jabalameli (2013)	✓		✓		
This thesis	✓		✓	✓	✓

2.4.4 Related to Economic Assessment of Siting Nuclear Power Plants

There is an opportunity to develop a novel approach to site nuclear power plants flexibly under long term uncertainty. The new approach should aim to site nuclear power plants in an optimal way based on the realization of uncertainty drivers, and over time and space, while past designs do not account well for long term uncertainty. More specifically, the concept of flexibility or real option can be considered as a means to deal with uncertainty pro-actively and increase the system's adaptability. Applicable strategic level flexibility includes investment deferral, phased deployment, capacity expansion, etc. Table 2.4 provides an overview of example applications using real option theory to deal pro-actively with uncertainty. The first column on the left represents the example studies, and the second one shows the uncertainty driver of social acceptance. As observed, real option strategy of deferral is in favor of siting nuclear power plants. Other flexibility strategies such as phased deployment and switching between technologies are also considered as alternatives to deal with uncertainty. To the author's knowledge, no study has yet focuses on the siting of nuclear power plant with the consideration of public/social acceptance as an important uncertainty driver potentially impacting future performance, in addition to standard uncertainty drivers like electricity demand.

Table 2.4 Example studies of capacity deployment in the nuclear engineering sector.

	Uncertainty	**Real Option Strategies**					
Example studies	Social acceptance	Deferral	Switching	Phased deployment	Capacity expansion	Life extension	No real options
Louberge *et al.* (2002)		✓					
Kiriyama and Suzuki (2004)		✓					
Rothwell (2006)		✓					
Abdelhamid *et al.* (2009)		✓					
Zhu (2012)			✓				
Cardin *et al.* (2012)			✓				
Siddiqui and Fleten (2010)				✓			
Jain *et al.* (2013)				✓			
Abudeif *et al.* (2015)							✓
Kojo and Richardson (2014)							✓
Erol *et al.* (2014)							✓
This thesis	✓			✓	✓	✓	

One reason creating this opportunity is that existing approaches consider long term demand fluctuation as a deterministic projection based on historical data, which could be incorrect. In addition to demand, there are many uncertainty factors involved in this siting problem, such as construction costs and time, price of electricity and fuel, etc. Deploying capacity once and all could underperform after a few years since the start of the project if the uncertainty parameters are far from expectations. This may cause a huge loss because the investment of nuclear power plants is capital intensive. On the one hand, deploying too much capacity which cannot be fully utilized in early stages will significantly increase construction costs, and expose system operator/owner to potential losses. On the other hand, deploying less capacity which cannot satisfy required demand of electricity will introduce contractual costs or penalty costs for losing demand. Both cases would affect the long term system performance in terms of the total costs.

Another reason for creating this opportunity is that the impacts on the investment caused by social or public acceptance of nuclear technology are typically not considered in the valuation of the investment. Studies show that nuclear accidents/disasters that occurred outside one region or country could have significant impacts on the investment of nuclear power plants inside a given region or country, even though how significant the impacts are is not confirmed yet, depending on which region or country it is (Joskow & Parsons, 2012). The social acceptance is somehow influenced by those nuclear events as they could harm public confidence on the safety issue with respect to nuclear technology. As a result, nuclear power may lose the support from the public and the funds for investing new plants from the government. Past research, however, has not paid enough attention to this issue. This thesis explicitly considers social acceptance in terms of cumulative INES in Chapter 6 and analyzes the value of flexibility under such assumptions.

One last reason for exploring this opportunity is based on the observation of past studies. Most of the work applying real option focuses on the issue of deferring investment until favorable market condition arises. These studies can be seen as examples of real option "on" projects. Nuclear power plants are partially deployed at the beginning of the life cycle of the project. Once the market condition is favorable, the remainder capacity will be deployed to satisfy the customers' need. There are two potential issues included in the past research. One issue is that the consideration of flexibility is correct but not comprehensive. Besides investment deferral, real option "in" projects such as capacity expansion and technology switch could also be valuable for the investment, as they did in other engineering contexts (de Neufville *et al.*, 2006; Marreco & Carpio, 2006). The other issue is that none of those studies, to the author's knowledge, consider social acceptance as an uncertainty driver in their analysis. As an important factor, social acceptance may call off a nuclear project when the public do not favor or even be scared with this technology. It is an uncertainty source that is exogenous to decision-makers, and may not be ignored. The valuation of the investment using real option/flexibility as a means to deal with uncertainty is also influenced by this lack of consideration.

2.5 Anticipated Contributions

This section explicitly discusses the main contributions of this thesis. The main contributions are twofold. First, this thesis proposes a novel design approach for infrastructure systems to further improve long term system performance under uncertainty. The proposed design approach incorporates the concept of flexibility instead of robustness, and the flexibility strategies are exercised based on managerial decision rules. Secondly, the proposed approach aims to address issues that exist in typical ROA approaches, such as path independency and the curse of dimensionality. The proposed design could then be more practical and easier for use by people lacking advanced mathematical knowledge. In addition, the proposed approach is applied to two engineering sectors – EMS and nuclear power plant systems – to analyze the value of flexibility systematically. The results shown later in Chapter 5 and Chapter 6 demonstrate that the proposed design approach can indeed help improve long term system performance.

2.5.1 An Approach for the Design of Infrastructure Systems

This thesis aims to provide system designers a novel design approach to developing infrastructure systems that can be adaptable to the rapid change of economic condition, demographics, technology, etc. The new approach addresses the issue that typical designs often function well for the first few years, but then may underperform later if the system cannot adjust itself to accommodate changing uncertainty patterns. The approach introduces the concept of flexibility, also known as real option, to make the system capable of this adjustment easily in the face of uncertainty. Strategic, tactical, and operational levels of flexibility strategies are possible to be embedded in an infrastructure system, depending on the design problems. The anticipated contributions discussed in this section could be considered as domain contributions.

The proposed approach addresses some of the concerns discussed in Section 2.4.1 as to how to improve expected performance of an infrastructure system in the long term under uncertainty. The proposed approach can be used in general to develop an infrastructure system and incorporate flexibility into the whole design according to the specific requirements and situations. Compared to the concept of robustness, not only flexibility limits downside risks by reducing exposure to possible losses, it also enables a system to pro-actively deal with uncertainty to gain on upside opportunities. For example, a flexible EMS system can, on the one hand, capture increasing emergency calls by deploying more capacity and emergency vehicles over time and space – thus improving response time. This system also can, on the other hand, reduce unnecessary costs by deploying less capacity and emergency vehicles in early periods, and limiting capacity deployment and/or vehicles if the number of emergency incidents increases slowly. As demonstration, two applications regarding an EMS system and a nuclear power plant system are described in Chapter 5 and Chapter 6, respectively. Typical designs for EMS systems focus on deploying capacity based on a fixed and short term plan, while designs for nuclear power plant systems focus on deploying capacity based on deterministic long term projection of electricity demand. As a contrary, the proposed design deploys the capacity over time and space, and considers long term uncertainty at the same time. Note that the flexible design may not always be better than non-flexible design due to the cost premium introduced by enabling flexibility. The flexibility is worth using only if the value of flexibility is greater than or equal to this cost premium. Also, the value of flexibility is not constant as well, depending on the assumption of significant parameters used in the design procedures.

2.5.2 A Modeling Framework for Flexible Infrastructure Systems

This section discusses the methodological contributions of this thesis. That is, a novel approach is described as a modeling framework based on the SAA scheme in Chapter 4 to help managers, planners, and/or designers developing flexible infrastructure systems incorporating ideas of flexibility. The flexibility used in the design is analyzed via managerial decision rules, and captured by non-anticipative constraints inside of the multi-stage stochastic integer programming model. The objective of the mathematical model for an infrastructure system is to find the stochastically optimal initial configuration as well as the best value of significant parameters for decision rules. The value of flexibility will be analyzed through a systematic and rigorous procedure extended upon the typical four-step procedure, which is applied in various engineering sectors (Cardin *et al.*, 2015c; Cardin *et al.*, 2012; Deng *et al.*, 2013).

The proposed framework addresses some concerns described in Section 2.4.2, with respect to the valuation of flexibility, as well as it implementation in operations. Compared to standard approaches based on dynamic programming and simulation, the proposed approach is developed based on mathematical modeling (more specifically stochastic programming), and analyzed via managerial decision rules. Compared to lattice-based approach, a multi-period stochastic programming has its advantages in dealing with many issues discussed above. Firstly, no path independence is strictly required in developing a mathematical model, and the objective value of such model of course could be negative if necessary. Uncertainty sources can be described either in analytical or approximation forms in the constraints. Secondly, the dimension of the mathematical model increases much slower than that of lattice-based approaches. So the curse of dimensionality could be eliminated in some sense. In addition, the exercise policy for flexibility strategies is based on managerial decision rules, emulating the actual decision-making process.

The mathematical model is theoretically able to find the stochastically optimal initial configuration as well as best decision rules among tens of thousands of combinations in a much shorter computational time, as compared to lattice-based approaches. The proposed approach is then better for estimating the value of flexibility strategies than lattice-based approaches from the pratical point of view. Moreover, the interactions existing in an infrastructure system can be explicitly captured in the constraints by properly defining sets of fire stations based on the vehicle fleet size. For example, N_i denotes the set of districts that can cover district i. The constraint $\sum_{j \in N_i} y_{ijt} \leq 1$ indicates that one district can only be assigned to one fire station. Lattice-based approaches however cannot account easily for such interactions, which is inherent to infrastructure systems. The proposed framework can provide managers, planners, and/or designers information on how to enable and use flexibility appropriately without relevant knowledge of more advanced mathematical techniques such as backward induction (used in DP-based approaches to determine when it is optimal to exercise a particular source of flexibility). In contrast to a simulation-based approach, it helps planners identify a stochastically optimal initial configuration and best decision rules so that the system can have better expected performance over its life cycle, compared to typical designs. Note that due to the complexity of the design problems regarding large-scale engineering systems, the global optimal solution may not always be found in finite or acceptable time. The proposed framework, however, is still helpful in finding good practical solutions within a short amount of time by applying heuristic algorithms (e.g., Tabu Search, Genetic Algorithm).

The valuation of flexibility for the proposed framework is based on an extension of an existing four-step methodology. The solutions obtained in the step of Uncertainty Analysis (i.e., step 2) and Flexibility Analysis (i.e., step 3) is by solving the stochastic model with consideration of multiple representative scenarios. The Post-optimality Analysis of step 4 consists of several parts, such as out-of-sample analysis, sensitivity analysis, and Pareto test. The solutions obtained in step 2 and 3 will be used in this step as the input, while significant parameters will be modified accordingly in order to find their impacts on the long term expected performance over system's life cycle. For instance, in EMS systems the requirement on incident coverage rate and the expected growth rate of emergency incidents are two significant parameters that could affect the system performance considerably. One can eventually find the value of flexibility and most sensible parameters for an infrastructure system though this valuation procedure.

Chapter 3 RESEARCH QUESTIONS AND APPROACHES

"If we knew what it was we were doing, it would not be called research, would it?" – Albert Einstein (1879 – 1955)

This chapter revisits and discusses the research questions investigated in this thesis, and the research approaches used to address them. To answer the research questions of interest, three research areas are identified and discussed in details. Both general and specific research questions are formulated in this chapter, indicating the anticipated contributions this thesis hopes to make. It is not expected that all questions can be fully answered here. The goal is to structure a novel framework for the design of infrastructure systems, and for inspiring further research on a longer timescale in this engineering sector.

The first research area focuses on the design of infrastructure systems over their life cycle, incorporating the idea of flexibility to pro-actively deal with long term uncertainty, discussed in Section 3.1. The second area is about the real options analysis approaches. The third and the fourth research areas are concerned with the development of a flexible system in the context of EMS systems and nuclear power systems, which are described in Sections 3.3 and 3.4, respectively. These two engineering systems are selected examples of infrastructure systems, used for demonstrating the feasibility and value of flexibility in design and management of such systems.

3.1 Area 1: Design of Flexible Infrastructure Systems

The first research area is motivated by the research opportunities identified in Chapter 2, focusing on the issue of how to design and manage infrastructure systems under uncertainty, especially by exploiting the emerging idea of flexibility. It is anticipated that the novel design can improve the system's expected performance in terms of KPIs over its life cycle. This broader general research interest is formulated as the following question:

"What is the best design for an infrastructure system to make it adaptable to the changing environment such as market condition and demographics, and to improve the anticipated performance over the life cycle of the system under uncertainty?"

It is postulated that incorporating the concept of flexibility or real options into the design of infrastructure systems and analyze flexibility via managerial decision rules, can help achieve these goals. To determine whether or not design the system for flexibility, one may want to find out the value for enabling flexibility. The more specific research question discussed in this thesis is:

"How can one design a flexible system with a stochastically optimal initial configuration as well as best decision rules, and evaluate the corresponding value of flexibility compared to benchmark design (i.e., typical design without flexibility)?"

The notion of "stochastically optimal" is necessary to measure the quality of a flexible design in the face of a large number of uncertainty scenarios. The initial configuration, on the one hand, is the decision-making process regarding how to deploy capacity at the beginning of the life cycle ($t = 0$). It is assumed that decision-making processes are only affected by observations about previous decisions, because the information in the future cannot be known before it really occurs. This initial configuration is considered as a fixed plan and independent from the realization of scenarios. The decision rules, on the other hand, are guidance for planners to change the system based on the realization of scenarios so that the system can perform well on average. The initial configuration and decision rules will be determined by the proposed modeling framework and solved to optimality. The framework is described explicitly in Chapter 4. In addition, the valuation procedure for flexibility is extended upon the typical four-step methodology, and is demonstrated through two applications in Chapter 5 and Chapter 6.

3.2 Area 2: Real Options Analysis Approaches

As discussed in Section 2.4.2, typical real options analysis approaches do not sufficiently account for long term uncertainty. The commonly used approaches have implicit assumptions that make them unrealistic, like path independence. Problems regarding the design and management of infrastructure systems usually have a high level of complexity, and long term uncertainty (even multiple uncertainty drivers) is inevitably involved. It is anticipated that the proposed novel approach could handle the design problem by relaxing those assumptions. To do so, one may ask the following question:

"Can one develop a different modeling framework that is suitable for the analysis of flexibility and real options in infrastructure systems?"

Throughout the literature review, it is found that operations research techniques could be a useful supplement to ROA approaches. Mathematical programming, Markov decision processes, and queuing theory are possible approaches to systematically analyze the value of flexibility. Also, these methodologies do explicitly account for uncertainty, but in different manners, and none of them rely on the assumption of path independence. In addition to the above question, one may want to know:

"Why is stochastic programming based on the sample average approximation selected as the approach to analyze the design problem in this thesis?"

Since the flexibility strategies are exercised via managerial decision rules, it is better for decision makers to have stochastically optimal rules with fixed parameters (e.g., when to exercise the rule). Modeling the design problem using stochastic programming could directly embed the decision rules into the model and then solve it to optimality to obtain the rules anticipated. Sample average approximation provides a relatively easy way to solve the stochastic programming by considering a few representative sample scenarios, which makes the proposed approach applicable.

3.3 Area 3: Design and Management of Flexible EMS Systems

EMS systems include all necessary features such as the resources (e.g., stations and emergency vehicles), links (e.g., roads), and customers (e.g., patients) to serve as good example for the analysis of infrastructure systems. Typical design planning and solutions for an EMS system is to deploy capacity all and once at the initial period based on a deterministic or stochastic analysis. Such designs probably perform well in the first few moment of operations, but may soon underperform because they cannot be adjusted to accommodate the changing uncertainty patterns. As demonstrated in a traditional engineering contexts in industry sectors like aerospace and real estate, flexibility can improve the expected system performance over its life cycle by 10-30% (de Neufville & Scholtes, 2011). One may ask the following question associated with the broader interest of this area:

"Can one develop an EMS system incorporating flexibility so that it can stochastically dominate existing rigid designs (i.e., the benchmark) in terms of KPIs such as total costs and/or incident coverage rate?"

To answer the above research question, a case study is described in Chapter 5 where a proposed flexible system is compared to two rigid systems. A flexible design "dominates stochastically" a rigid design if the expected performance of this flexible system obtained via out-of-sample analysis is better than that of the rigid design. Note that improving multiple KPIs may require considerations of a trade-off between the different objectives. For instance, higher incident coverage rate may imply higher total costs, while it actually worsens the performance associated with these costs. The more specific research question therefore is:

"From an economic view, is a flexible design always better than a non-flexible design?"

The reason that this question arises is because there is a cost premium for enabling flexibility. The flexible design may not be the preferred one from an economic, if this premium is greater than the expected benefit from design for flexibility. To answer this question, one may need to assess the value of flexibility for a specific design problem with given assumptions on the main uncertainty drivers characterized by a specific set of parameters. The value of flexibility can be obtained through a systematical and rigorous procedure, as described in Chapter 5.

3.4 Area 4: Design and Management of Flexible Nuclear Systems

Energy systems also represent another typical example of infrastructure systems. The resource is the electricity generated by the power plants in the system, and the links are the power grids connected between cities, towns, and individual houses. This thesis selected nuclear power as the source for generating electricity instead of traditional thermal power, mostly due to its ability to limit CO_2 emissions, and ensure better sustainability. In particular, the design problem considered focuses on siting nuclear power plants in a region or country, as described in Chapter 6. Social acceptance of nuclear technology is considered as one of the major uncertainty drivers in this thesis. Similar to the EMS system, a broader research interest is summarized with the following research question:

"Can one site nuclear power plants flexibly so that this energy system can have good anticipated performance in terms of total costs under uncertainty?"

The notion "good" anticipated performance is necessary to measure the plan (or design) of siting plants incorporating the concept of flexibility. A design is good in siting nuclear power plants if it minimizes the expected total costs or expected LCOE over the lifetime of the energy system. Flexibility is introduced in the design in light of its past performance in other engineering sectors. The more specific research questions addressed in this thesis are:

"What flexibility strategy can one consider in the design to improve the anticipated performance of the energy system? If multiple real options strategies are taken into account, which one benefits the system mostly?"

The above research questions are concerned with the flexibility specifically embedded in the design, and are explicitly discussed in Chapter 6. A particular flexibility can be defined as a strategic-, tactical-, or operational-level one depending on which level of decision it belongs to. Due to the huge expenses for siting nuclear power plants, it is necessary to find out the most valuable flexibility as the most preferred option if budget is limited. Besides, one distinctive feature of the novel design proposed is the consideration of social acceptance. The more specific research question therefore is:

"What is the influence of social acceptance on the expected performance for a nuclear system, as well as decision-making processes?"

The hypothesis is that the social acceptance at least has significant impacts on the expected performance, for both rigid and flexible systems. To answer this question, a comprehensive case study is described in Chapter 6, where social acceptance is considered as an uncertainty that can be turned on/off. It is also anticipated that the flexible system is affected less than the rigid system when social acceptance turns on. This is due to the nature of flexibility as demonstrated in past studies.

Chapter 4 MODELING FRAMEWORK FOR INFRASTRUCTURE SYSTEMS

“Good tools are essential to do the job well.” – Confucius (c. 551 – 479 BC, the Analects)

This chapter explains explicitly the modeling framework for designing infrastructure systems exploiting the idea of flexibility under long term uncertainty. The proposed framework is represented as a multi-stage mixed integer stochastic programming based on a sample average approximation scheme, where decision-making processes and decision rules are captured by integer/binary decision variables and non-anticipative constraints, respectively. This framework builds upon and extends the typical resource allocation model with considerations of flexibility, and suggests that it can be applied more generally to support the design of infrastructure systems. The purpose of this framework is to provide system planners a novel design approach that can be used to find the stochastically optimal design configuration as well as best decision rules for enabling and implementing flexibility easily and appropriately in infrastructure systems.

The proposed framework consists of defining design variables, identifying modeling restrictions, and performing the numerical analysis. Section 4.1 describes the major decisions that are involved in an infrastructure system, shown as design variables in the mathematical model. Section 4.2 introduces internal and/or external factors that may influence the regular operations of an infrastructure system. These factors are then captured by constraints in the model. The numerical analysis procedure is described in Section 4.3. this procedure is based upon the standard four-step methodology used in various applications (de Neufville & Scholtes, 2011). It aims at finding the stochastically optimal initial configuration as well as best decision rules for the flexible design, the value of flexibility under specific assumptions of uncertainty drivers, and determining the parameters which affect the expected performance and the value of flexibility mostly.

4.1 Defining Design Variables

This section explicitly describes the issue regarding defining design or decision variables which are used for representing different levels of decisions in general involved in the design of infrastructure systems. The decision variable can be either discrete or continuous, depending on the variable itself. If the variable denotes whether to install a new station, it should be a binary variable. The variable could also be continuous if it denotes the capacity to be deployed. It should be noted that the type of decision variable does not depend on the point in time that the decision is made. In particular, general decisions are described as follows.

1. Decisions regarding where and/or when to allocate infrastructures used

for generating and/or storing resources. This is one of the most important decisions made for the design of infrastructure systems, because the availability of resources depends on the availability of infrastructures. This strategic-level decision is about determining the appropriate locations to install infrastructures and the best time to start this installation. Infrastructures in such systems may have different functions. For instance, a station in an EMS system is used for parking emergency vehicles when they are idle, while a nuclear power plant is used for generating electricity. One may formulate this decision either as an integer or a binary variable in each candidate site.

2. Decisions regarding how much capacity to be deployed for each available infrastructure and when to do that. This is also an important decision, as it directly influences the output of an infrastructure system. The term "capacity" has different meanings in different systems. It represents the number of emergency vehicles in an EMS system, or represents the power output of a plant in an energy system. One may model this as an integer variable.
3. Decisions regarding infrastructure closure over the life cycle. The infrastructure of an infrastructure system always has a lifetime in practice, and it is also feasible to close an infrastructure before its service life if needed. This issue about infrastructure closure would be considered when the life cycle of the system is longer than the lifetime

of the infrastructure. One may use a binary variable to describe this decision.

4. Decisions regarding how much capacity of resource will be deployed to a specific customer or a group of customers. This is a tactical level decision that directly links to the customers who are involved in the system. Note that the resource (e.g., electricity) from an infrastructure could be shared by different customers at any one time period. It is also true that in some systems resources can only serve one customer at a time (e.g., emergency vehicles). Formulating this decision either as an integer or a binary variable is acceptable.
5. Decisions regarding the concept of flexibility incorporated in the system. This decision is applicable only in the flexible system due to the special constructional feature, for instance, the strong pillar of a parking garage for enabling capacity expansion (de Neufville *et al.*, 2006). It is introduced by the flexibility embedded in the design, and can be represented by an integer or a binary variable.

4.2 Identifying Design Restrictions

The constraints used in the model represent the real-world requirements and limits involved in the design and management of infrastructure systems. Typical requirement is about the service satisfaction with the system, e.g., the incident coverage rate in an EMS system. Such requirement usually introduces penalty costs if the system cannot respond in a timely manner. Besides, there are various limits in practice if one would like to design an infrastructure system. Those limits are inevitable due to the boundary of current technology and/or physical structure. For instance, one station with three units' capacity cannot operate and maintain four vehicles at a time. The causal relationship between decision-making processes also needs to be represented as constraints so that the model may capture reality. A nuclear power plant, for example, cannot be closed if it is not installed yet.

The entire constraints can be separated into two groups. The first group of constraints consists of numbers of basic requirements and common limits which are not related to flexibility in the design of infrastructure systems. This group of constraints is fundamental for designing infrastructure systems in general, whilst the specific representations of the requirement between systems are not necessarily the same.

The other group of constraints represents the decision-making processes regarding the flexibility considered in the design. The constraints are formulated as logical expressions since it is assumed that flexibility is analyzed via managerial decision rules. The "IF-THEN-ELSE" statements can be explicitly described in an integer-based programming model. More detailed examples can be found in the case studies introduced in Chapter 5 and Chapter 6, where mathematical formulations about the decision rules are explicitly discussed. Note that the complexity of developing constraints associated with flexibility heavily depends on the definition of the corresponding decision rules.

4.3 Numerical analysis

The approach used for numerical analysis in this thesis follows a standard four-step methodology. In addition, Cardin *et al.* (2015a) proposed a design catalog as a systematic approach to improve the design and evaluation of engineering systems by exploiting the concept of flexibility. A similar four-step methodology was applied in their study as well. This four-step methodology is therefore used in this thesis for the evaluation of flexibility strategies in the design of infrastructure systems. It is anticipated that after the numerical analysis there is a clear understanding on how to design and manage a flexible infrastructure system if the value of flexibility is favorable. This approach can be summarized as follows.

1. Deterministic analysis. This step focuses on finding the best rigid design based on a deterministic projection of the main uncertainty drivers. The expected or most likely value of those uncertain parameters is usually chosen to create the projection. The design is normally developed based on professional or past experience, while it is found by solving the simplified model to optimality in this thesis.
2. Uncertainty analysis. In contrast to considering deterministic values for the uncertainty modeling parameters, this step focuses on finding

the best rigid design based on various scenarios. The optimal design is determined by solving the stochastic model, which may differ from typical methods used in practice. The optimal solution of the rigid design under uncertainty will be analyzed in the out-of-sample test to gain significant information on the distribution of the design solution for a large number of uncertainty sample scenarios, as well as the mean and standard deviation. The purpose of this out-of-sample analysis is to estimate the performance of the solution based on forecasts calibrated using historical data.

3. Flexibility analysis. This step aims at finding the stochastically optimal initial configuration as well as the best decision rules. The procedure is similar to step 2, while the difference is in the model analyzed in this step. The mathematical model incorporating flexibility will be solved to optimality by considering multiple representative scenarios at a time based on the SAA scheme. Similar to step 2, the optimal solution will also be subject to an out-of-sample analysis due to the same reason as stated in last paragraph. The value of flexibility will then be calculated by comparing the results (i.e., expected performance) for the flexible and deterministic designs in the out-of-sample analysis. This comparison is fair because it uses a considerably large sample size (e.g., 1,000 sample scenarios) to avoid the situation where one design may be favorable for particular

scenarios.

4. Sensitivity analysis. A comprehensive sensitivity analysis will be conducted to find out which parameter influence most the expected performance of a given design solution. The purpose of the sensitivity analysis is to find the most significant parameter influencing the variability in the inputs' assumptions. That is, how the objective value will change when we change the value of the input parameters. In this thesis, the sensitivity analysis consists of two consecutive sub-analyses: 1) evaluate the influence of several input parameters of interest based on a one-factor-at-a-time approach, and then 2) evaluate the most influential input parameter by slightly adjusting its value. The most influential parameters will be selected for a Pareto test to see how the expected performance evolves alongside with possible changes in those underlying parameters.

Chapter 5 CASE STUDY ONE – DESIGN AND MANAGEMENT OF FLEXIBLE EMS SYSTEMS

"The truth of the sea, and let not found things to lie down in front of my eyes, let I to explore." – Sir Isaac Newton (1642 – 1726)

This chapter describes a case study about the design and management of an EMS system under uncertainty, considering the concept of flexibility. It describes explicitly the proposed multi-stage set covering location problem (MSCLP) considering multiple periods and the numerical analysis, as explained generically in Chapter 4. Section 5.1 introduces the design problem in details, while Section 5.2 explicitly discusses the mathematical model including notations, variables, and constraints. Section 5.3 describes a specific analysis of the numerical analysis based on a hypothetical city.

5.1 Step 1: Design Problem Description

In recent decades, much research has been dedicated to the issue of designing a cost-effective EMS system to quickly and efficiently respond to patients' emergency calls. An EMS system is a type of emergency services that is dedicated to providing out-of-hospital acute medical care, and transporting the patient to the nearest available hospital for definitive care. Ong *et al.* (2009) show that delivering fast defibrillation in out-of-hospital care to emergency patients could significantly increase their survival rate. Population, demographic structure, and the environment influence the spatial distribution of emergency incidents. This distribution is therefore uncertain and changes over time in terms of incident rates. The current designs of EMS infrastructure and vehicle systems, however, often focus on an optimal configuration that may not account well for fluctuations in emergency incidents in the long term (i.e., 5-10 years). Consequently, an EMS system may function well for the first few years, but then start underperforming later if adjustments are not made to accommodate changing incident patterns in light of changing demographics, economic situations, regulatory environments, and technology. The EMS system might be called to adapt flexibly to achieve better life cycle performance and reduce costs over time. This may pose a challenge when the system has been designed and optimally planned based on rigid deterministic projections of future needs, so that flexible adaptations may be more costly.

This case study is about the design and management of an EMS system in the context of a hypothetical city. This city could be thought of a small or medium-sized one like many cities in the central and western regions of China or other emerging and urbanizing countries. The system planner focuses on developing a flexible EMS system that can satisfy the requirements on patients over life cycle, and optimizing the expected performance in terms of KPIs. In this case study, it is assumed that land can be "reserved" for the possibility of deploying a new station in the future. Even though this may not be always true in reality, it is still applicable because the land reserved right now could still be used to build other facilities in the future, as long as the facility has the ability to house emergency vehicles (e.g., reserved position for parking). Also, the set of candidate sites may or may not be fixed during the system life cycle, depending on the real situation. This set is assumed to be fixed in this case study for simplicity. The proposed model, however, is clearly able to account for cases where the set of candidate sites changes over time.

5.2 Step 2: Analytical Model

The proposed flexible design alternative considers strategic- and tactical-level flexibility and is analyzed under a stochastic programming framework. Compared to typical models, strategic- and tactical-level decision-making should occur on different time scales. As shown in Figure 5.1, it is assumed that tactical-level decisions occur four times as frequently (e.g., once every quarter) as strategic-level decisions (e.g., once every year). This asymmetry reflects reality based on discussions with a local EMS provider, and is not limited to EMS systems. To make the alternative practical, these two types of decisions are distinguished.

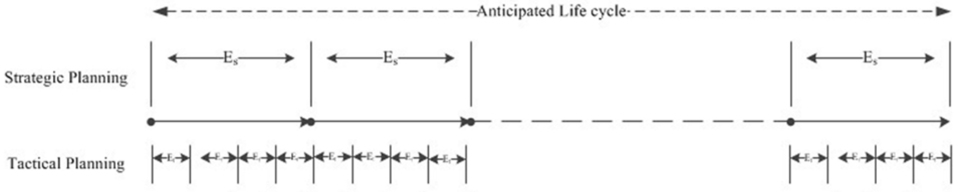


Figure 5.1 An illustration of strategic and tactical decision horizons.

5.2.1 Symbols and Terminology

The flexible alternative deals pro-actively with uncertainty by leveraging the concept of flexibility. Here, the term "flexible" has the meaning that the infrastructure and vehicles in an EMS system can change and adapt according to variations in emergency incident rates over time. To realize this, strategic-level flexibility is introduced. First, the model accounts for a flexible phased deployment of station capacity over time and space. Instead of deploying all stations at once, the flexible design allows the deployment of stations over time, depending on the realization of the uncertainty scenario.

Second, flexible capacity expansion at any given strategic period is studied. Once a station is deployed, the capacity can be increased when needed, which benefits from special attention to the design of the infrastructure (e.g., designing for smaller capacity first, and carefully planning for expansion in the future by buying a collocated piece of land, for instance). The phased structure of EMS stations allows the system to expand an existing station to a particular phase, or deploy a new station with pre-determined phase, as needed. Each phase in this model indicates a specific capacity to a station, which is related to its ability to maintain and operate a given number of emergency vehicles. The following list summarizes the modeling notation (sets and parameters):

$S =$ the set of strategic periods in a complete lifecycle $(s \in S)$;

$S_1 =$ the set of strategic periods in which $s = 1$;

$S_2 =$ the set of strategic periods in which $s \geq 2$;

$S_3 =$ the set of strategic periods in which $s = |S|$;

$T =$ the set of tactical periods $(t \in T)$, partitioned into $|S|$ subsets, noted as T_s;

$T^1 =$ the set of tactical periods in which $t = 1$;

$E_s =$ the length of a strategic period. It may range from one to several years;

$E_t =$ the length of a tactical period. It may ranges from several weeks to several months;

$E =$ the set of length of review periods $(e \in E)$. The maximal length is denoted as a_m, which equals to $\lfloor E_s/E_t \rfloor$;

$I =$ the set of districts to be covered in the system $(i \in I)$;

$J =$ the set of candidate sites to allocate stations $(j \in J)$;

$N =$ the set of emergency incident scenarios $(n \in N)$;

$L =$ the set of station phases $(l \in L)$. The corresponding capacity, installation cost and operation cost of a phase l station are denoted as U_l, c_l and m_l, respectively;

$L_1 =$ the set of station phases with the minimum capacity (i.e., $U_l = 1$);

$L_2 =$ the set of station phases with the maximum capacity (i.e., $U_l = |L|$);

$L_3 =$ the set of station phases when $U_l < |L|$;

$N_i =$ the set of candidate sites that can cover district i. That is, the ambulances allocated in station $j \in N_i$ can respond to the incidents in i within a predetermined time (e.g., 11 minutes in Singapore), usually to satisfy the EMS provider's required key performance indicator (KPI) (e.g., serve a given percentage of call within a given time);

$M_j =$ the set of districts that can be covered by candidate site j. The number of elements in this set is denoted as $|M_j|$;

$c_v, m_v =$ the unit cost of the medical vehicle and its corresponding maintenance cost;

$c_u =$ the unit cost per phase for capacity expansion;

$q_{ij} =$ the corresponding cost of assigning district i to candidate site j;

$CoV =$ the required incident coverage rate;

$r_t, r_s =$ the discount rate for tactical period t and strategic period s, respectively;

$p_n =$ the corresponding probability of scenario n;

$h =$ the number of hours in a tactical period;

$\varepsilon, M =$ a small tolerance and an arbitrary large integer, used to ensure a given constraint is always or never satisfied.

In addition, the following random variables are considered in the model:

$d_{itn} =$ the average number of emergency incidents per hour (i.e., the incident arrival rate) in district i within tactical period t under scenario n. This is a random variable that will be captured by the GBM process.

5.2.2 Mathematical Formulations

There are several assumptions in the flexible design alternative. Firstly, two levels of decision-making processes are implemented in different time periods. Secondly, opening a station is instantaneous, while closing an open station is not considered. In reality, this new installation is not instantaneous due to construction delays. However, some buildings or facilities can be used as emergency stations and the time for this change is negligible. For instance, emergency vehicles could be temporarily or permanently allocated in fire stations, such as is done in Singapore. Closure is not allowable in this model because simply closing a station is definitely not a good choice because installation costs too much. The decision making regarding open – closure – reopen, however, makes the design problem too complicated. The mathematical model can be too complex to be solved if all these possibilities are considered together. Also, every decision is made at the beginning of either a strategic or tactical period. The number of missing incidents or the total number of incidents in a district is considered as a criterion for implementing the corresponding decision rules and exercising the real option strategies. The following notations denote the decision variables considered in the MSCLP model:

$o_{jl}^{1} = 1$ if a phase l station is opened at site j when strategic period $s \in S_1$, and zero otherwise;

$o_{jlsn}^{2} = 1$ if a phase l station is opened at site j when strategic period

$s \in S_2$ under scenario n, and zero otherwise;

$o_j^o =$ the capacity to be deployed at site j if $o_{jlsn}^2 = 1$;

$x_{jlsn} = 1$ if a phase l station at site j at strategic period s under scenario n;

$u_{jlsn} = 1$ if the station at site j is expanded l unit when strategic period $s \in S_2$ under scenario n, and zero otherwise;

$y_{ijtn} = 1$ if district i is assigned to site j at tactical period t under scenario n, and zero otherwise;

$o_j^u =$ the capacity to be expanded at site j if $u_{jlsn} = 1$;

$v^1 =$ number of vehicles purchased when strategic period $s \in S_1$;

$v_{sn}^2 =$ number of vehicles purchased when strategic period $s \in S_2$ under scenario n;

$w_{jtn} =$ number of vehicles allocated at site j at tactical period t under scenario n;

$\delta_j^d =$ the amount of lost incidents to trigger the flexibility of capacity expansion at site j;

$\delta_j^o =$ the amount of incidents for phased deployment at site j when strategic period $s \in S_2$;

$\omega_{jtn}, \xi_{jtn} =$ non-negative variables, used to express absolute value of other variables.

The objective of the MSCLP model is to minimize the expected total costs over the life cycle. This model consists of two groups of constraints. The first group is typically found in EMS systems resource allocation models, which is used to represent the relationship between stations and vehicles. The other group consists of the so-called “IF-THEN-ELSE” decision rule statements, modeling strategic-level flexibility via non-anticipative constraints. The mathematical formulation of the MSCLP model is as follows:

$$\begin{aligned} \min \ & \sum_n p_n [\sum_s (r_s \sum_j \sum_l (c_l o^2_{jlsn} + c_u u_{jlsn}) + r_s c_v v^2_{sn} \\ & + \sum_{t \in T_s} m_l r_t w_{jtn}) + \sum_t (\sum_j [m_v w_{jtn} + \sum_i q_{ij} y_{ijtn}]) \\ & + \sum_j \sum_l c_l o^1_{jl} + c_v v^1] \end{aligned} \tag{5.1}$$

Subject to:

$$\sum_l o^1_{jl} \le 1 \; j \in J; \tag{5.2}$$

$$\sum_l \left(o^2_{jlsn} + u_{jlsn} \right) \le 1 \; j \in J, \; s \in S_2, \; n \in N; \tag{5.3}$$

$$x_{j,l+1,sn} \le x_{jlsn} \; j \in J, \; l \in L_3, \; s \in S, \; n \in N; \tag{5.4}$$

$$\sum_l x_{jlsn} = \sum_l U_l o^1_{jl} \; j \in J, \; s \in S_1, \; n \in N; \tag{5.5}$$

$$\sum_l x_{jlsn} = \sum_l x_{jl,s-1,n} + \sum_l U_l (o^2_{jlsn} + u_{jlsn}) \; j \in J, \; s \in S_2, n \in N; \tag{5.6}$$

$$\sum_l u_{jlsn} \le x_{jl^*,s-1,n} \; j \in J, \; l^* \in L_1, \; s \in S_2, \; n \in N; \tag{5.7}$$

$$v^1 = \sum_j \sum_l x_{jlsn} \; s \in S_1, \; n \in N; \tag{5.8}$$

$$v^2_{sn} = \sum_j \sum_l U_l (o^2_{jlsn} + u_{jlsn}) \; s \in S_2, \; n \in N; \tag{5.9}$$

$$w_{jtn} = \sum_l x_{jlsn} \; j \in J, \; s \in S, \; t \in T_s, \; n \in N; \tag{5.10}$$

$$y_{ijtn} \le w_{jtn} \; i \in I, \; j \in J, \; t \in T, \; n \in N; \tag{5.11}$$

$$\sum_{j \in N_i} y_{ijtn} \le 1 \; i \in I, \; t \in T, \; n \in N; \tag{5.12}$$

$$\sum_t \sum_j w_{jtn} - \xi_{jtn} \ge CoV \sum_t \sum_i d_{itn} \; n \in N; \tag{5.13}$$

$$\sum_{i \in M_j} d_{itn} y_{ijtn} - w_{jtn} = \omega_{jtn} - \xi_{jtn} \; j \in J, \; t \in T, \; n \in N; \tag{5.14}$$

$$h\sum_{t\in T_s}\omega_{jtn}-\delta_j^d \le M\sum_l u_{jlsn}-\varepsilon+Mx_{jl^{'},s-1,n}\ j\in J,\ l^{'}\in L_2,\ s\in S_2,\ n\in N; \tag{5.15}$$

$$h\sum_{t\in T_s}\omega_{jtn}-\delta_j^d \ge M(\sum_l u_{jlsn}-1)-Mx_{jl^{'},s-1,n}\ j\in J,\ l^{'}\in L_2,\ s\in S_2,\ n\in N; \tag{5.16}$$

$$\sum_{t\in T_s} d_{jtn} \le \delta_j^o-\varepsilon+M\sum_l o_{jlsn}^2+Mx_{jl^*,s-1,n}\ j\in J,\ s\in S_2,\ n\in N; \tag{5.17}$$

$$\sum_{t\in T_s} d_{jtn} \ge \delta_j^o-M(\sum_l o_{jlsn}^2-1)-Mx_{jl^*,s-1,n}\ j\in J,\ s\in S_2,\ n\in N; \tag{5.18}$$

$$\sum_l o_{jlsn}^2 \le U_l\sum_l o_{jlsn}^2 \le |L|\sum_l o_{jlsn}^2\ j\in J,\ s\in S_2,\ n\in N; \tag{5.19}$$

$$1-\sum_l o_{jlsn}^2 \le o_j^o-U_l\sum_l o_{jlsn}^2 \le |L|(1-\sum_l o_{jlsn}^2)\ j\in J,\ s\in S_2,\ n\in N; \tag{5.20}$$

$$\sum_l u_{jlsn} \le U_l\sum_l u_{jlsn} \le |L|\sum_l u_{jlsn}\ j\in J,\ s\in S_2,\ n\in N; \tag{5.21}$$

$$1-\sum_l u_{jlsn} \le o_j^u-U_l\sum_l u_{jlsn} \le |L|(1-\sum_l u_{jlsn})\ j\in J,\ s\in S_2,\ n\in N; \tag{5.22}$$

$$y_{ijtn^1}=y_{ijtn^2}\ i\in I,\ j=1,\ldots J,\ t\in T^1,\ n^1,n^2\in N; \tag{5.23}$$

The problem (5.1)-(5.23) describes the flexible design under the multi-stage stochastic programming framework. The objective function is the expected total costs over the life cycle, including costs for installation, capacity expansion, vehicle purchase as well as operations. Inequalities (5.2) -(5.3) indicate the boundary for making decision regarding installation and capacity expansion. Formulations (5.5)-(5.6) show how to calculate the capacity of a station at one site over the life cycle. Inequality (5.7) captures the condition in which the decision about capacity expansion cannot be made. Specially, such decision making is impossible if there is no station yet.

Formulations (5.8)-(5.9) indicate the relationship between strategic-level decisions, while (5.10)-(5.12) introduce the boundary for assigning districts to candidate sites. The tactical-level flexibility in this model is that each station should be fully utilized. In other words, the number of vehicles allocated to a station is equal to its capacity, which can be found in (5.10). Note that operational-level decisions (e.g., dispatch policy) are not explicitly considered in the model. The policy for vehicles to respond to incidents is similar to the one introduced by Beraldi and Bruni (2009), i.e., each district is assigned to no more than one station that is able to respond to it within a predetermined time (see (5.11)-(5.12)). Inequality (5.13) is the restriction regarding predetermined KPI that is required to be satisfied under different scenarios. The loss of demand in each tactical period is shown in (5.14). Inequalities (5.15)-(5.16) and (5.17)-(5.18) are model the "IF-THEN-ELSE" statements used in managerial decision rules regarding capacity expansion and station installation (i.e., phasing), respectively. More specifically, variable $u_{jlsn} = 1$ if the number of missing incidents per strategic period at site j (i.e., $h\sum_{t\in T_s}\omega_{jtn}$) is greater than or equal to the threshold δ_j^d, and the station has not reached its upper bound (i.e., $x_{jl',s-1,n} \neq 0$). Similarly, a new station would be installed at site j if the number of incident arrivals exceeds the threshold δ_j^o and this site has not been used before (i.e., $x_{jl^*,s-1,n} = 0$). The rule for capacity expansion is that station j will be expanded by o_j^u unit capacity if it loses δ_j^d unit number of incidents in the assigned district over a strategic period.

The rule for phased deployment is that a new station will be installed with o_j^o unit capacity at site j if the demand of a strategic period at that site is greater than or equal to δ_j^o. The remaining four constraints represent the mathematical technique used to model a general formulation $y = xd$, where variables y and x are continuous or integer while variable d is binary. These constraints are used to guarantee that the unit capacity deployed at site j for phasing at different time is identical, as well as for capacity expansion. The last constraint indicates the consistency in the decision-making processes, i.e., the demand should be assigned to the same station at the first tactical period. It should be noted that the decision rules considered in the model are not necessarily applicable everywhere. The application of the decision rules relies on available data and specific operation policy. For example, if an emergency vehicle could respond to an incident outside the coverage of the station where it is located, the decision rule regarding capacity expansion would then be meaningless for focusing on missing incidents for a particular station. If that is the case, one possible decision rule could be: IF *the total number of missing incidents is greater than or equal to a specific number* δ, THEN *the station that has the most covered incidents (i.e.,* $\sum_{t \in T_s} d_{jtn}$ *is the largest among all candidate sites) should be expanded until it reaches the upper bound*; OTHERWISE, *do nothing*.

5.2.3 Rigid Designs Formulations

This section introduces the mathematical formulations of two other design alternatives that were compared in numerical analysis – referred to as rigid designs. The rigid designs represent sensible strategies to deploy capacity and manage such system over time, and aim to capture best practices. The first design deploys all resources (i.e., stations and vehicles) at once at the beginning of the project, and reallocates vehicles accordingly over the system life cycle in an optimal manner. Thus, this alternative can be called the "least flexible" design (i.e., it embeds some level of operational flexibility in terms of vehicle reallocation). The least flexible design is a simple multi-period extension of Beraldi and Bruni (2009) where only emergency vehicles can be deployed over time. The second alternative is also a multi-period extension of Beraldi and Bruni (2009) in which both stations and vehicles can be deployed over time instead of all at once. It can be called the "less flexible" design because the resources are deployed gradually based on a calculated plan, but not according to specific realization of uncertainty. The less flexible design captures to some extent a robust design philosophy, finding the deployment path that can deal best with a wide range of scenarios, without requiring the system to be changed dynamically over time (Jugulum & Frey, 2007). Compared to the above designs, the flexible design gradually deploys and allocates resources according to the decision rules, which is more dynamic as it adapts to the uncertainty realizations. The MSCLP model described in

Section 5.2.2 can be generalized as Problem (5.24)-(5.27):

$$\min\ f(x) \tag{5.24}$$

Subject to:

$$A_1 x_1 \le b_1 \tag{5.25}$$

$$Cx_1 = d \tag{5.26}$$

$$\begin{aligned} A_2 x_2 &\le b_2, \text{ if } x_2 = 1 \\ A_2 x_2 &\ge b_3, \text{ if } x_2 = 0 \end{aligned} \tag{5.27}$$

Formulations (5.25)-(5.26) represent the typical constraints in such problem, and inequality (5.27) represents the decision rules-related constraints. $x = (x_1, x_2)$ denotes the solution vector of the optimization problem, where x_1 and x_2 represent regular and decision rule variables, respectively. Constraint (5.27) shows the coupled formulations to illustrate the managerial decision rules, as shown and discussed in constraints (5.15)-(5.18). For the two rigid designs, (5.27) is not included in the models and the original decision variables, such as o_{jlsn}^2 and x_{jlsn}, will be redefined. More specifically, the problem (5.28)-(5.36) describes the least flexible design under uncertainty, where most constraints are similar to ones in the problem (5.1) -(5.23) (e.g., (5.10) and (5.30), (5.11)-(5.14) and (5.32)-(5.35)). $x_{jl} = 1$ if there is a phase l station opened at site j over the life cycle.

$$\begin{aligned} \min\ & \sum_j \sum_l c_l x_{jl} + c_v v^1 + \sum_j \sum_t r_t (m_v w_{jt} + \sum_l m_l x_{jl}) \\ & + \sum_n p_n \sum_t r_t \sum_j \sum_i q_{ij} y_{ijtn} \end{aligned} \tag{5.28}$$

Subject to:

$$\sum_l x_{jl} \le 1\ j \in J; \tag{5.29}$$

$$w_{jt} = \sum_l U_l x_{jl}\ j \in J,\ t \in T; \tag{5.30}$$

$$\sum_{t} w_{jt} \leq v^{1} \; j \in J, \; t \in T; \tag{5.31}$$

$$y_{ijtn} \leq w_{jt} \; i \in I, \; j \in J, \; t \in T, \; n \in N; \tag{5.32}$$

$$\sum_{j \in N_i} y_{ijtn} \leq 1 \; i \in I, \; t \in T, \; n \in N; \tag{5.33}$$

$$\sum_{i \in M_j} d_{itn} y_{ijtn} - w_{jt} = \omega_{jtn} - \xi_{jtn} \; j \in J, \; t \in T, \; n \in N; \tag{5.34}$$

$$\sum_{t} \sum_{j} w_{jt} - \xi_{jtn} \geq CoV \sum_{t} \sum_{i} d_{itn} \; n \in N; \tag{5.35}$$

$$y_{ijtn^1} = y_{ijtn^2} \; i \in I, \; j \in J, \; t \in T^{1}, \; n^{1}, n^{2} \in N; \tag{5.36}$$

The less flexible design is captured in problem (5.37)-(5.46). In this design, the stations can be deployed in phases over time based on a fixed plan. Decision variable x denotes the decision-making processes regarding station installation over the life cycle, and it thus has the subscript s. Formulation (5.39) indicates that a phase l station opened at site j since strategic period s will remain open until the end of the project. As can be seen, this less flexible system is developed based on the same basic restrictions with small modifications in order to realize more specific features. The three alternatives (two rigid and one flexible) are analyzed numerically in the following section.

$$\begin{aligned} \min? \sum_{s} r_s \left[\sum_{j} \sum_{l} c_l \left(x_{jl,s} - x_{jl,s-1} \right) + c_v v_s \right] + \sum_{s} \sum_{t \in T_s} r_t m_l x_{jls} \\ + \sum_{t} r_t \sum_{j} m_v w_{jt} + \sum_{n} p_n \sum_{t} r_t \sum_{j} \sum_{i} q_{ij} y_{ijtn} \end{aligned} \tag{5.37}$$

Subject to:

$$\sum_{l} x_{jls} \leq 1 \; j \in J, s \in S; \tag{5.38}$$

$$x_{jl,s-1} \leq x_{jl,s} \; j \in J, l \in L, s \in S; \tag{5.39}$$

$$w_{jt} = \sum_{t \in T_s} \sum_{l} U_l x_{jls} \; j \in J, s \in S; \tag{5.40}$$

$$\sum_{j} w_{jt} \leq \sum_{s=1}^{s^t} v_s \; t \in T, s^t = \left\{ s \mid t \in T_s \right\}; \tag{5.41}$$

$$y_{ijtn} \le w_{jt}\ i \in I, j \in J, t \in T; \tag{5.42}$$

$$\sum_{j \in N_i} y_{ijtn} \le 1\ i \in I,\ t \in T,\ n \in N; \tag{5.43}$$

$$\sum_{i \in M_j} d_{itn} y_{ijtn} - w_{jt} = \omega_{jtn} - \xi_{jtn}\ j \in J,\ t \in T,\ n \in N; \tag{5.44}$$

$$\sum_t \sum_j w_{jt} - \xi_{jtn} \ge CoV \sum_t \sum_i d_{itn}\ n \in N; \tag{5.45}$$

$$y_{ijtn^1} = y_{ijtn^2}\ i \in I,\ j \in J,\ t \in T^1,\ n^1, n^2 \in N; \tag{5.46}$$

The least and less flexible designs are extensions of the existing model because it may not be fair to compare the proposed model based on long term planning with models based on short term planning. The consideration of the uncertainty drivers is limited in short term models due to the planning horizon, and thus may bias the results of a comparison.

5.3 Step 3: Numerical Analysis

The proposed case study is about the design and management of an EMS system in the context of a hypothetical city. This city could be thought of a small or medium-sized one like many cities in the central and western regions of China or other emerging and urbanizing countries. The current EMS system operating in this city was designed based on historical demographics, and operates only one station. Since the population grows annually, system capacity needs to be increased in order to have an acceptable incident coverage rate. The city is divided into 10 districts and, for the purpose of this analysis, each district is geographically abstracted as a demand node – see

Table 5.1. The coordinates of the nodes are listed as follows, which are adapted from the 36 node example problem in Batta *et al.* (1989). The problem considers how to design and manage the EMS system capacity over the next 10 years.

Table 5.1 The (x, y) coordinates for the 10 nodes of the city.

Node	**Location (x, y)**	**Node**	**Location (x, y)**
1	(32, 31)	6	(27, 29)
2	(29, 32)	7	(24, 33)
3	(27, 36)	8	(34, 30)
4	(29, 29)	9	(29, 21)
5	(32, 29)	10	(33, 28)

The current station operates at node no.3 with 1 unit capacity – meaning that it can operate and maintain only one emergency vehicle. This station is a typical one and thus its capacity cannot be expanded flexibly. The other nine nodes are able to open new stations if needed. To calculate the coverage rate of each node, a Euclidean metric is used for distance measurements. The distance standard for coverage is 10 units. The assignment cost q_{ij} is proportional to the distance between two nodes. Table 5.2 summarizes the assumptions of parameters used in the MSCLP model and rigid designs. The order of magnitude and relative scale of the parameter values are based on discussions with a collaborating EMS provider – although the actual values cannot be disclosed.

The two rigid designs share these assumptions together as well as the flexible design. The construction cost of a flexible station is normally higher than that of a rigid station. This cost is assumed to be the same in this study in order to reveal the value of enabling flexibility. That is, the difference between the expected total costs of a rigid design and a flexible design is the upper bound of the value of enabling flexibility if the difference is positive, see (5.48). This value is the most that decision-makers should be willing to pay to enable the flexibility (e.g. purchase extra land, pay for shared infrastructures in an expandable station, etc.) The specific formula used to calculate the installation and expansion costs is as follows:

$$c_l = c_0 U_l^b \tag{5.46}$$

Eq. (5.46) incorporates economies of scale (*EoS*) into the calculation of costs, where b is the *EoS* factor. The average number of incidents per hour is used as the value of d_{itn}. This value is generated through a Geometric Brownian Motion (GBM) process with a particular expected growth rate (μ) and volatility (σ), as shown in Eq. (5.47). W_t is a Wiener process or so-called Brownian motion.

Table 5.2 List of assumptions for the flexible alternative.

Parameters	Value	Definition
$\boldsymbol{S}$	10	Number of strategic periods
$\boldsymbol{T}$	40	Number of tactical periods
$\boldsymbol{E_s}$	1 year	The length of a strategic period
$\boldsymbol{E_T}$	3 months	The length of a tactical period
$\boldsymbol{N}$	10	Number of scenarios
$\boldsymbol{L}$	4	Number of station phases
$\boldsymbol{U_l}$	1, 2, 3, 4	Unit capacity of a phase l station ($l = 1 \dots 4$)
$\boldsymbol{c_l}$ (\$million)	2, 3.73, 5.38, 6.96	Installation costs for a phase l station
$\boldsymbol{m_l}$ (\$million)	0.01, 0.02, 0.03, 0.04	Operation costs for a phase l station per tactical period
$\boldsymbol{c_u}$ (\$million)	2, 3.73, 5.38	Costs for expanding l unit phase
$\boldsymbol{c_v}$ (\$million)	0.1	Costs for purchasing one unit emergency vehicle
$\boldsymbol{m_v}$ (\$million)	0.01	Maintenance costs per emergency vehicle per tactical period
$\boldsymbol{CoV}$	0.95	Required incident coverage rate
$\boldsymbol{r_s}$	12%	Discount rate per strategic period
$\boldsymbol{r_t}$	2.83%	Discount rate per tactical period
$\boldsymbol{h}$	2160	Number of hours per tactical period
$\boldsymbol{\varepsilon, M}$	10^{-3}, 10^{6}	Small tolerance and large integer
$\boldsymbol{R}$	10	Coverage radius for a station
$\boldsymbol{t_{exp}}$	4	Number of tactical periods for flexibility regarding capacity expansion
$\boldsymbol{t_{dep}}$	4	Number of tactical periods for flexibility regarding phased deployment

GBM is a continuous-time stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion (also called a Wiener process) with drift (Ross, 2014). It is fairly safe to generate uncertain demand using the GBM process, since it consists of a mean growth rate (the drift – which is based on the idea that populations grow slowly in the long term) plus a random shock (the volatility – which can be normally distributed around the mean growth rate to simulate the *iid* process, since each evolutionary scenario might differ). The assumption of a growing population in the long term is reasonable because an emerging country or a developing district is considered in the thesis. Also, GBM is a Markov process, which means that *the future, given the present state, is independent of the past.* This process is thus useful for modeling cases where one thinks that the percentage changes (and not the absolute changes) are independent and identically distributed (Ross, 1995), e.g., the incident arrival rate in year 5 may have some connections with that in year 4, but it may have little or even no connection with that in year 1. Therefore, it is safe to make such assumptions and use the GBM process to capture the randomness of the uncertainty drivers (e.g., incident arrival rate, electricity demand, etc.) of interest.

$$d_{itn} = d_{i,0} \exp\left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right) \tag{5.47}$$

For modeling incident arrivals, it is assumed that the incident arrivals in each district are *iid* (independently and identically distributed), except for the initial starting point $d_{i,0}$. In addition, each scenario is assigned the same weight, meaning that the probability of occurrence is the same. Table 5.3 shows the values of the corresponding parameters in sequences, which are fictitious data. Figure 5.2 shows 1,000 samples generated for the out-of-sample test (described later), illustrating some outcomes of the GBM process.

Table 5.3 Parameters of uncertainty (i.e. incident rate per hour).

Parameter	Value	Definition
$\boldsymbol{\mu}$	5%	Quarterly expected growth rate
$\boldsymbol{\sigma}$	10%	Volatility
$\boldsymbol{p_n}$	$1/\vert N\vert$	Probability of scenario n $(N = 10)$

The value of flexibility is the difference between the expected total costs of the flexible design and a rigid design, as shown in Eq. (5.48).

$$
\begin{aligned}
E\left[Value\left(\text{Flex}\right)\right] = max\{E\left[Total\left(\text{Rigid}\right)\right] \\
- E\left[Total\left(\text{Flex}\right)\right], 0\}
\end{aligned}
\tag{5.48}
$$

where the expected value of total costs can be obtained through an out-of-sample analysis as introduced later. Since there is no cost premium for a flexible station considered here, this difference indicates the upper bound on the cost premium the decision-maker should be willing to pay to enable the flexibility in the EMS system. This cost may vary depending on the system, location, economic conditions, and therefore cannot be detailed here. The max condition captures the fact that if flexibility does not create positive value (e.g. cost savings), it will not be embedded – and therefore the lower bound is 0.

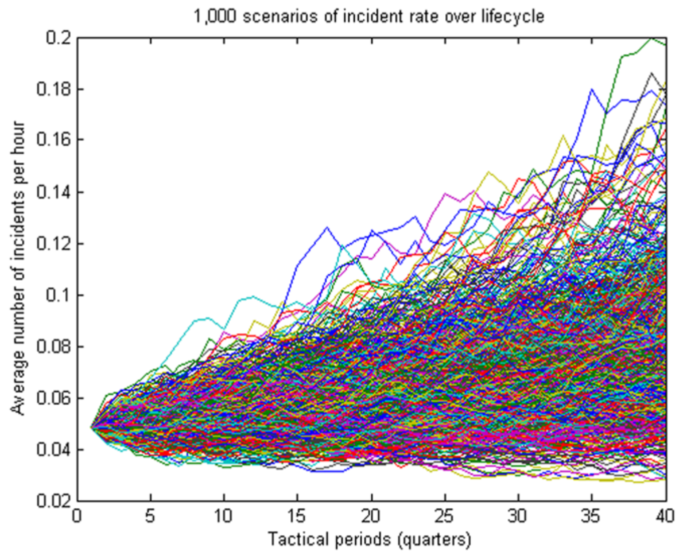


Figure 5.2 1,000 samples of incident rate in district No.1 throughout the system life cycle.

The computational analysis is run on a desktop machine with a 3.30 GHz CPU and 8.0 GB of random access memory (RAM). The design alternatives are coded in AIMMS 4.2 (AIMMS, 2014) and CPLEX 12.6 is used as the linear programming (LP) and mixed integer programming (MIP) solver. The default algorithm for MIP is the standard Branch and Bound (B&B), which terminates either when the relative gap is satisfied by the predetermined one (e.g., 1%) or the number of iterations reaches an upper bound (e.g., four millions).

5.3.1 Deterministic Analysis

This section describes the results of the analysis for the three designs based on the deterministic projection. The demand considered in this analysis is the expected growth over life cycle. Table 5.4 shows the characteristics and results of the three designs. The flexible design is somewhat (i.e., 1.1%) worse than the less flexible design, while the least flexible design performs the worst of three designs. The solving procedure for all designs terminate when the relative gap is less than or equal to the boundary before running out of time. It is observed that the flexible design is not necessarily the best in a deterministic case, because this analysis do not account for uncertainty and the ability to adapt. Based on decision rules, phased deployment and capacity expansion may be exercised if the rule is triggered. This may result in a higher incident coverage rate than expected (i.e., 0.95) and also more costs.

Table 5.4 Comparison for characteristics and results of three design alternatives based on the deterministic projection.

Alternatives	**No. of constraints**	**No. of integer variables**	**Best LP bound ($million)**	**Best Solution ($million)**	**Gap** (%)	**CPLEX Time (sec)**
Least Flex	5,256	4,441	13.92	14.05	0.98	38.34
Less Flex	5,706	4,810	13.65	13.79	0.99	55.58
Flex	6,841	5,444	13.80	13.94	1.00	230.41

5.3.2 Uncertainty Analysis

This section explicitly describes the analysis for rigid designs under uncertainty. For ease of comparison, the characteristics of the three alternatives in terms of the problem size and optimization results are shown together in Table 5.5. The expected total costs for the less flexible design is 0.7% worse than that of the least flexible design, considering the best solutions. The flexible design performs the worst across all three designs. The solution gap of the flexible design is consequently greater (e.g., 9.39% vs. 0.98% or 7.01%). This may be because the default algorithm terminates before finding a good solution.

Table 5.5 Comparison for characteristics and results of three design alternatives based on the stochastic projection.

Alternatives	**No. of constraints**	**No. of integer variables**	**Best LP bound ($million)**	**Best Solution ($million)**	**Gap** (%)	**CPLEX Time (sec)**
Least Flex	49,365	40,441	15.64	15.79	0.98	3,368.8
Less Flex	49,815	40,810	14.78	15.90	7.01	9,031.4
Flex	69,175	53,846	14.61	15.98	9.39	24,754.0

Table 5.6 and Table 5.7 summarize the design solutions for rigid systems, where columns 2 to 11 show the capacity (i.e., number of emergency vehicles) deployed in each district over the system's life cycle. As can be seen in Table 5.6, since the least flexible design deploys all capacity at once at the beginning of the life cycle, the capacity for a district does not change over time (no matter whether it is zero or nonzero). The less flexible design, in contrast, can deploy capacity based on a fixed plan. For instance, the planner may open a station in district 5 at the beginning of the second strategic period (year 2), and opens a station in district 7 at the beginning of the seventh strategic period (year 7) (see Table 5.7).

Table 5.6 Summary of the output for the least flexible design.

District	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1
5	2	2	2	2	2	2	2	2	2	2
6	0	0	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0

The out-of-sample analysis for rigid designs was conducted in order to find out the statistical properties about the optimal solutions. The corresponding results are shown in Table 5.10 and Figure 5.3 in Section 5.3.3. In generic statistics, an out-of-sample analysis is used to estimate a mathematical or statistical model based on the forecast of historical data. In this thesis, this analysis evaluates the optimal solution under various sample scenarios based on the same GBM process that are generated by Monte-Carlo simulation and were not considered in the solving procedure.

Table 5.7 Summary of the output for the less flexible design.

District	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	1	1
4	2	2	2	2	2	2	2	2	2	2
5	0	1	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0

5.3.3 Flexibility Analysis

The analysis for the flexible design is described in this section. Specifically, Table 5.8 summarizes the design solution for the flexible design. As can be seen, columns 2 to 5 list the values of the corresponding parameters in the decision rules for the flexible system. For example, to summarize the flexible solution for district no. 8, the planner should deploy initially $o_8^1 = 1$ unit capacity at the beginning of the life cycle in period 1 ($S = 1$). The value of o_j^1 here is simply equal to the summation of $U_l o_{jl}^1$ over l. If the number of missed incidents per year $\delta_8^d = 50$ is exceeded in 4 consecutive tactical periods (i.e., one strategic period), the station will deploy $o_8^u = 1$ unit of additional capacity. If the planner did not open a station in period 1 (which is recommended), they then should open a station in a later strategic period when the sum of the incident arrival rate in this period is greater than or equal to $\delta_8^o = 3.676$. For the other districts where $o_j^1 = 0$, initial capacity is deployed based on satisfying the flexible decision rule with the corresponding parameters. That is, if the sum of the incident arrival rate at site j in one strategic period is greater than or equal to δ_j^o and there is no station open yet, then the planner should consider opening one at this site at the beginning of the next strategic period. Note that the value of 0.001 in column 2 is the lower bound of δ_j^o and the value of 50 in column 4 is the upper bound of δ_j^d, respectively. Table 5.9 shows the solution for the flexible design in a particular sample scenario.

Table 5.8 Summary of the output for the flexible design.

District	$\boldsymbol{\delta_j^o}$	$\boldsymbol{o_j^o}$	$\boldsymbol{\delta_j^d}$	$\boldsymbol{o_j^u}$	$\boldsymbol{o_j^1}$
1	0.486	1	50	1	0
2	0.001	1	50	1	1
4	0.793	1	50	1	0
5	0.85	1	50	1	0
6	3.319	1	50	1	1
7	2.492	1	50	1	0
8	3.676	1	50	1	1
9	1.106	1	50	1	0
10	0.001	1	50	1	1

Table 5.9 Summary of the output for the flexible design in sample scenario 1.

District	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	2	2	2	2	2	2	2
3	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	1
7	0	0	0	0	0	0	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0
10	1	1	1	1	1	1	1	2	2	2

The out-of-sample analysis summarized below shows that the flexible design performs the best of all alternatives (see column 4 in Table 5.10). The results show that the flexible system is more adaptable to unforeseen scenarios – which is one of the purposes of a flexible solution – even though it is not the best in the optimization analysis. The expected total costs of the flexible design are 4.6% less than for the least flexible design, while it is quite close to the expected total costs of the less flexible design. This study shows, nevertheless, that the flexible design is more adaptable to all 1,000 out-of-sample scenarios. The least and the less flexible designs are infeasible in some scenarios, leading the solver to output value cost 0. The expected value of solutions thus decreases artificially but non-negligibly due to the reduction in the percentage of feasible solutions, and affects the performance measures (e.g., mean, standard deviation (STD)). To make a fair comparison, 1,000 scenarios are selected to be feasible for all three design alternatives. As can be seen in Table 5.10, the expected total cost of the flexible design is 4.9% and 0.9% less than those of rigid design solutions. The expected cost differences between the design alternatives are all statistically significant according to the corresponding t-test. For example, the p-value for the comparison between the flexible and less flexible designs is 0.003023, which is much less than 0.05. In fact, the number of samples considered in the out-of-sample analysis (i.e., $n = 1{,}000$) is much higher than t-tests and other statistical tests are designed to deal with (i.e., $n \leq 30$ samples). The flexible design also has the lowest costs for P5 and P95 of three designs. The terms P5 and P95 are abbreviation of the 5th and 95th percentiles. A percentile is a statistical measure indicating the value below which a given percentage of observations in a group of observations fall. For example, the 95th percentile is the value below which 95 percent of the observations are found. The reason that the flexible design has the greatest standard deviation is because the flexible system could allocate budget anywhere if necessary. Rigid designs, on the other hand, do not have much flexibility to allocate budget since the capacity is deployed based on a fixed plan.

Figure 5.3 provides an illustration of the out-of-sample analysis in terms of cumulative density function. As can be seen, there is a clear benefit from designing for flexibility in terms of the whole cumulative distribution and the mean value, compared to the least flexible design. The value of flexibility between the flexible and the less flexible designs is, however, not as much as expected under current assumptions of uncertainty drivers and parameters (i.e., $14.91 - 14.77 = 0.14$ \$million).

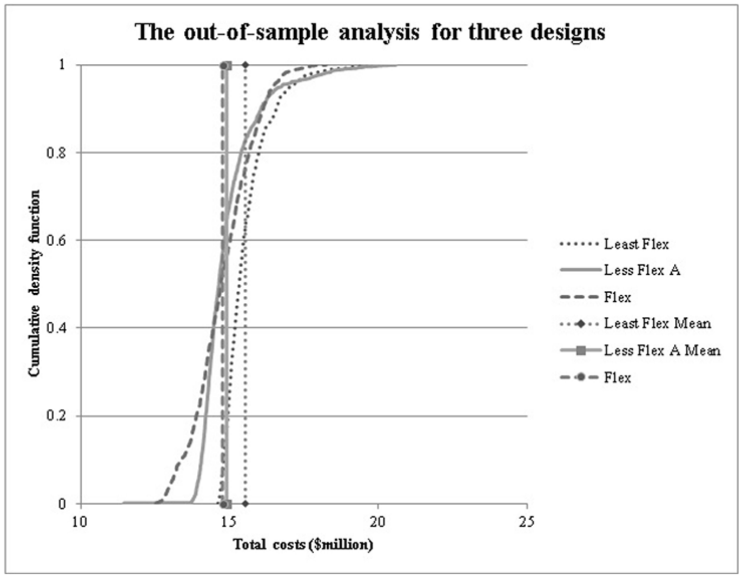


Figure 5.3 Cumulative density function of the out-of-sample analysis (CoV = 0.95, $|N|$ = 1,000).

Table 5.10 Characteristics and results for the out-of-sample analysis ($CoV = 0.95$, $|N| = 1{,}000$).

Alternatives	**No. of constraints**	**No. of integer variables**	**Mean ($million)**	**STD ($million)**	**STD error ($million)**	**P5 ($million)**	**P95 ($million)**	**Total time (sec)**
Least Flex	4,082	4,000	15.53	0.776	0.025	14.74	17.03	866.4
Less Flex	4,802	4,000	14.91	0.988	0.031	13.95	16.74	1,184.7
Flexible	6,179	4,972	14.77	1.023	0.032	13.06	16.49	2,470.4
Best?	-	-	Flex	Least Flex	Least Flex	Flex	Flex	Lease Flex

5.3.4 Sensitivity Analysis

The sensitivity analysis next evaluates how uncertainty in the performance output is affected by variability in input parameter assumptions. To see how the significant parameters (e.g., incident coverage rate, station coverage radius) influence the expected system performance, a one-factor-at-a-time (OFAT) approach is used and the results are shown as Tornado diagrams in Figure 5.4. The three terms "base", "low", and "high" represent different values for different parameters, which are shown in Table 5.11. The low and high values equal $-/+$ 50% the base value, except for incident coverage rate. This is because such variation makes little sense in the coverage rate when $CoV \leq 0.8$.

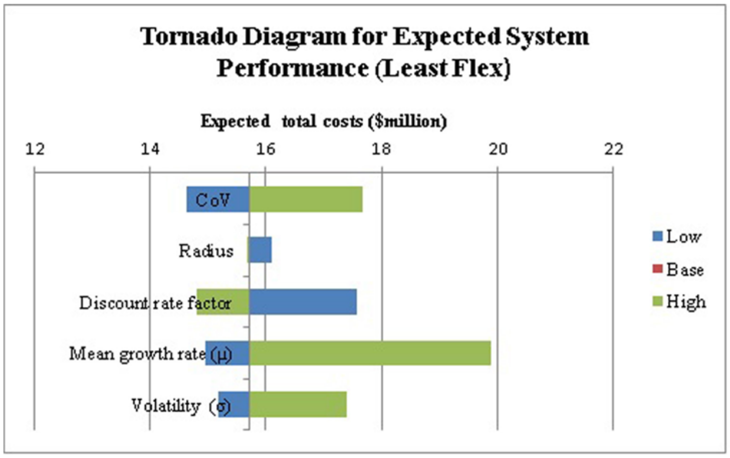
Table 5.11 Values for the significant parameters in sensitivity analysis.

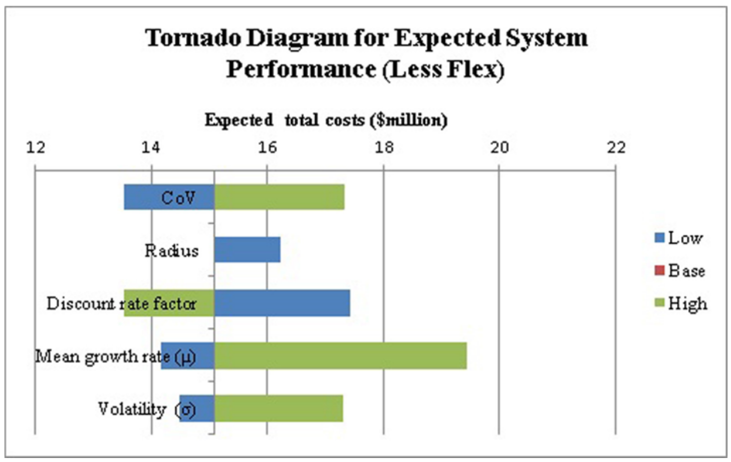
	Low	**Base**	**High**
Incident coverage rate ($\boldsymbol{CoV}$)	0.90	0.95	0.98
Coverage radius	5	10	15
Discount rate factor per year	6%	12%	18%
Mean growth rate ($\boldsymbol{\mu}$)	2.5%	5%	7.5%
Volatility ($\boldsymbol{\sigma}$)	5%	10%	15%

Of the five parameters, the mean growth rate (μ) is the most influential for all three designs (refer to Figure 5.4). This makes sense because when incident rates grow fast (slow), the planner has to spend more (less) on assignment costs (in rigid designs) or may enable more (less) flexibility (in the flexible design) in order to satisfy the requirement on incident coverage rate (i.e., 0.95).

It indicates that the flexible design may be more favorable than the other designs if the population of the region or city of interest grows quite rapidly (i.e., more people leads to more incidents). This also implies that the rigid systems may not perform as well as it was designed to due to the increasing incident rate. This is because the capacity of stations is deployed based on the previous forecast and it cannot be changed over time. When demand increases, maintaining the required coverage rate may be impossible in some cases. The coverage rate CoV and discount rate are also influential on both low and high sides for all three designs, especially for rigid ones. Rigid designs cannot change the system capacity accordingly, and thus lack the ability to meet higher coverage rates if requested. As CoV increases (decreases), expected total cost naturally increases (decreases). An increasing (decreasing) discount rate reduces (increases) the net present value of costs. On the other hand, the flexible design is less sensible to the change of volatility σ than rigid designs. As the flexible design has the ability to change according to the realization of different uncertainty scenarios, it manages to save costs through deploying and/or expanding capacity in later periods when needed, thereby mitigating unnecessary capacity deployment. It again indicates that the flexible design may be a better choice when dealing with a situation that the emergency incident arrival rate is highly unstable, either it grows fast or it fluctuates frequently, or even both. The above analysis also shows that the flexible design may be more suitable for a developed city as those input parameters are usually less stable (and thus difficult to forecast accurately) in such region.

Figure 5.5 shows the results for sensitivity analysis focusing on the value of flexibility. This value is defined in Eq. (5.48), which is the difference between the expected total costs of a rigid design and the flexible design. The value of flexibility is affected mostly by the mean growth rate, and then by the volatility, and CoV. Again, the faster the incident rate grows, the more valuable flexibility is in the face of growing uncertainty. As volatility increases, flexibility is also worth more, as it enables the system to deal better with changing conditions. At the other end of the spectrum, little volatility brings little value of flexibility, since there is no need for change. Interestingly, the results show that flexibility is more valuable when there is a higher requirement on coverage. When comparing the flexible to the least flexible, the value of flexibility increases with an increasing discount rate because a higher discount rate provides more incentives to defer capacity deployment to later – thereby making more use of the flexibility. The impact of the discount rate is reduced when compared to the less flexible design, since capacity is deployed over time in both cases.





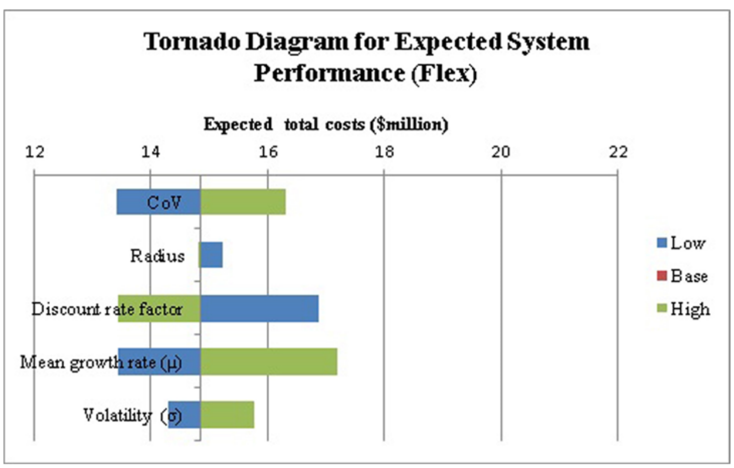


Figure 5.4 Tornado charts for expected total costs of three designs.

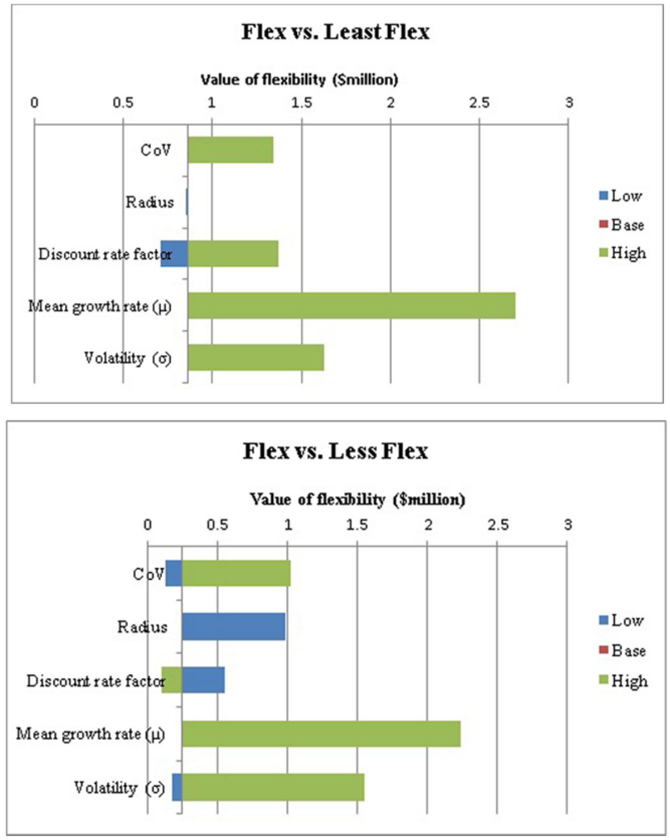


Figure 5.5 Tornado charts for value of flexibility.

The results from the main analysis are obtained under the assumption that the required coverage rate is 0.95. There is a clear tradeoff between coverage and budget (e.g., in the limit, an infinite budget would converge towards 100% coverage). For a comprehensive comparison among the three designs, a sensitivity analysis is also conducted by increasing gradually coverage rate requirement from 0.8 to 0.99, and a Pareto set is introduced in Figure 5.6 where the coverage rate and the amount of the total costs are the x- and y-axis, respectively. The dots represent the average total costs correlating to a specific design and coverage rate, whose value is obtained through the exact same procedure as the example out-of-sample test above. The curve shows the corresponding shape of the total costs for each design alternative.

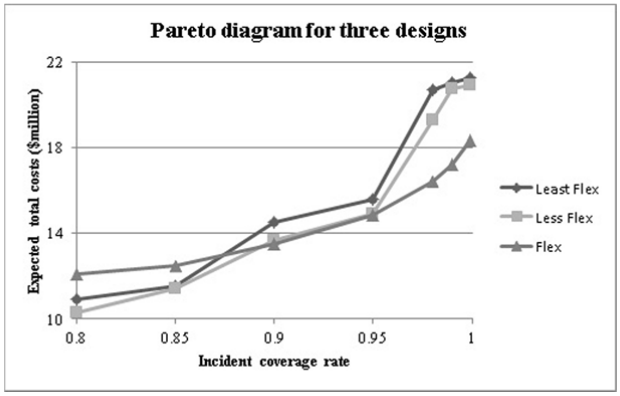


Figure 5.6 The outcome of the Pareto analysis for three designs.

The flexible design dominates the other two alternatives when the required coverage rate is greater than 0.9. Thus, a flexible EMS design performs much better than more rigid alternatives when the required coverage rate is considerably high (i.e., CoV from 0.95 to 0.99). To achieve such a high-level coverage rate, one needs to deploy capacity at different sites with more than one unit (i.e., the least flexible design), or deploy capacity over time and space (i.e., less flexible and flexible designs). The flexible design performs better than the two rigid alternatives because it deploys capacity based on uncertainty realizations. In contrast, the performance of the flexible design is worse than the less flexible design when the required coverage rate is less than 0.9, and it is even worse than the least flexible design when the rate is less than 0.88. In this model, it is assumed that a decision regarding station installation must be implemented if the rule is triggered, as well as for capacity expansion. This makes the flexible design more expensive when the system requires less coverage rate (thus less capacity). This assumption makes sense, however, in the system management process, as decisions can only be made considering information available at any given time. In addition, one may expect that the least and less flexible designs would not reach the coverage requirements if there are out-of-sample scenarios (i.e., not used in the optimization process). Those are clearly handled better by the flexible design, as demonstrated by the results of the out-of-sample study.

Chapter 6 CASE STUDY TWO – STRATEGIC FLEXIBILITY ANALYSIS REGARDING THE DEPLOYMENT OF NUCLEAR POWER PLANTS UNDER UNCERTAINTY

"Mind and hand." – Wang Shouren (1472 – 1529)

This chapter explicitly describes an engineering application about how to best site nuclear power plants under uncertainty under demand and public/social acceptance by taking into account the expected long term life cycle performance, and the concept of flexibility as a means to deal with such uncertainty drivers. In addition to the requirement on electricity demand, public or social acceptance is considered as a significant factor that can terminate the project in this study. This nuclear system with regards to site nuclear power plants was analyzed based on the methodology described in Chapter 4, as the energy system is an example of infrastructure systems. More specifically, a two-stage multi-period mathematical model based on the SAA scheme is developed to find out the stochastically optimal initial configuration (e.g., how much capacity to be deployed at the beginning of the project), and the best decision rules for exercising the corresponding flexibility (e.g., when to expand the capacity for a nuclear power plant). Section 6.1 explicitly describes the background information about the nuclear project. Section 6.2

presents the mathematical model in details, with explanations for notations, the objective function, and constraints. The results are explicitly shown in the remaining sections. In Section 6.3, electricity demand is the main uncertainty driver, while both electricity demand and public acceptance are uncertainty drivers in Section 6.4. The numerical analysis follows the procedure described in Section 4.3.

6.1 Step 1: Design Problem Description

The global demand for electricity has rapidly increased in recent decades. The world electricity consumption was 13,174 TWh in 2000, and increased to 20,301 TWh (54% growth) in 2014 (Enerdata, 2015). According to the Global Energy Statistical Yearbook 2015, China and the United States are the only two countries that have consumed more than 1,000 TWh per year since 2000, and the BRIC (Brazil, Russia, India, and China) countries shared 37% of world electricity consumption in 2014. Also, there is clear growth in the demand for electricity in developing countries such as China and India. Traditional thermal power plants generate electricity from the combustion of fossil fuels and thus contribute significantly to CO_2 emissions. In the hope of reducing global warming caused by greenhouse effects, renewable or green energy sources with little or zero CO_2 emissions have been receiving much attention. Compared to expanding renewable energy sources like solar or wind power, nuclear power has a relatively stable output of electricity because it requires no specific weather or environmental conditions to generate electricity. Furthermore as a generator with high fixed costs and very low marginal costs baseload generation is favored even if load-following operation is technically possible (Pouret *et al.*, 2009). Even though it has potential safety issues, for example concerning the large amount of energy stored in the nuclear reactor core and the need to cool the reactor system even after shutdown. And noting ongoing concerns around long term waste management, nuclear power is still a significant energy source for generating electricity in most countries. For now, nuclear power is one of the two main sustainable, dependable, and low-emission sources of energy (the other is hydropower).

There is, however, much uncertainty and risks associated with siting nuclear power plants from an economic standpoint (Kessides, 2010). This uncertainty will ultimately affect the total costs or the levelized cost of electricity (LCOE) generation, which are metrics for evaluating economic performance. More specifically, the total costs consist of four major components: construction costs, operations and management (O&M) costs, fuel costs, and back-end costs, where construction costs represent approximately 60% of the total costs (Joskow, 2006). In addition to the economic and safety factors discussed above, there are several other factors that may be involved in the problem of siting nuclear power plants. Compared to technological or economic factors, the effect of social factors is difficult to measure and quantify. Social factors that may affect the decision-making processes include public acceptance of nuclear technology and government support. Public or social acceptance is “essential for any activity that affects large sectors of a nation” (Golay, 2001). It represents the public intention of siting more/less nuclear power plants. Although nuclear power is said to be safe, nuclear accidents with extremely low probability, such as the ones in Chernobyl and Fukushima, can be worldwide disasters. Such catastrophes can affect public acceptance of nuclear energy in distant countries. For example, after the Fukushima accident occurred in Japan, both Germany and Switzerland announced plans to gradually shut down all nuclear power plants in use by 2025 (Joskow & Parsons, 2012).

6.2 Step 2: Analytical Model

The flexible design deals pro-actively with uncertain drivers by leveraging strategic-level flexibility. More specifically, three strategic-level real options are introduced in the design so that the nuclear power system can change and adapt according to variations in demand as well as external nuclear events over time. Firstly, the model accounts for a flexible phased deployment of plant capacity over time and space. The flexible design allows the deployment of plants over time, depending on the realization of the uncertainty scenario. In addition to flexible phased deployment, flexible capacity expansion at any strategic period is studied. Once a plant is deployed, the capacity can be increased (or uprated) when needed and when external circumstance is allowable. This flexible expansion benefits from special attention in terms of architecture and the design of the infrastructure (e.g., designing for smaller capacity first, and carefully planning for expansion in the future by installing new nuclear reactor, for instance). The capacity of a nuclear power plant is related to its ability to generate a given amount of electricity. Thirdly, the flexible design allows the life extension of a plant when it is supposed to be closed, depending on the external circumstance at that moment (i.e., whether the cumulative INES (international nuclear events scale) is below the threshold or not). INES is a tool for promptly and consistently communicating the safety significance of events associated with sources of ionizing radiation to the public. The cumulative INES is the summation of INES indices for past several years. Compared to a single INES index, this cumulative index helps represent the historical impact of nuclear events in a reasonable timeframe.

6.2.1 Symbols and Terminology

The following list summarizes the modeling notation (sets and parameters):

T = the set of strategic periods in a complete life cycle ($t \in T$);

T_1 = the set of strategic periods excluding the first period (i.e., $t \geq 1$).

J = the set of candidates to site nuclear power plants ($j \in J$);

S = the set of uncertainty scenarios ($s \in S$);

V = the set of plant phases ($v \in V$);

r = discount rate factor per strategic period;

p_s = the corresponding probability of scenario s;

v^*, v^{**} = the smallest and biggest phases of a plant, respectively;

U', U^* = the smallest and biggest capacity of a plant, respectively;

$C_{fix}, C_{rac}, C_{var}, C_{fuel}$ = the fixed costs, unit cost of a nuclear reactor, unit variable cost, and unit cost for fuel, respectively;

C_p = the unit cost of losing demand for electricity. The penalty cost would occur when power plants cannot generate enough electricity to meet the requirement on demand;

t_c, t_e, t_l = the delay periods for new plant construction and capacity expansion, and the additional period for life extension, respectively. For instance, a plant is decided to be deployed at strategic period $t = 0$, and this plant will be available for use starting from $t = t_c$;

t_d = the review period for the decision rules regarding phased deployment and capacity expansion;

t_{LE} = the first period that life extension is available;

θ_j = transmission loss factor of electricity for a plant at site j;

ε, M = a small tolerance and an arbitrary large integer, used to ensure a given constraint is always or never satisfied.

The random variables considered in the model are as follows:

d_{ts} = the demand for electricity within strategic period t under scenario s;

I_{ts} = the cumulative INES factor in strategic period t under scenario s.

6.2.2 Mathematical Formulations

Several assumptions must be made when describing a flexible design via a mathematical model. First, a plant will not be used again if it is closed, and this closure is effective immediately. Physically the plant cannot be shut down immediately because there are many things needs to do after the normal closure (e.g., waste management). This closure however can be considered effective immediately from the economic standpoint of view, as the plant will not generate electricity after its closure. Second, decisions are made at the beginning of a strategic period and only observed information can be used for the decision-making processes. When the cumulative INES falls into the dead zone, the whole project will be terminated immediately. When the cumulative INES falls into the warning zone, the project can still operate, but strategic-level decisions can only be made when the cumulative index returns to safe zone. In other words, flexible phased deployment, capacity expansion, and life extension of a plant are only available when public acceptance is fairly positive (i.e., the cumulative INES falls into the safe zone). Moreover, the land is assumed to be reserved for a fairly long period because the geographical requirement for installing a nuclear power plant is strict. There will be only a few candidate sites that are acceptable, and thus they need to be reserved for further investment. The specific decision variables used in this model are as follows:

o_{jv}^{1} = 1 if a phase v plant is deployed at site j in strategic period $t = 0$;

o_{jts}^2 = 1 if a plant is deployed at site j in strategic period $t \geq 1$ under scenario s;

u_{jts} = 1 if the capacity of a plant is expanded at site j in strategic period $t \geq 1$ under scenario s;

l_{jts} = 1 if the life cycle of a plant is extended at site j in strategic period $t \geq 1$ under scenario s;

x_{jvts} = 1 if a phase v plant is open at site j in strategic period t under scenario s;

$x_d \in (0,1)$, the threshold for triggering decision rules regarding phased deployment and capacity expansion;

q_1, q_2 = integers, the thresholds for partitioning the cumulative INES index into three zones, e.g., if this index is greater than or equal to q_2, it indicates that the cumulative index in t has fallen into the dead zone;

n_c, n_e = integers, the number of plants to be deployed or expanded if the decision rule is triggered, respectively;

m_c, m_e = integers, the capacity of a plant that should be deployed or expanded if the decision rule is triggered, respectively;

CO_{jts}, CE_{jts} = integers, the capacity of a plant to be deployed or expanded at site j in strategic period t under scenario s;

y_{ts}^1, y_{ts}^2 = 1 if the cumulative INES factor falls into the warning or dead zones in strategic period t under scenario s, respectively;

DR = binary, decision rules related variables;

ω_{ts}, ξ_{ts} = non-negative, used to express the difference between the demand for electricity and the electricity generated by plants;

δ = binary, indicator variable.

The objective of the model is to minimize the expected total costs over the life cycle of the system. The total costs comprise fixed costs, variable operation and maintenance costs, variable fuel costs, and penalty costs. The costs spent since period 1 are discounted back to period 0.

$$\begin{aligned} \min \ & \sum_j (C_{fix} o^1_{jv^*} + \sum_v C_{rac} o^1_{jv}) + \sum_s p_s \sum_t r_t (\sum_j [C_{fix} o^2_{jts} \\ & + C_{rac}(CO^2_{jts} + CE_{jts}) + \sum_v (C_{\mathrm{var}} + C_{fuel}) x_{jvts}] + C_p w_{ts}) \end{aligned} \tag{6.1}$$

Formulations (6.2)-(6.4) describe the relationships and limits of the decision variables. More specifically, inequalities (6.2) and (6.4) indicate that a plant has the ability to generate an amount of electricity that is less than or equal to the capacity of this plant. Inequality (6.3) shows that a plant can only be installed once at a site.

Subject to:

$$o^1_{j,v+1} \le o^1_{j,v} \ j \in J, v \in V; \tag{6.2}$$

$$o^1_{jv^*} + \sum_t o^2_{jts} \le 1 \ j \in J, s \in S; \tag{6.3}$$

$$x_{j,v+1,ts} \le x_{jvts} \ j \in J, v \in V, t \in T, s \in S; \tag{6.4}$$

The following inequalities imply the relationship within the decision-making processes. Formulations (6.5)-(6.8) indicate that the capacity to be initially deployed at a candidate site from period 1 is a constant, and so is the capacity of expansion. The number of plants to be deployed or expanded at different time should be consistent, as shown in (6.9)-(6.12). $DR_{ts}^{c} = 1$ indicates that the decision rule regarding flexible phased deployment is triggered.

$$U'o_{jts} \le U'CO_{jts} \le U^{*}o_{jts} \quad j \in J, t \in T_1, s \in S; \tag{6.5}$$

$$U'\left(1-o_{jts}^{2}\right) \le U'\left(m_c - CO_{jts}\right) \le U^{*}\left(1-o_{jts}^{2}\right) \quad j \in J, t \in T_1 \;\; s \in S \tag{6.6}$$

$$U'u_{jts} \le U'CE_{jts} \le U^{*}u_{jts} \quad j \in J, t \in T_1, s \in S; \tag{6.7}$$

$$U'\left(1-u_{jts}\right) \le U'\left(m_e - CE_{jts}\right) \le U^{*}\left(1-u_{jts}\right) \quad j \in J, t \in T_1 \;\; s \in S \tag{6.8}$$

$$DR_{ts}^{c} \le \sum\nolimits_{j} o_{jts}^{2} \le \left|J\right| DR_{ts}^{c} \quad t \in T_1, s \in S; \tag{6.9}$$

$$1 - DR_{ts}^{c} \le n_c - \sum\nolimits_{j} o_{jts}^{2} \le \left|J\right|\left(1 - DR_{ts}^{c}\right) \quad t \in T_1, s \in S; \tag{6.10}$$

$$DR_{ts}^{e} \le \sum\nolimits_{j} u_{jts} \le \left|J\right| DR_{ts}^{e} \quad t \in T_1, s \in S; \tag{6.11}$$

$$1 - DR_{ts}^{e} \le n_e - \sum\nolimits_{j} u_{jts} \le \left|J\right|\left(1 - DR_{ts}^{e}\right) \quad t \in T_1, s \in S; \tag{6.12}$$

Constraints (6.13)-(6.18) indicate how to determine the value of variables regarding the cumulative INES factors. When this cumulative factor falls into the dead zone, the entire project is terminated immediately. Therefore, the restriction regarding the decision-making processes when $y_{ts}^{1} = 1$ still holds for $y_{ts}^{2} = 1$, as shown in (6.17) and (6.18). Variable y_{ts}^{1} can be 0 only if the current cumulative INES is less than the threshold q_1 and the last cumulative INES did not fall into the dead zone ((6.13)-(6.14)).

$$I_{ts} \le q_1 - \varepsilon + My_{ts}^{1} + My_{ts}^{2} \quad t \in T, s \in S; \tag{6.13}$$

$$I_{ts} \ge q_1 - M\left(1 - y_{ts}^{1}\right) - My_{ts}^{2} \quad t \in T, s \in S; \tag{6.14}$$

$$I_{ts} \le q_2 - \varepsilon + My_{ts}^{2} + My_{t-1,s}^{2} \quad t \in T, s \in S; \tag{6.15}$$

$$I_{ts} \geq q_2 - M\left(1 - y_{ts}^2\right) - My_{t-1,s}^2 \;\; t \in T, s \in S; \tag{6.16}$$

$$y_{ts}^2 \leq y_{ts}^1 \;\; t \in T, s \in S; \tag{6.17}$$

$$y_{t-1,s}^2 \leq y_{ts}^2 \; t \in T, s \in S; \tag{6.18}$$

There are several cases when the decision rules regarding flexible phased deployment and capacity expansion are not allowed even though the condition of the rules are satisfied. First, the cumulative acceptance of the system (or project) should be within the safe zone (see (6.19)). Besides, any construction of new plants or expansion of current plants would make the decision rules unavailable, as can be found in (6.20) and (6.21). If the capacity of the system is fully deployed, or all the plants expire after life extension, the embedded decision rules are not allowed either. Inequality (6.24) shows the condition when decision rules are available for this system.

$$DR_{ts}^A \leq 1 - y_{t-1,s}^1 \;\; t \in T_1, s \in S; \tag{6.19}$$

$$DR_{ts}^A \leq 1 - \delta_{ts}^{DR1} \;\; t \in T_1, s \in S; \tag{6.20}$$

$$DR_{ts}^A \leq 1 - \delta_{ts}^{DR2} \;\; t \in T_1, s \in S; \tag{6.21}$$

$$DR_{ts}^A \leq 1 - \delta_{ts}^{DR3} \;\; t \in T_1, s \in S; \tag{6.22}$$

$$DR_{ts}^A \leq 1 - \delta_{ts}^{DR4} \;\; t \in T_1, s \in S; \tag{6.23}$$

$$DR_{ts}^A \geq 1 - y_{t-1,s}^1 - \delta_{ts}^{DR1} - \delta_{ts}^{DR2} - \delta_{ts}^{DR3} - \delta_{ts}^{DR4} \;\; t \in T_1, s \in S; \tag{6.24}$$

The following formulations explicitly describe the conditions when the indicator variables, with respect to decision rule availability, can be 1. For example, constraints (6.25) and (6.26) show that indicator variables will be 1 if an expansion was implemented in previous t_e periods, and 0 otherwise.

Similarly, indicator variables will be 1 if an initial installation or phased deployment was implemented within previous t_c periods, see (6.27)-(6.30). When all candidate sites are fully occupied by plants with maximal capacity, no decision rule regarding phased deployment and capacity expansion needs to be considered. When all plants are finished operation after life extension, the decision rules will also be unavailable. These cases can be found in (6.31) -(6.34).

$$\sum\nolimits_{t-t_e+1}^{t-1} \sum\nolimits_j u_{jts} \leq M\,\delta_{ts}^{DR1} \quad t \in T_1, s \in S; \tag{6.25}$$

$$\sum\nolimits_{t-t_e+1}^{t-1} \sum\nolimits_j u_{jts} \geq \varepsilon - M\left(1-\delta_{ts}^{DR1}\right) \quad t \in T_1, s \in S; \tag{6.26}$$

$$\sum\nolimits_{t-t_c+1}^{t-1} \sum\nolimits_j \left(o_{jts}^2 + o_{jv^*}^1\right) \leq M\delta_{ts}^{DR2} \quad t = 0,\ldots,t_c, s \in S; \tag{6.27}$$

$$\sum\nolimits_{t-t_c+1}^{t-1} \sum\nolimits_j (o_{jts}^2 + o_{jv^*}^1) \geq \varepsilon - M\left(1-\delta_{ts}^{DR2}\right) t = 0,\ldots,t_c, s \in S; \tag{6.28}$$

$$\sum\nolimits_{t-t_c+1}^{t-1} \sum\nolimits_j o_{jts} \leq M\,\delta_{ts}^{DR2} \quad t = t_c+1,\ldots,T, s \in S; \tag{6.29}$$

$$\sum\nolimits_{t-t_c+1}^{t-1} \sum\nolimits_j o_{jts}^2 \geq \varepsilon - M\left(1-\delta_{ts}^{DR2}\right) \quad t = t_c+1,\ldots,T, s \in S \tag{6.30}$$

$$\sum\nolimits_j x_{jv^{**}ts} \leq |J| - \varepsilon + M\,\delta_{ts}^{DR3} \quad t \in T, s \in S; \tag{6.31}$$

$$\sum\nolimits_j x_{jv^{**}ts} \geq |J| - M\left(1-\delta_{ts}^{DR3}\right) \quad t \in T, s \in S; \tag{6.32}$$

$$\sum\nolimits_j \delta_{jts}^{LE3} \leq |J| - \varepsilon + M\,\delta_{ts}^{DR4} \quad t = t_{LE},\ldots,T, s \in S; \tag{6.33}$$

$$\sum\nolimits_j \delta_{jts}^{LE3} \geq |J| - M\left(1-\delta_{ts}^{DR4}\right) \quad t = t_{LE},\ldots,T, s \in S; \tag{6.34}$$

The difference between the demand for electricity and the capacity of the plants is described in (6.35)-(6.37). If the system loses $x_d\%$ demand for t_d consecutive periods, then the decision rules regarding phased deployment and capacity expansion will be triggered (see (6.38)-(6.41)). However, these rules can only be implemented when they are both necessary available and triggered. This idea is represented in (6.42)-(6.44).

$$d_{ts} - \sum_j \theta_j U' \sum_v x_{jvts} = \omega_{ts} - \xi_{ts} \ t \in T, s \in S; \tag{6.35}$$

$$\omega_{ts} \le M\delta_{ts}^{D} \ t \in T, s \in S; \tag{6.36}$$

$$\xi_{ts} \le M\left(1 - \delta_{ts}^{D}\right) \ t \in T, s \in S; \tag{6.37}$$

$$\omega_{ts} \le d_{ts} x_d - \varepsilon + M\delta_{ts}^{DM} \ t \in T, s \in S; \tag{6.38}$$

$$\omega_{ts} \ge d_{ts} x_d - M\left(1 - \delta_{ts}^{DM}\right) \ t \in T, s \in S; \tag{6.39}$$

$$\sum_{t-t_d}^{t-1} \delta_{ts}^{DM} \le t_d - \varepsilon + MDR_{ts}^{T} \ t \in T, s \in S; \tag{6.40}$$

$$\sum_{t-t_d}^{t-1} \delta_{ts}^{DM} \ge t_d - M\left(1 - DR_{ts}^{T}\right) \ t \in T, s \in S; \tag{6.41}$$

$$DR_{ts}^{Tr} \le DR_{ts}^{T} \ t \in T, s \in S; \tag{6.42}$$

$$DR_{ts}^{Tr} \le DR_{ts}^{A} \ t \in T, s \in S; \tag{6.43}$$

$$DR_{ts}^{Tr} \ge DR_{ts}^{T} + DR_{ts}^{A} - 1 \ t \in T, s \in S; \tag{6.44}$$

In several cases, flexible phased deployment and capacity expansion cannot possibly be implemented. If there is no plant in the system, capacity expansion is of course unavailable. Similarly, if all sites are occupied, phased deployment becomes unavailable. These are represented in (6.45)-(6.46) and (6.50)-(6.51). When these conditions are false and the decision rules can be implemented, planners will first consider implementing capacity expansion. Flexible phased deployment will be implemented if capacity expansion is forbidden, as shown in (6.47)-(6.49) and (6.52)-(6.55).

$$\sum_j \left(x_{jv^*ts} - x_{jv^{**}ts}\right) \ge \varepsilon - M\left(1 - \delta_{ts}^{E}\right) \ t \in T, s \in S; \tag{6.45}$$

$$\sum_j \left(x_{jv^*ts} - x_{jv^{**}ts} \right) \leq M\delta_{ts}^E \;\; t \in T, s \in S; \tag{6.46}$$

$$DR_{ts}^e \leq \delta_{ts}^E \;\; t \in T, s \in S; \tag{6.47}$$

$$DR_{ts}^e \leq DR_{ts}^{Tr} \;\; t \in T, s \in S; \tag{6.48}$$

$$DR_{ts}^e \geq \delta_{ts}^E + DR_{ts}^{Tr} - 1 \;\; t \in T, s \in S; \tag{6.49}$$

$$\sum_j \left(x_{jv^*ts} + \delta_{jts}^{LE3} \right) \leq |J| - \varepsilon + M\delta_{ts}^C + M\delta_{ts}^E \;\; t \in T, s \in S; \tag{6.50}$$

$$\sum_j \left(x_{jv^*ts} + \delta_{jts}^{LE3} \right) \geq |J| - M\left(1 - \delta_{ts}^C\right) - M\delta_{ts}^E \;\; t \in T, s \in S; \tag{6.51}$$

$$\delta_{ts}^C \leq 1 - \delta_{ts}^E \;\; t \in T, s \in S; \tag{6.52}$$

$$DR_{ts}^c \leq \delta_{ts}^C \;\; t \in T, s \in S; \tag{6.53}$$

$$DR_{ts}^c \leq DR_{ts}^{Tr} \;\; t \in T, s \in S; \tag{6.54}$$

$$DR_{ts}^c \geq \delta_{ts}^C + DR_{ts}^{Tr} - 1 \;\; t \in T, s \in S; \tag{6.55}$$

Formulations (6.56)-(6.60) show the condition when life extension can be implemented. Specifically, if public acceptance is favorable and a plant is ready for extension, planners should consider implementing this flexibility. Constraints (6.61)-(6.68) indicate that site-specific life extension is ongoing, while constraints (6.69)-(6.70) imply that such extension is finished.

$$l_{jts} \leq 1 - y_{t-1,s}^1 \;\; j \in J, t = t_{LE}, \ldots, T, s \in S; \tag{6.56}$$

$$l_{jts} \leq o_{jv^*}^1 + My_{t-1,s}^1 \;\; j \in J, t = t_{LE}, s \in S; \tag{6.57}$$

$$l_{jts} \geq o_{jv^*}^1 - My_{t-1,s}^1 \;\; j \in J, t = t_{LE}, s \in S; \tag{6.58}$$

$$l_{jts} \leq o_{j,t-t_{LE}+1,s}^2 + My_{t-1,s}^1 \;\; j \in J, t = t_{LE}+1, \ldots, T, s \in S; \tag{6.59}$$

$$l_{jts} \geq o_{j,t-t_{LE}+1,s}^2 - My_{t-1,s}^1 \;\; j \in J, t = t_{LE}+1, \ldots, T, s \in S; \tag{6.60}$$

$$\delta_{jts}^{LE2} \leq o_{jv^*}^{1} + l_{jts} \;\; j \in J, t = t_{LE}, s \in S; \tag{6.61}$$

$$\delta_{jts}^{LE2} \geq o_{jv^*}^{1} - l_{jts} \;\; j \in J, t = t_{LE}, s \in S; \tag{6.62}$$

$$\delta_{jts}^{LE2} \geq l_{jts} - o_{jv^*}^{1} \;\; j \in J, t = t_{LE}, s \in S; \tag{6.63}$$

$$\delta_{jts}^{LE2} \leq 2 - o_{jv^*}^{1} - l_{jts} \;\; j \in J, t = t_{LE}, s \in S; \tag{6.64}$$

$$\delta_{jts}^{LE2} \leq o_{j,t-t_{LE}+1,s}^{2} + l_{jts} \;\; j \in J, t = t_{LE}+1,\ldots,T, s \in S; \tag{6.65}$$

$$\delta_{jts}^{LE2} \geq o_{j,t-t_{LE}+1,s}^{2} - l_{jts} \;\; j \in J, t = t_{LE}+1,\ldots,T, s \in S; \tag{6.66}$$

$$\delta_{jts}^{LE2} \geq l_{jts} - o_{j,t-t_{LE}+1,s}^{2} \;\; j \in J, t = t_{LE}+1,\ldots,T, s \in S; \tag{6.67}$$

$$\delta_{jts}^{LE2} \leq 2 - o_{j,t-t_{LE}+1,s}^{2} - l_{jts} \;\; j \in J, t = t_{LE}+1,\ldots,T, s \in S; \tag{6.68}$$

$$\delta_{jts}^{LE1} \leq \delta_{jt-1,s}^{LE1} + l_{jt-t_l,s} \;\; j \in J, t = t_{LE}+t_l,\ldots,T, s \in S; \tag{6.69}$$

$$\delta_{jts}^{LE1} \geq \frac{1}{2}\left(\delta_{jt-1,s}^{LE1} + l_{jt-t_l,s}\right) \;\; j \in J, t = t_{LE}+t_l,\ldots,T, s \in S; \tag{6.70}$$

Constraints (6.71)-(6.74) show when regular operation (including life extension) is done at a site. Formulations (6.75)-(6.79) are used to describe the conditions when decision-making regarding deployment and expansion are impossible at a specific site.

$$\delta_{jts}^{LE3} \leq \delta_{jts}^{LE1} + \delta_{jts}^{LE2} + \delta_{jt-1,s}^{LE3} \;\; j \in J, t \in T, s \in S; \tag{6.71}$$

$$\delta_{jts}^{LE3} \geq \delta_{jts}^{LE1} \;\; j \in J, t \in T, s \in S; \tag{6.72}$$

$$\delta_{jts}^{LE3} \geq \delta_{jts}^{LE2} \;\; j \in J, t \in T, s \in S; \tag{6.73}$$

$$\delta_{jts}^{LE3} \geq \delta_{jt-1,s}^{LE3} \;\; j \in J, t \in T, s \in S; \tag{6.74}$$

$$u_{jts} \leq x_{jv^*ts} \;\; j \in J, t \in T_1, s \in S; \tag{6.75}$$

$$u_{jts} \leq 1 - x_{jv^{**}ts} \;\; j \in J, t \in T_1, s \in S; \tag{6.76}$$

$$o_{jts}^{2} \leq 1 - x_{jv^*ts} \;\; j \in J, t \in T_1, s \in S; \tag{6.77}$$

$$u_{jts} \leq 1-\delta_{jts}^{LE3} \quad j \in J, t \in T_1, s \in S; \tag{6.78}$$

$$o_{jts}^2 \leq 1-\delta_{jts}^{LE3} \quad j \in J, t \in T_1, s \in S; \tag{6.79}$$

The last inequalities explicitly show the relationship between the capacity of a plant and corresponding decisions, such as deployment and expansion. The capacity of a plant in period t will be equal to the capacity of this plant in the last period plus the expected capacity by deployment and/or expansion, when public acceptance is favorable and the plant (system) is still in operation. Otherwise, the capacity will be 0 until the end of the project.

$$\begin{aligned}&\sum_v x_{jvts} \leq \sum_v x_{jvt-1,s} + CO_{jt-t_c,s} + CE_{jt-t_e,s} + My_{t-1,s}^2 \\ &+M\delta_{jts}^{LE1} + M\delta_{jts}^{LE2} \quad j \in J, t \in T_1, s \in S;\end{aligned} \tag{6.80}$$

$$\begin{aligned}&\sum_v x_{jvts} \geq \sum_v x_{jvt-1,s} + CO_{jt-t_c,s} + CE_{jt-t_e,s} - My_{t-1,s}^2 \\ &-M\delta_{jts}^{LE1} - M\delta_{jts}^{LE2} \quad j \in J, t \in T_1, s \in S;\end{aligned} \tag{6.81}$$

$$\sum_v x_{jvts} \leq M\left(1-y_{t-1,s}^2\right) \quad j \in J, t \in T_1, s \in S; \tag{6.82}$$

$$\sum_v x_{jvts} \leq M\left(1-\delta_{jts}^{LE1}\right) \quad j \in J, t = t_{LE}, \ldots, T, s \in S; \tag{6.83}$$

$$\sum_v x_{jvts} \leq M\left(1-\delta_{jts}^{LE2}\right) \quad j \in J, t = t_{LE} + t_l, \ldots, T, s \in S. \tag{6.84}$$

6.3 Step 3a: Numerical Analysis – One Uncertainty Driver

In the next two sections, four alternative design solutions are analyzed and compared in two numerical studies. Study 1 accounts for uncertainty in electricity demand only, and rely on an optimistic projection for social acceptance. Study 2 accounts for uncertainty in both electricity demand and social acceptance. Table 6.1 summarizes the strategic-level decisions involved in the four designs alternatives considered across two case studies. The first design is termed the "rigid design" and it, deploys all resources (i.e., nuclear power plants) at once at the beginning of the system's life cycle in an optimal manner in light of anticipated scenarios in social acceptance and electricity demand. This one represents a robust approach to capacity deployment under uncertainty, and aims at maximizing cost effectiveness considering a wide range of uncertainty scenarios. The second and third designs are simplified versions of a fully flexible design and strategic-level flexibility is only partially embedded. More specifically, the second design only considers the ability to deploy capacity in phase along with flexible capacity expansion subject to electricity demand uncertainty, while the third design only takes life extension into account as flexibility in the face of uncertainty in social acceptance. Thus, the second and third options are referred to as "flexible design A" and "flexible design B", respectively. The flexible design alternative that considers all strategic-level flexibility is called "flexible design C". It should be noted that rigid and Flex A designs do not incur any extension under both case studies. Designs Flex B-C incur a "forced" life extension in case study 1 (i.e. social acceptance is assumed satisfactory so operations can continue longer), while under case study 2 life extension is flexible, and governed by decision rules. In case study 2, all designs have an added flexibility of early shutdown if social acceptance is not satisfactory. The purpose of comparing flexible design C to the three other alternatives is to determine: 1) the significance of embedded flexibility strategies and 2) the value of flexibility based on uncertainty realization. The assumptions for the case studies are listed below and have been adapted from the paper by Steer *et al.* (2012). Note that several figures are valid as of 2006, and some cost figures might have changed since then. It is assumed that the salvage value of the equipment is approximately equal to the disposal costs for radioactive waste management.

Table 6.1 The design features of the four design alternatives across the two case studies.

Alternatives	**Case 1**	**Case 2**
Rigid	All capacity deployed at time 0, no life extension	All capacity deployed at time 0, early shutdown, no life extension
Flex A	Phased deployment + capacity expansion, no life extension	Phased deployment + capacity expansion, early shutdown, no life extension
Flex B	All capacity deployed at time 0, forced life extension	All capacity deployed at time 0, early shutdown, flexible life extension
Flex C	Phased deployment + capacity expansion, forced life extension	Phased deployment + capacity expansion, early shutdown, flexible life extension

There are two uncertainty drivers considered in this model, namely electricity demand (d_{ts}) and public acceptance of nuclear technology (I_{ts}). In this analysis, the demand is generated through a Geometric Brownian Motion (GBM) process with a particular expected growth rate (μ) and volatility (σ), as shown in Eq. (5.47). Public acceptance is captured by the INES factor. Since this factor is intended to be logarithmic, it is assumed that it is generated through Eqs. (6.85) and (6.86). Parameter i denotes the level of INES factor, and $p(i)$ is the corresponding probability. More specifically, the probability of one level is approximately $1/\beta$ that of the previous level. Parameter β is referred here as a magnification factor on the logarithmic INES scale.

$$p(i) = p(0)\beta^{-i} \quad i = 1,\ldots 7; \tag{6.85}$$

$$\sum_i p(i) = 1 \quad i = 0,\ldots 7. \tag{6.86}$$

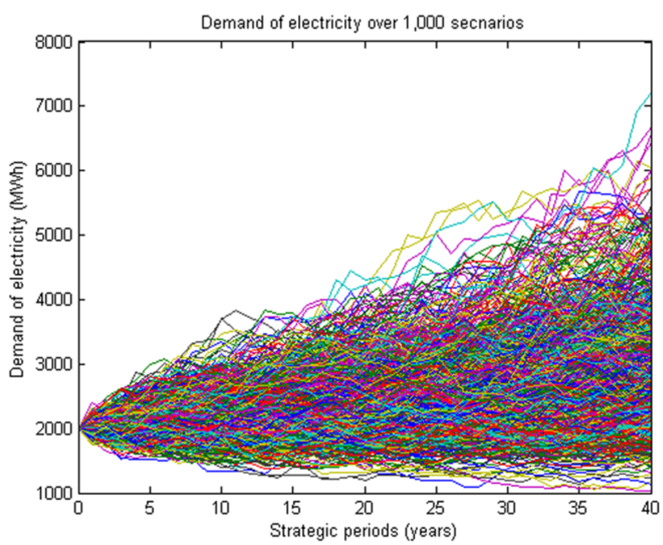


Figure 6.1: Demand for electricity throughout plant lifetime over 1,000 scenarios.

Table 6.2 List of assumptions for the numerical analysis (£million).

Parameter	Assumption
Declared net capacity (DNC) of a nuclear reactor	600 MWe per reactor
Pre-development costs (PD)	£250 in 2006 money
Construction period of initial deployment ($\boldsymbol{t_c}$)	6 years
Construction period of capacity expansion ($\boldsymbol{t_e}$)	3 years
Additional period of life extension ($\boldsymbol{t_l}$)	10 years
Construction costs of a nuclear power plant (w/o reactor) ($\boldsymbol{C_{fix}}$)	Nominally £524 = £250 (PD) + £274 (IDC)
Costs of reactor ($\boldsymbol{C_{rac}}$)	Per reactor: nominally £975
O&M costs of reactors ($\boldsymbol{C_{var}}$)	Nominally £3.85/MWh per plant, followed by a £3.85/MWh per reactor
Fuel supply costs for thorium ($\boldsymbol{C_{fuel}}$)	Nominally £1.1/MWh
Contractual costs for losing demand ($\boldsymbol{C_p}$)	Nominally £200/MWh
Number of uncertainty scenarios ($\boldsymbol{\lvert S \rvert}$)	1,000 scenarios
Corresponding probability for each scenario ($\boldsymbol{p_s}$)	0.001 (1/1,000)
Maximal phase (or number of reactors) for a nuclear plant ($\boldsymbol{\lvert V \rvert}$)	4 reactors
Review period for flexible phased deployment and capacity expansion ($\boldsymbol{t_d}$)	3 years
Review period for life extension	6 years
Costs of capital or discount rate ($\boldsymbol{r}$)	Per period: nominally 10%
Small tolerance ($\boldsymbol{\varepsilon}$)	Nominally 10^{-3}
Large integer ($\boldsymbol{M}$)	Nominally 10^{6}

Table 6.3 summarizes the assumptions for the significant parameters in Eqs., (6.85) and (6.86), as well as in the GBM process. The values of these parameters are based on communications with domain experts. Changes in these values assumptions are analyzed in the sensitivity analysis done later in this section. Figure 6.1 illustrates some outcomes of the GBM process

generated for the out-of-sample analysis.

Table 6.3 List of assumptions for uncertainty drivers.

Parameter	Assumption
Initial demand per hour in period 0 ($\boldsymbol{d_0}$)	Nominally 2,000 MWh
Expected growth rate ($\boldsymbol{\mu}$)	Per period: nominally 1%
Volatility ($\boldsymbol{\sigma}$)	Nominally 5%
Magnification ($\boldsymbol{\beta}$)	Nominally 3.5

Two case studies are presented in the following two sections. In the first case study, the long term demand for electricity is the only uncertainty driver considered; in the second case study, both long term demand and public acceptance are considered. The purpose of the case studies is to find out the value of flexibility under different assumptions for uncertainty. This value is thus the difference between the expected total costs of any flexible design and the rigid design, as shown in Eq. (5.48). The assumption about cost premium for enabling flexibility is the same as the design for EMS systems in Chapter 5. This cost may vary depending on various issues like the system; location, technology, the need for shared infrastructure to plan for capacity expansion, purchase additional land, and other socio-economic conditions that are difficult to detail here, and are out of scope. The expected value of total costs can be obtained through out-of-sample analysis, which will be introduced later.

This case focuses on how to site and manage the nuclear power system capacity over the next 40 years, considering long term demand for electricity as the main uncertainty driver. It aims to site nuclear power plants in a hypothetical developing country that wishes to further develop and expand nuclear power capacity. The number of candidate sites for the plants is relatively small due to the availability in the country. This project is intended to generate sufficient electricity to satisfy the electricity needs of this hypothetical country. Table 6.4 shows the assumptions for case study 1. The electricity generated by a plant can be transmitted via the power grid network to customers. The corresponding transmission loss factor is simply assumed to be inversely proportional to the distance.

Table 6.4 List of specific assumptions for case study 1.

Parameter	**Assumption**
Number of candidates ($\|\boldsymbol{J}\|$)	3
Operational lifetime of the project ($\|\boldsymbol{T}\|$)	40 years
Operational lifetime of a power plant	20 years
Transmission loss factor of electricity ($\boldsymbol{\theta_j}$)	1 (site 1), 0.98 (site 2), and 0.95 (site 3)

The computational analysis was run on a desktop machine with a 3.30 GHz CPU and 8.0 GB of random access memory (RAM). All design alternatives were coded in AIMMS 4.2 (AIMMS, 2014) and CPLEX 12.6 was used as the MIP solver. The default algorithm for MIP is Branch and Bound (B&B), which terminates when the relative gap is smaller than or equal to 0.01, or the number of iterations reaches an upper bound (e.g., 20 million), or the computation time reaches an upper bound (e.g., 10,000 seconds).

6.3.1 Deterministic Analysis

This section describes the analysis based on a deterministic projection (e.g., the expected value) of demand for electricity over life cycle. As can be seen, flexible design B performs the best of the four designs, which is about 9.95% less than that of the rigid design. It is interesting that flexible design C is not the best. This is because the expected demand for electricity is relatively flat, and the phased deployment and capacity expansion thus are not as much valuable as expected.

Table 6.5 Characteristics and results for deterministic analysis in case study 1.

Alternative	**No. of constraints**	**No. of integer variables**	**Best LP bound (£billion)**	**Best solution (£billion)**	**Gap (%)**	**CPLX time (sec)**
Rigid	1,966	976	25.47	25.47	0.00	0.06
Flex A	4,302	1,464	25.45	25.45	0.00	2.29
Flex B	1,888	976	23.19	23.19	0.00	0.19
Flex C	4,393	1,479	23.66	23.66	0.00	1.47

6.3.2 Uncertainty Analysis

Table 6.6 shows the characteristics of four alternatives (one rigid design plus three flexible designs) in terms of the problem size and optimization results, for the ease of comparison. The rigid design performs the worst of four design alternatives in terms of the expected total costs.

Table 6.6 Characteristics and results for design alternatives under uncertainty in case study 1.

Alternative	**No. of constraints**	**No. of integer variables**	**Best LP bound (£billion)**	**Best solution (£billion)**	**Gap (%)**	**CPLX time (sec)**
Rigid	19,543	9,652	27.04	27.04	0.00	1.28
Flex A	42,903	14,496	24.78	24.78	0.00	224.17
Flex B	19,543	9,652	24.76	24.76	0.00	3.39
Flex C	43,813	14,646	24.28	24.28	0.00	299.6

6.3.3 Flexibility Analysis

This section describes the expected performance of flexible designs. The results are shown in Table 6.6. The expected total costs for flexible design C are 10.21% less than for the rigid design, considering the best solutions. The performance measures (e.g., mean, STD) can be found in Table 6.8. It is also better than flexible designs A-B. Figure 6.2 and

Table 6.7 illustrate the solutions obtained for the four designs. As can be seen for the rigid design solution, four units of capacity is deployed at site 1 at year 0 (which will be available at year 6), and retired in year 26. For the flexible design A, 3 unit capacity is deployed at site 2 (i.e., $o_{2,3}^{1} = 1$) at the beginning of the life cycle, while 4 unit capacity is deployed at site 1 (i.e., $o_{1,4}^{1} = 1$) for flexible designs B-C. For flexible design A (C), if installed capacity is not able to provide for more than $x_d = 0.0001\%$ (2.16%) of the electricity demand for 3 consecutive years, the planner may expand capacity by $m_e = 1$ (1) unit capacity at $n_e = 1$ (1) non-empty sites, or deploy $m_c = 3$ (1) unit capacity at $n_c = 1$ (1) empty sites, depending on the feasibility of the expansion phase. That is, if there are power plants that are not fully deployed, the capacity expansion will be implemented. Otherwise, new power plants will be installed at empty sites. Note that if all candidate sites are non-empty and the power plants are fully deployed, the flexibility strategies will not be implemented even if the condition is satisfied. Since no social acceptance is considered in case study 1, the life extension strategy is always implemented. The performance measures for the out-of-sample analysis (e.g., mean, STD) can be found in Table 6.8 based on 1,000 scenarios generated through the GBM process.

Flexible design C is the best of the four designs in terms of all performance measures except P5. This is because the benefit of phased deployment and capacity expansion may not be greater than the cost for enabling those flexibility strategies when the electricity demand increases slowly or even decreases in the long term. However, the difference of the expected total costs between flexible design C and flexible design B is not statistically significant (p-value = 0.07067) under current assumptions. This indicates that flexible design B may be as good as flexible design C when the discount rate is equal to 10%. The expected value of flexibility for flexible design A is £26.38 – £25.97 = £0.41 billion pounds and the expected value of flexibility for flexible design B is £26.38 – £24.24 = £2.14 billion pounds. This demonstrates that life extension is more valuable than flexible phased deployment and capacity expansion under the current assumptions.

Furthermore, the value of flexibility for flexible design C is less than the linear sum of the values of flexibility for flexible designs A and B (i.e., £2.32 < £0.41 + £2.14 = £2.55). This shows that enabling different real option strategies in one system can decrease the value of flexibility because those may interact in some cases (e.g. flexible design A may be lagging or deploying capacity not fast enough under some scenarios, and therefore life extension may not be as profitable in some cases). Figure 6.3 provides an illustration of the out-of-sample analysis in terms of cumulative density functions.

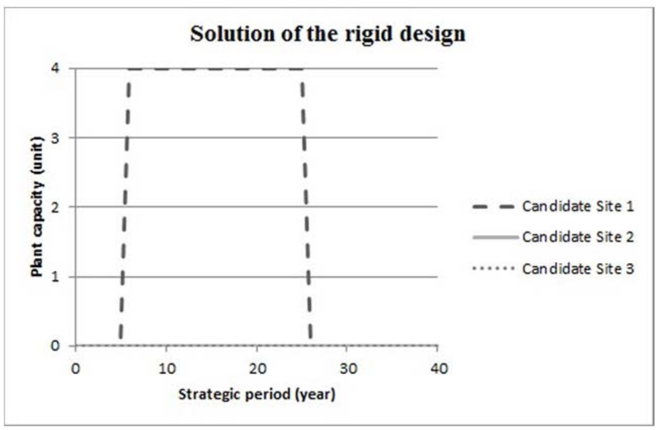


Figure 6.2 Graphic illustration of the solution obtained for the rigid system in case study 1 ($r = 10\%$).

Table 6.7 Tabular illustration of solutions obtained by the flexible designs in case study 1.

Alternatives	**Initial configurations**	x_d	m_e	m_c	n_e	n_c
Flex A	$o_{2,3}^{1} = 1$	0.000001	1	3	1	1
Flex B	$o_{1,4}^{1} = 1$	N.A	N.A	N.A	N.A	N.A
Flex C	$o_{1,4}^{1} = 1$	0.021573	1	1	1	1

Table 6.8 Results for out-of-sample analysis in case study 1 (£billion).

Alternative	Mean	STD	STD error	P5	P95	Value of flexibility
Rigid	26.38	2.83	0.089	23.12	32.03	-
Flex A	25.97	2.42	0.077	22.49	30.35	0.41
Flex B	24.24	2.71	0.086	21.30	29.75	2.14
Flex C	24.06	1.58	0.050	22.10	27.21	2.32
Best?	Flex C	Flex C	Flex C	Flex B	Flex C	Flex C

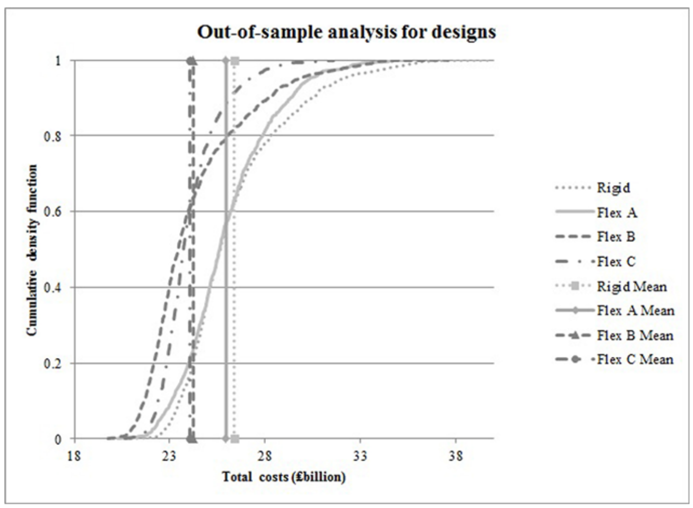


Figure 6.3 Cumulative density functions for out-of-sample analysis in case study 1.

6.3.4 Sensitivity Analysis

The sensitivity analysis shows how uncertainty in the output is affected by variability in the input parameter assumptions. A one-factor-at-a-time (OFAT) approach is applied to identify how the significant parameters (e.g., expected growth rate, discount rate) influence the expected system performance. The corresponding results are shown as Tornado diagrams in Figure 6.4. The terms "base", "low", and "high" represent different values for the significant parameters, which are listed in Table 6.9. The low and high values equal -/+50% of the base value, except for the magnification. This is because such variation makes little sense in the magnification when $\beta \leq 3$ is taken into account (i.e., the probability of occurrence for the high level events is considerably large). Of the four parameters, the penalty costs for losing demand are the most influential for all the design alternatives. This is because the greater (lower) the penalty costs, the greater (lower) the expected total costs spent on the system. The rigid design only allows the system to operate for a certain period, and it cannot satisfy the contractual demand in the last 14 years. Flexible designs B and C have the ability to operate 10 more years, and these designs are thus penalized much less than the rigid one. The discount rate is also influential for the designs, especially the rigid design (see Figure 6.4). This makes sense because for the rigid design, all capacity is deployed when the project begins, while for flexible designs A and C, such capacity can be deployed in later periods, and thus lowering expected total costs. Although flexible design B cannot deploy capacity over time and space as the other two flexible designs, it allows for life extension with 0 costs at a fixed later period, thereby reducing the total expected costs because the operational cost is much less than the penalty cost (i.e., the contractual penalty cost for losing demand). Since flexible design C is the combination of flexible designs A and B, it has all three real option strategies and can as well reduce the total expected costs. The expected growth rate μ is also influential on both the low and high sides for designs without options for flexible phased deployment and capacity expansion.

Table 6.9 Values for the significant parameters in the sensitivity analysis.

Parameter	**Low**	**Base**	**High**
Discount rate	5%	10%	15%
Expected growth rate ($\boldsymbol{\mu}$)	0.5%	1%	1.5%
Volatility ($\boldsymbol{\sigma}$)	2.5%	5%	7.5%
Magnification ($\boldsymbol{\beta}$)[1]	3	3.5	4
Contractual or penalty costs for losing demand ($\boldsymbol{C_p}$)	100	200	300

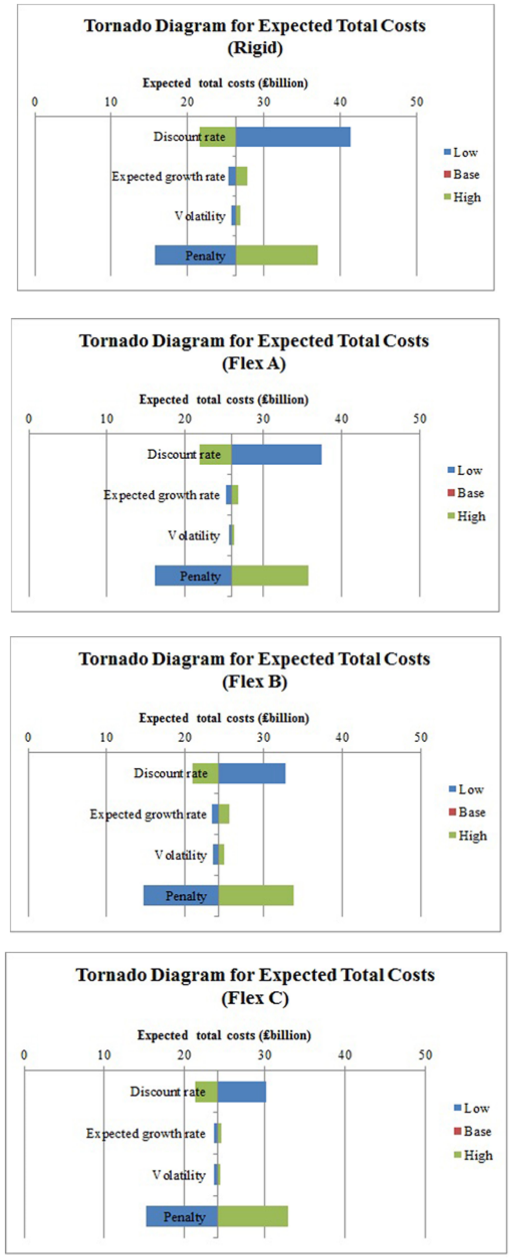
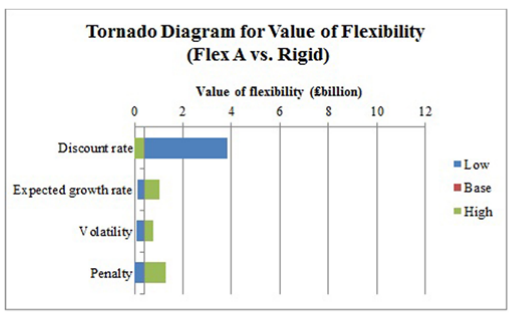
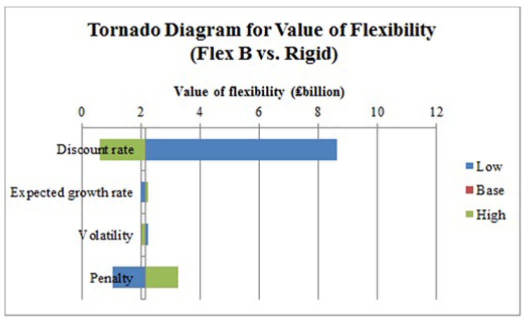


Figure 6.4 Tornado charts for the expected performance in case study 1.

Figure 6.5 shows the results of the sensitivity analysis focusing on the value of flexibility, which is defined in Eq. (5.48). The value of flexibility is considered as the difference between the expected total costs of the flexible design and the rigid design. As can be seen, the value of flexibility is most affected by the discount rate, then by the expected growth rate and volatility. Again, the lower (higher) the discount rate, the more (less) valuable flexibility is in the face of uncertainty. When the discount rate is lower (high), costs occurring later on weigh more (less) in the present value term – see Eq. (6.1). Therefore, the ability to avoid unnecessary cost deployment as enabled by the flexible alternatives should have a bigger (lower) impact in terms of expected discounted cost savings when the discount rate is lower (higher). Because the amount of flexibility increases in flexible designs A-C as compared to the rigid design, this is also why the value of flexibility increases between the flexible designs when the discount rate is varied by the same amount across all three flexible alternatives.



The value of flexibility also decreases when the penalty cost decreases. This makes sense because the planner can consider losing some demand instead of deploying new capacity to satisfy the demand. When the penalty cost is low, it eventually results in more savings for the rigid design, and thus reduces the value of flexibility. As the expected growth rate increases, flexibility also becomes worth more as it enables the system to deal better with changing conditions. At the other end of spectrum, a small expected growth rate results in lower value for flexibility, since there is no need for change. The results show that flexible design B is slightly affected by expected growth rate and volatility. This is because in case study 1, decision-making regarding life extension is not related to the demand for electricity. Since there is no such public acceptance (and therefore no INES index) considered in this first case study, this strategy will always be exercised if the nuclear power plant is available for the extension.



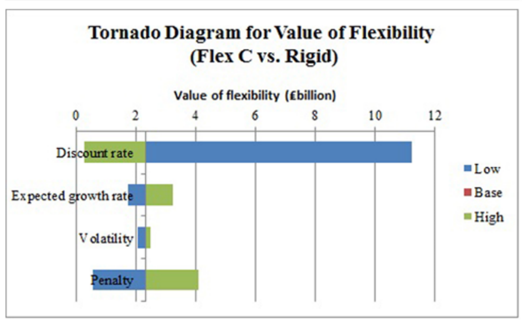


Figure 6.5 Tornado charts for value of flexibility in case study 1.

The results of the main analysis were obtained under the assumption that the discount rate is 10% and the penalty cost is £200/MWh. The sensitivity analysis demonstrates that the value of flexibility is significantly affected by the discount rate. Also, there is a clear tradeoff between implementing flexibility and paying penalty costs. In order to make a clear comparison of these designs, a sensitivity analysis was also conducted by gradually decreasing the discount rate from 10% to 0%. Figure 6.6 shows the results of this sensitivity analysis when the contractual or penalty costs for losing demand are equal to £100/MWh, £200/MWh, and £300/MWh. The x coordinate denotes the discount rate, while the y coordinate denotes the expected total costs. The dots represent the expected total costs correlating to a specific design and discount rate under the same assumption of penalty costs. The value of the dots is obtained by the exact same procedure as the example for the out-of-sample analysis shown above.

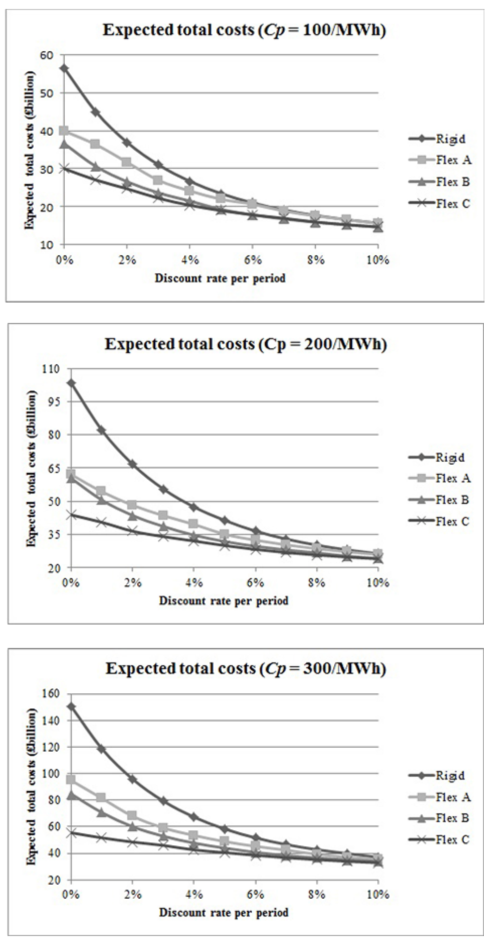


Figure 6.6 The outcomes of the Pareto analysis in case study 1.

As can be seen, flexible design C performs the best of the four designs. More specifically, flexible design C performs much better than the others when the discount rate is less than or equal to 4%. Compared to the rigid design, the value of flexibility (i.e., the difference between a given curve and the curve for the rigid system) is considerably high when the discount rate is less than or equal to 8%, and it increases along with increasing penalty costs. This makes sense because the embedded flexibility can reduce the costs due to losing demand. In addition, life extension is more valuable than flexible phased deployment and capacity expansion in this system, as shown in Figure 6.5. This is because life extension can be implemented with no additional costs (i.e., others than continuing to pay existing costs longer), while implementing phased deployment and capacity is not free (e.g., construction costs, reactor costs). The value of flexibility for phased deployment and capacity expansion is especially negligible when the penalty costs are quite low (e.g., less than 100/MWh) and the discount rate is fairly small (e.g., less than 5%). This also makes sense because the planner would prefer to pay the penalty costs for losing demand than to implement flexibility in such conditions. Moreover, the rigid design performs the worst across all four designs, even though the gap between the rigid design and flexible designs decreases along with the decreasing penalty costs. It is possible that the rigid design may be the solution when the penalty cost is fairly small (e.g., equal to the selling price for electricity), which may not be realistic.

6.4 Step 3b: Numerical Analysis – Two Uncertainty Drivers

The background information for case study 2 is similar to case study 1, except for the assumptions about the uncertainty drivers. In this case study, both electricity demand and public acceptance are considered. As a significant input, the impact of public acceptance of nuclear technology on decision-making regarding nuclear power plants is our major concern. In the proposed model, decision-making processes are heavily affected by public acceptance. That is, ongoing processes may immediately be terminated if nuclear technology is no longer in favor with the public (e.g., typically occurs when the cumulative INES falls into the dead zone). Also, the lower bounds for the warning and dead zones are fairly low at 5 and 7, respectively. This means that any nuclear event that reaches level 7 can end the project. The assumptions about other significant parameters, such as correlated costs and annual demand for electricity, are the same as in case study 1.

6.4.1 Deterministic Analysis

This section describes explicitly the results based on a deterministic analysis, considering the expected value of demand for electricity and INES as inputs. The results are shown in Table 6.10. It is found that the results are the same as ones in Table 6.5. This is because the expected INES for each year is too small, which can be considered as 0 in general. Therefore, the impact of INES is not fully revealed and the deterministic version of case study 2 is the same as that of case study 1.

Table 6.10 Characteristics and results for deterministic analysis in case study 2.

Alternative	No. of constraints	No. of integer variables	Best LP bound (£billion)	Best solution (£billion)	Gap (%)	CPLX time (sec)
Rigid	1,966	976	25.47	25.47	0.00	0.06
Flex A	4,302	1,464	25.45	25.45	0.00	2.29
Flex B	1,888	976	23.19	23.19	0.00	0.19
Flex C	4,393	1,479	23.66	23.66	0.00	1.47

6.4.2 Uncertainty Analysis

As did in Section 6.3.2, the characteristics and results of uncertainty analysis and flexibility analysis are combined in a single table for comparison shown in Table 6.11 in this section, while Figure 6.8 in Section 0 is used for illustrating the out-of-sample analysis of rigid and flexible designs.

Table 6.11 Characteristics and results for design alternatives in case study 2.

Alternative	No. of constraints	No. of integer variables	Best LP bound (£billion)	Best solution (£billion)	Gap (%)	CPLX time (sec)
Rigid	24,383	13,015	27.59	27.59	0.00	1.75
Flex A	54,453	16,698	27.60	27.60	0.00	4.57
Flex B	24,723	11,044	25.76	25.76	0.00	2.35
Flex C	55,353	16,698	25.77	25.77	0.00	5.35

6.4.3 Flexibility Analysis

This section describes explicitly the analysis regarding flexible designs. Table

6.11 lists the characteristics of the four design alternatives for case study 2, while Figure 6.7 and Table 6.12 illustrate the solutions obtained for the four designs. More specifically, Figure 6.7 compares the capacity evolution at site 1 under two sample scenarios, out of the ten scenarios considered for the optimization. The plant at site 1 is closed at period 18 (which is earlier than its life time) because public acceptance is not favorable in sample scenario 2, as the cumulative INES falls into the dead zone at that time. For flexible design A, 3 units of capacity are deployed at site 1 (i.e., $o_{1,3}^{1} = 1$) at the beginning of the life cycle, while 4 units of capacity are deployed at site 1 (i.e., $o_{1,4}^{1} = 1$) for flexible designs B-C. For flexible design A (C), if capacity installed loses more than $x_d = 0.0001\%$ (1.796%) of electricity demand for 3 consecutive years, the planner may expand capacity $m_e = 1$ (1) unit capacity at $n_e = 1$ (1) non-empty sites, or deploy $m_c = 4$ (1) unit capacity at $n_c = 1$ (1) empty sites, depending on the feasibility of the expansion. If the cumulative INES is greater than or equal to $q_1 = 5$ but less than $q_2 = 6$, the system can still be operated but no strategic-level decisions can be implemented. When the cumulative INES is greater than or equal to $q_2 = 6$, the system will be shut down immediately.

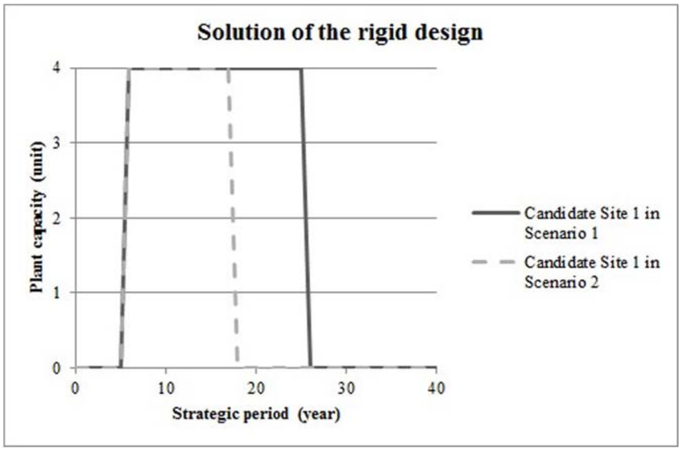


Figure 6.7: Graphic illustration of the solution at site 1 obtained for the rigid system in case study 2 ($r = 10\%$).

Table 6.12 Tabular illustration of solutions obtained by the flexible designs in case study 2.

Alternatives	**Initial configurations**	x_d	m_e	m_c	n_e	n_c	q_1	q_2
Flex A	$o_{1,3}^1 = 1$	0.0001	1	4	1	1	5	7
Flex B	$o_{1,4}^1 = 1$	N.A	N.A	N.A	N.A	N.A	5	7
Flex C	$o_{1,4}^1 = 1$	0.018	1	1	1	1	5	7

Table 6.13 shows the values of the significant performance measures. As can be seen, flexible design B has the lowest expected total costs for all designs, which are 6.43% less than the highest cost. The expected costs for flexible design C, however, are close to the expected costs for flexible design B and are about the same for flexible design A and the rigid design (see Figure 6.8). The statistical analysis shows that the difference between flexible design C and flexible design B, as well as the difference between flexible design A and the rigid design, are not statistically significant. The difference between

flexible design C and flexible design A, however, is statistically significant in terms of its p-value (2.03×10^{-21}). In other words, both "couples" are statistically different from one another. If one compares to case study 1, this shows that flexible phased deployment and capacity expansion does not add as much value to the system when public acceptance is taken into account. This makes sense because the whole project may be stopped whenever acceptance is unfavorable. The implemented flexibility may or may not result in enough financial return (i.e. cost savings) from the investment. If that is the case, the planner may consider not implementing flexibility in the first place. Life extension, in contrast to other flexibility, is assumed to incur very low costs to implement and thus is always valuable. Besides, the CPLEX time is much shorter as compared to case study 1. This is because the consideration of public acceptance reduces the size of the solution space, and thus decreases the time for searching the optimal solution.

Table 6.13 Results for out-of-sample analysis in case study 2 (£billion).

Alternative	**Mean**	**STD**	**STD error**	**P5**	**P95**	**Value of flexibility**
Rigid	27.64	5.41	0.171	23.14	38.04	-
Flex A	27.67	5.47	0.173	23.14	38.5	0
Flex B	25.89	5.83	0.184	21.39	37.81	1.75
Flex C	25.98	5.88	0.186	21.39	38.04	1.66
Best?	Flex B	Rigid	Rigid	Flex B-C	Flex B	Flex B

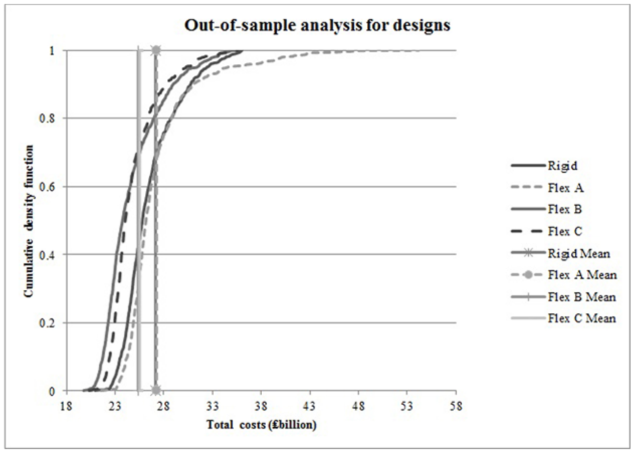


Figure 6.8 Cumulative density functions for the out-of-sample analysis in case study 2.

6.4.4 Sensitivity Analysis

The following sensitivity analysis in this section focuses on the evaluation of the performance affected by variability in the input assumptions. The terms in the Tornado diagrams are listed in Table 6.9 and were discussed in case study 1. As can be seen in Figure 6.9, the expected total costs are mostly affected by the discount rate and penalty costs for losing demand, then by the magnification and expected growth rate. This situation is similar to that of case study 1. It is interesting that the change of magnification did not affect the performance as much as expected. This is because the probability for events higher than or equal to level 4 are considerably small in the three cases (i.e., 0.0122 (low) vs. 0.0066 (base) vs. 0.0039 (high)). Since the lower and upper bounds for the different zones are predetermined by the planner, the change of magnification could only influence the expected performance in a fairly small way.

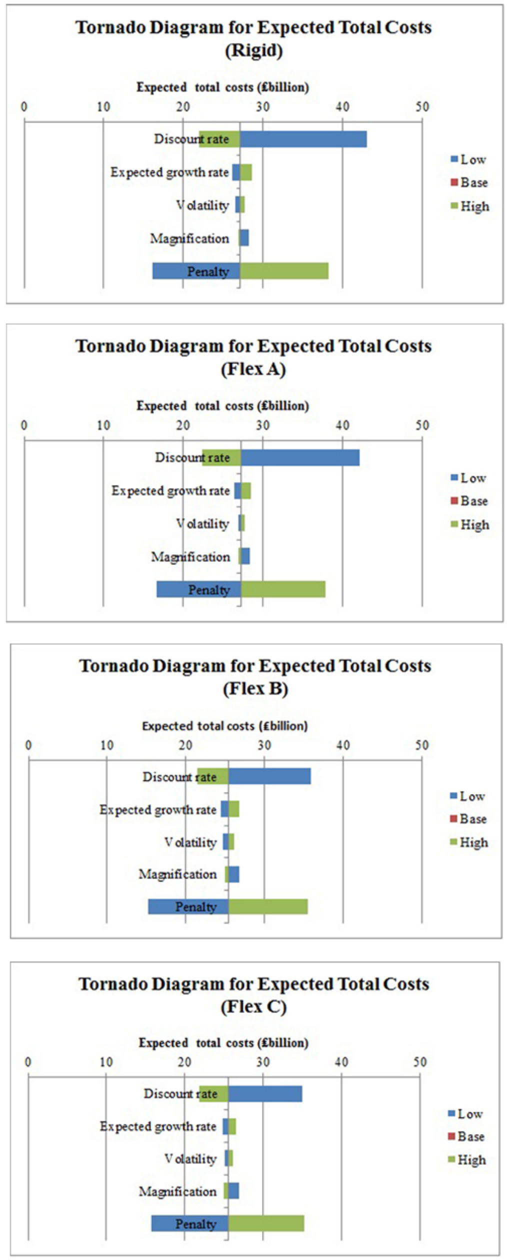


Figure 6.9 Tornado charts for the expected performance of design alternatives in case study 2.

Figure 6.10 illustrates variations in the value of flexibility between the rigid and flexible designs. The discount rate affects the value of flexibility most for the three designs. This makes sense because a smaller discount rate can increase the weight of future cost in the discounted cash flow analysis, and therefore the ability to avoid unnecessary capacity deployment with the flexible systems is worth increasingly more. Once again, the most flexible design B is affected most by the discount rate, as shown by the largest variations in value for flexibility. The rigid design may not necessarily be the worst in terms of expected total costs if the discount rate is large enough (e.g., discount rate > 15%). The penalty costs for the rigid design increase more than those of the flexible designs when the penalty increases, since flexibility can reduce the amount of lost demand while the rigid design can do nothing with it. As magnification decreases, the probability of high level events (level 4 and above) decreases, and flexible designs B and C can reap more rewards by exercising the flexibility strategies, especially life extension. Overall, flexible design B is the best alternative under the current assumptions.

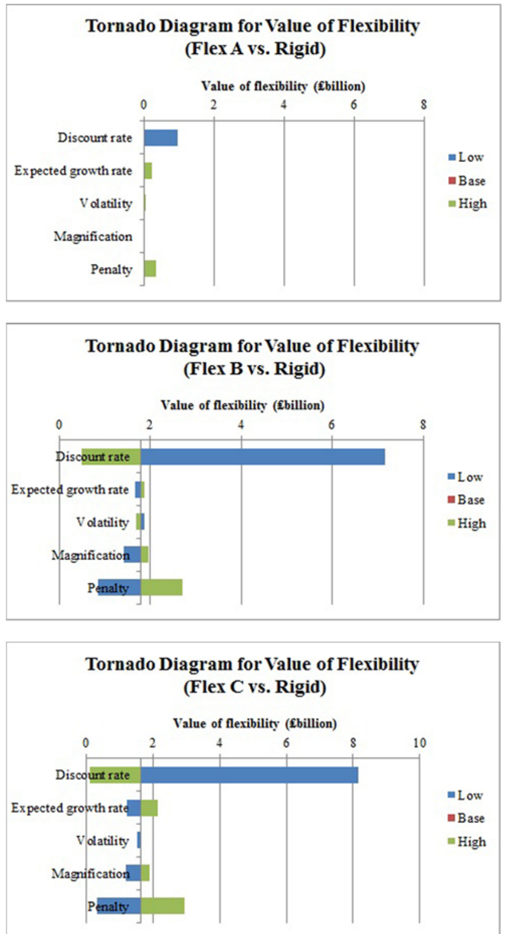


Figure 6.10 Tornado charts for the value of flexibility in case study 2.

In case study 2, the lower bounds for determining the warning and dead zones are significant to the value of flexibility, as discussed above and shown in Figure 6.10. The following sensitivity analysis aims to investigate how significant these lower bounds are. The lower bounds under the original assumption were 5 and 7, respectively. The bounds become 6 and 9 in the mid-level case, and 7 and 11 in the high-level case. The higher the lower bounds, the more public are in favor of nuclear technology. As can be seen in Figure 6.11, the expected total costs decrease when the lower bounds of public acceptance increase. This makes sense because flexibility – such as phased deployment, capacity expansion, and life extension – could be implemented as needed if public environment is more conducive. If the lower bound is large enough, the system does not have to consider the effect of public acceptance as it will never fall into the warning and/or dead zones, which is similar to case study 1 where public acceptance is not a concern. The boundaries of public acceptance require input from the public, and so they may not be completely determined by experts. However, this analysis helps decision makers to determine the final boundaries in consideration of profit and public emotions.

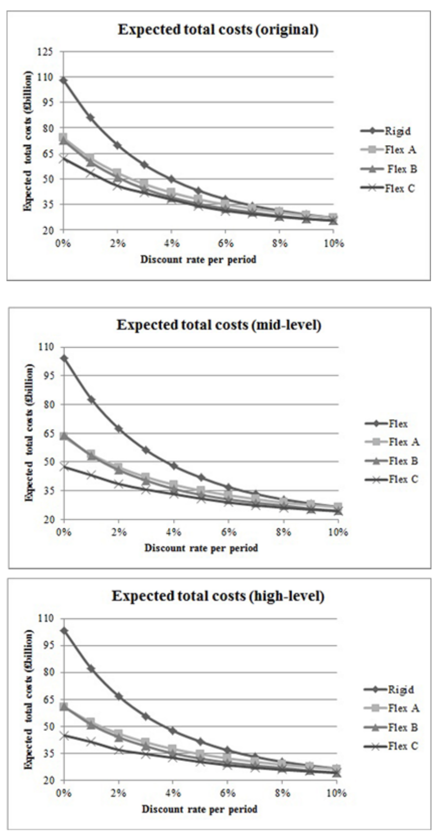


Figure 6.11 The outcomes of the Pareto analysis in case study 2.

Chapter 7 HYBRID HEURISTIC

"The unexamined life is not worth living for a human being." – Socrates (c. 470 – 399 BC)

This chapter explicitly describes the heuristic framework used to solve the two-stage multi-period stochastic models discussed in Chapter 5 and Chapter 6. The purpose is to 1) validate the solutions found to the two-stage multi-period stochastic programming problems using the standard Branch and Bound algorithm in AIMMS, and 2) to determine whether such solutions can be found faster, and with better quality. The framework is based on a two-phase structure, aiming to search for the optimal solution within an acceptable amount of time. The two-phase framework consists of two search algorithms: one regarding the initial configuration and managerial decision rules, and one regarding the capacity deployment over system's life cycle. Section 7.1 introduces the reasons for creating this framework and reviews the current applications of heuristic methods. Section 7.2 describes the inner and outer search algorithms and determines the criteria for stopping the algorithm and avoiding cycling. To validate the framework, this thesis compared it to the default algorithm (which is the standard Branch and Bound) embedded in commercial optimization software (e.g., AIMMS®). The design problem for EMS systems discussed in Chapter 5 is used as the benchmark. The results of this comparison are shown and discussed in Section 7.3.

7.1 Purpose of the Heuristic Algorithm

This thesis proposes a novel design approach for infrastructure systems. This approach incorporates the concept of flexibility and managerial decision rules, and describes the design problem using a two-stage multi-period stochastic model based on the SAA scheme. The model thus consists of a large number of constraints and decision variables, which makes the problem difficult to solve. More specifically, the design problems for EMS systems (also known as location or relocation problems) are usually considered NP-complete problems because the size of the solution space for locating v stations/emergency vehicles in i districts is i^v (Saydam *et al.*, 1994). The complexity of this combinational nature requires various attempts to explore and exploit near-optimal solutions for practical use by meta-heuristic search methods, as the default optimization solvers like CPLEX® and Gurobi® cannot obtain optimal solutions within a reasonable time (Rajagopalan *et al.*, 2007).

Heuristic methods play an important role in solving complex problems. As an example of infrastructure systems, the EMS system has been studied and systematically analyzed using heuristic search methods by many researchers. For example, Rajagopalan *et al.* (2007) considered four meta-heuristic search methods to help identify good solutions for MEXCLP. Their work investigated the evolutionary algorithm (EA), Tabu search (TS), simulation annealing (SA), and hybridizing hill-climbing (HC) through an experimental comparative study of the expected performance of an EMS system. Tabu search was developed by Glover (1990) and is a unique meta-heuristic method that uses a list to store past searched or so called “violated” solutions during the search procedure. Revisiting those unsatisfactory solutions may cause a cycling problem and the search may not be able to get close to the global optima. To get rid of this cycling, “violated” solutions are stored in a list so that the algorithm will not take them as possible alternatives, and the algorithm thus is able to explore and exploit solution space more and deeper in a limited amount of time. For examples of TS applications in EMS systems, see Diaz and Rodriguez (1997), Gendreau *et al.* (1997), Gendreau *et al.* (2001), Rajagopalan *et al.* (2008), and Başar *et al.* (2009). TS accepts a worse solution due to pressure given by the Tabu list, while SA accepts a worse solution with some probability pre-defined by programmers (Kirkpatrick *et al.*, 1983). This ability of SA to “go downhill” helps it escape local optima and thus explore more of the solution space. As a significant meta-heuristic method, SA has been applied in EMS systems in recent decades (Aboueljinane *et al.*, 2013; Syam & Côté, 2010). Evolutionary algorithm is another meta-heuristic method that aims to explore very large search spaces using the concept of evolution (of which genetic algorithm (GA) is an example). There are various examples of how EA can be applied to complex systems (Aytug *et al.*, 2003), including EMS systems (Aytug & Saydam, 2002; Jia *et al.*, 2007b). Unlike the above methods, HC is not a meta-heuristic based search method; rather, it is an iterative improvement method to exploit better solutions within a given search space. The mutation operators in HC help it to explore solution spaces and thus obtain better solutions.

7.2 Algorithm Description

This section describes the framework of the hybrid heuristics. This framework aims to obtain good practical solutions for the design and management of flexible infrastructure systems in a relatively short and acceptable time. Wang (2015) described a similar search heuristic and demonstrated it in a numerical analysis where a two-phase search approach can have better performance than the default algorithm of the solver (i.e., CPLEX) for an expected maximal covering location problem in the EMS sector, in terms of quality of solution and time to best solution. The benchmark problem discussed in Wang (2015) is an older version of the MSCLP proposed in Chapter 5 and an example of designing an infrastructure system for flexibility. This thesis modifies that approach by replacing the inner algorithm while keeping the outer part, and then compares the proposed hybrid heuristic to meta-heuristics such as genetic algorithms and simulated annealing.

The framework consists of inner and outer search algorithms developed for different purposes. The inner search mainly focuses on finding the optimal capacity deployment plan given a group of decision rules and an initial configuration. The outer search, in contrast, focuses on finding the stochastically optimal initial configuration and the best decision rules by changing corresponding parameters. The structure of the framework is shown in Figure 7.1. As can be seen in this flow chart, one may initially determine a basic feasible solution (BFS) for the outer search, and then send this BFS into the inner search to find the corresponding optimal solution for capacity deployment. The solutions of the outer and inner search are combined to calculate the objective value for the design problem. This objective value will be stored in RAM as a benchmark. The procedure to find this objective value continues and terminates due to some stopping criteria (e.g., time and/or iteration limit) or when no more progress can be made in a certain time. The inner and outer search methods are explicitly described in terms of procedure steps in Section 17.2.1 and Section 17.2.2, respectively.

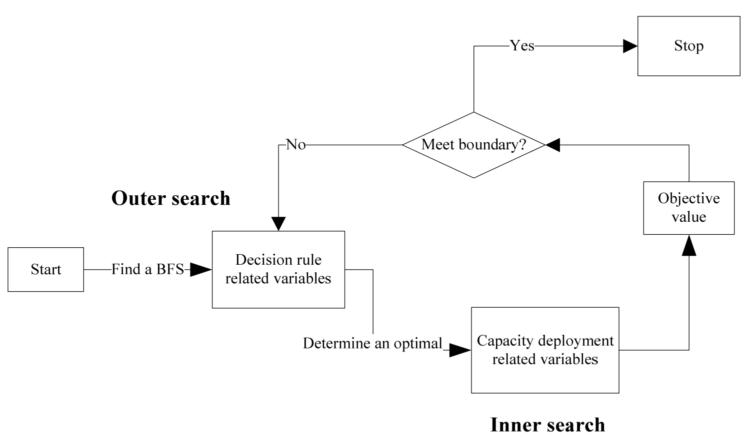


Figure 7.1 Flow chart for the two-phase search framework.

7.2.1 Inner Search Algorithm

The purpose of the inner search heuristic is to optimize the design of an infrastructure system in terms of a specific measurement (e.g., NPV, expected incident coverage rate) based on the given initial configuration and decision rules. It is assumed that the scenarios considered in the modeling framework discussed in Chapter 4 are *iid* (i.e., identically and independently distributed). The original design problem can thus be divided into N sub-problems, and each sub-problem focuses on optimizing the objective function without taking expectation. For example, the objective function of MSCLP can be generalized as $\min \sum_n p_n C_n$, where C_n is the nth total discounted cost under scenario n. The objective function of the nth sub-problem is then $\min C_n$.

The inner algorithm is based on the typical Branch and Boundmethod, which is an exact method for finding the optimal solution. This method was first proposed by Land and Doig (1960) and searches the solution space by enumerating new sub-problems (branching), deleting unsatisfactory sub-problems (pruning), and retaining acceptable sub-problems with an optimal solution (partitioning). It is known that the Branch and Bound method can be time-consuming, depending on the size of the problems and the efficiency of the estimation for the lower and upper bounds of a branch of the search space. In Wang (2015), the inner search heuristic consists of four steps: initialization, evaluation, generation of moves, and recording of results. The heuristic may be faster at finding a good solution than an exact method, but it cannot guarantee the optimality of the solution. The reason that branch and bound was selected as an alternative method for the inner algorithm was because the sub-problem of the MSCLP could be solved to optimality efficiently within an acceptable amount of time (e.g., a few seconds to minutes).

7.2.2 Outer Search Heuristic

The outer search heuristic is dedicated to finding the optimal solution for the initial configuration and decision rules. It is anticipated that the combinations of parameters of decision rules within the possible range is considerably large. Thus, an exhaustive search is time-consuming and may be impractical. An experimental approach known as adaptive One-Factor-at-a-Time (aOFAT) (Frey & Wang, 2006) was applied in the framework to deal with this challenge. Figure 7.2 illustrates the specific process of the aOFAT method and shows it visually by a cube.

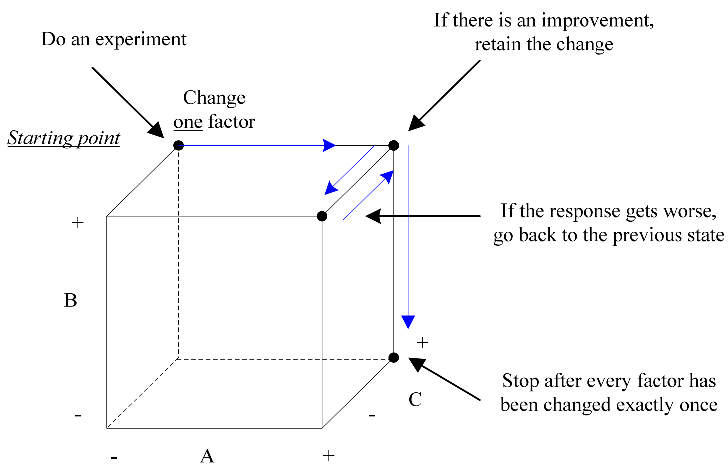


Figure 7.2 Adaptive OFAT as applied to a system with three two-level factors (A, B, and C) (Frey & Wang, 2006).

Unlike typical OFAT, the interactions between different factors are considered in the adaptive OFAT approach. It is assumed that there are n factors, and that there is a response y which is a function of the factors (e.g., the objective function). Furthermore, it is assumed that the factors have two levels each, coded as $x_i \in \{-1, +1\}$. The initial point (or the baseline observation) can be $O_0 = y(\tilde{x}_1, \tilde{x}_{2,} \ldots, \tilde{x}_n)$. After the first step, x_1 is toggled and we have the corresponding observation $O_1 = y(-\tilde{x}_1, \tilde{x}_{2,} \ldots, \tilde{x}_n)$. The final value of x_1 will be determined by taking the difference between O_0 and O_1, and is denoted as x_1^*. Note that none of these interactions are exploited after this step. However, all of the factors except for x_1 are toggled in the subsequent steps of the adaptive OFAT; their final state may be different from the state after the first step. In that case, the contributions due to interactions β_{1j} may potentially be reversed as the process continues. In contrast, the first main effect β_1 is exploited and its contribution is permanent. In the second step, we toggle x_2 to see whether there is improvement in the observation. The interaction β_{12} will not be affected in any way by subsequent experiments. The probability of exploiting the interaction β_{12} will be greater than 50% for all systems with nonzero interactions (Frey & Wang, 2006), which is better than that provided by random chance. If β_{12} is the largest interaction, this probability is no less than 75%. Because the number of experiments taken by aOFAT is only a small fraction to that of OFAT, if the number of factors is large $(n > 5)$, this probability is remarkably high.

Furthermore, Frey and Wang (2006) proved that the probability of exploiting any interaction β_{ij} is no less than the probability of exploiting β_{12}. The aOFAT approach accounts for the interaction at a rate of at least 50%. If a good initial point is selected by experience, the probability can be as good as 75%.

Taking the design problem of EMS systems as an example, there are five factors (i.e., $o^1, o^u, \delta^d, o^o, \delta^o$) and their interactions that need to be considered. Instead of randomly selecting an initial point for o^1, a preprocess based on the demand inputs and incident coverage rate (i.e., fleet size) was implemented to determine a feasible initial point. For example, if the coverage requirement is no less than 90% of the total incident calls and the total hourly arrivals are 10, then the system requires at least nine units of capacity as an initial configuration. These nine units of capacity could be deployed in different combinations at the candidate sites. The outer algorithm will then find the best combination of the initial configuration using the aOFAT approach, as shown in the following diagram. The starting point is randomly selected subject to the capacity requirement. If the system requires nine units of capacity at the beginning of the life cycle, multiple points will be chosen to construct the initial configuration.

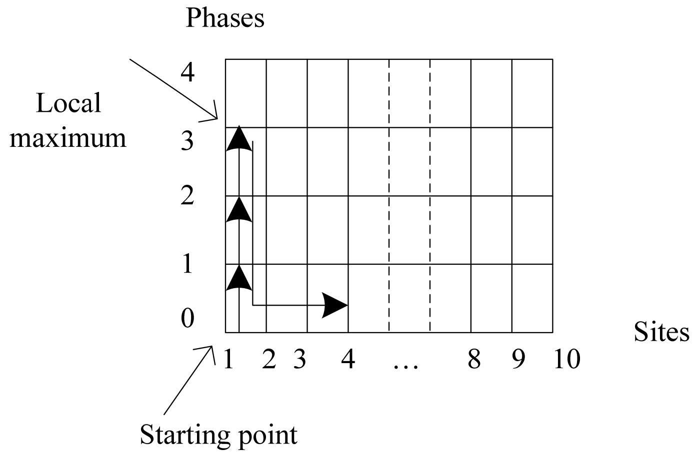


Figure 7.3 Graphic illustration of the search process w.r.t. o^1, the initial configuration (Wang, 2015).

The variables regarding the units to be expanded (i.e., o^u) and installed (i.e., o^o) are explored next. A stepwise increase was used to search the optimal solution for both variables, where the value of the stepwise increase could be independent or dependent of the candidate sites. Note that o^o only affects a subset of sites that were excluded from the initial configuration (i.e., the sites were not deployed capacity at time 0). The method used to search for the optimal δ^d and δ^o was the same as the one used by Wang (2015); the value of δ^d decreases or increases by "a step value which is halved each iteration". Since δ^o is not necessarily an integer, the step value is considerably small (e.g., 0.05) and thus the number of iterations is large relative to δ^d. Any change in the factors can influence the objective value. If this change improves the value, it will be stored as a benchmark. A change will not be accepted if the objective value gets worse, and this factor will not be changed any further. The overall search terminates after every factor has been changed exactly once.

7.3 Numerical Analysis

This section describes the performance of the hybrid heuristic compared to a commercial solver using the branch and bound algorithm (i.e., CPLEX) as well as generic meta-heuristics such as simulated annealing and genetic algorithms. The B&B method is considered the benchmark, which was introduced in Section 5.3, while SA and GA are considered alternative outer heuristics. All outer search heuristics were coded in AIMMS 4.2 and CPLEX 12.6 was selected as the solver for implementing the inner search. All alternative methods were run on the same high performance workstation with 2.60 GHz CPU and 32 GB RAM to make the comparison fair. Two important metrics – Quality of Solution (QoS) and Time-to-Best-Solution (TBS) – were used to measure the performance of the proposed heuristic. These metrics were discussed in Sections 7.3.1 and 7.3.2, respectively. Section 7.3.3 discusses the results for the out-of-sample test of over 1,000 sample scenarios.

7.3.1 Quality of Solution (QoS)

The quality of solution is defined as the relative gap between the solution obtained from the optimization problem and the best LP bound. This bound is determined by the B&B method in the commercial solver (i.e., CPLEX). It is known that the smaller the gap is, the better the quality of the solution. Figure 7.4 shows a comparison of the results of four runs across different incident coverage rates. The form of the solutions is exactly the same as the one shown in Table 5.9 in Chapter 5, which consists of the initial configuration and the parameters for decision rules. As can be seen, the results for all the alternatives and the best LP bound increased alongside the incident coverage rate. This result makes sense because the system requires a greater budget to achieve a higher coverage rate.

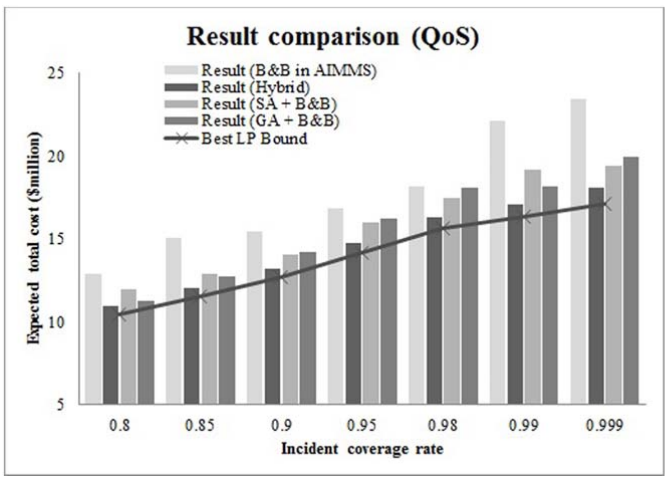


Figure 7.4 Comparison of results in terms of Quality of Solution across different incident coverage rates.

As observed in the figure, the hybrid heuristic was generally the best of the four alternatives in terms of the gap between its solutions and bounds, while the standard B&B algorithm in AIMMS performed the worst in all cases. This may be because the standard B&B algorithm is a general algorithm and is thus not well-suited for this particular design problem. Besides, the results for the other two meta-heuristics were roughly the same when the coverage rate was between 85% and 98%, and both results were a bit worse than those for the hybrid heuristic. This may be because the mechanism of SA and GA is based on random search, and there was no such initial treatment as there was for o^1 in the hybrid method. The quality of solution thus heavily depends on the number of iterations defined initially; in general, the larger this number is, the better quality of solution will be obtained from SA and GA.

7.3.2 Time to Best Solution (TBS)

The time to best solution (TBS) is the exact time spent solving the optimization problem subject to iterations or gap limits which, as indicated in Section 5.3, were four million iterations and a 1% relative gap, respectively. The default upper bound of the solving time was 100,000 seconds (i.e., approximately 1 day and 4 hours), which may not be an acceptable time in practice. Figure 7.5 shows a comparison of the TBS results for the four alternative methods. Overall, it indicates that the hybrid method was generally faster than the other two meta-heuristics and the default algorithm. The solving time for the hybrid method was never more than 3,000 seconds, while only the SA method with a 80% coverage rate was close to this value (i.e., 2,980.47 seconds). The TBS for the meta-heuristics was also less than that for the default method when the coverage rate did not exceed 98%, while SA cost less than GA when the incident coverage rate was relatively small (i.e., less than 95%).

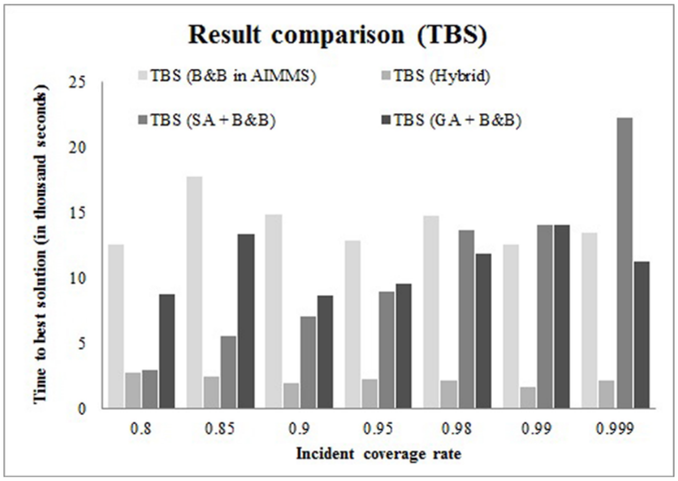


Figure 7.5 Result comparison in terms of Time to Best solution across different incident coverage rates.

It is interesting that only the TBS for SA increased alongside the incident coverage rate. This makes sense because SA searches the neighborhood solution in every iteration, while the solving time for each solution increased with the coverage rate. For GA, TBS did not consistently increase due to the randomness of making crossovers between the generated solutions and mutating new solutions. Although the TBS of the hybrid method was not dependent on the coverage rate and thus intractable, its value was relatively stable, fairly small (no more than one hour), and applicable in practice.

7.3.3 Out-of-Sample Analysis

An out-of-sample analysis was conducted to ensure that the results obtained from these heuristics, as compared to those obtained from the commercial solver, were reliable. This test is normally used to evaluate the performance of results in untested sample scenarios and it was applied in the engineering two application studies on EMS and nuclear power systems in Chapter 5 and Chapter 6, respectively. Figure 7.6 shows the results for the four alternative methods across different incident coverage rates. As can be seen, the optimization solution obtained from the hybrid heuristic outperformed the other methods overall. The solution obtained from the B&B method in AIMMS also performed better than the meta-heuristics when the coverage rate was higher than or equal to 90%. Results for SA and GA followed the same order as in the analysis of QoS in Section 7.3.1.

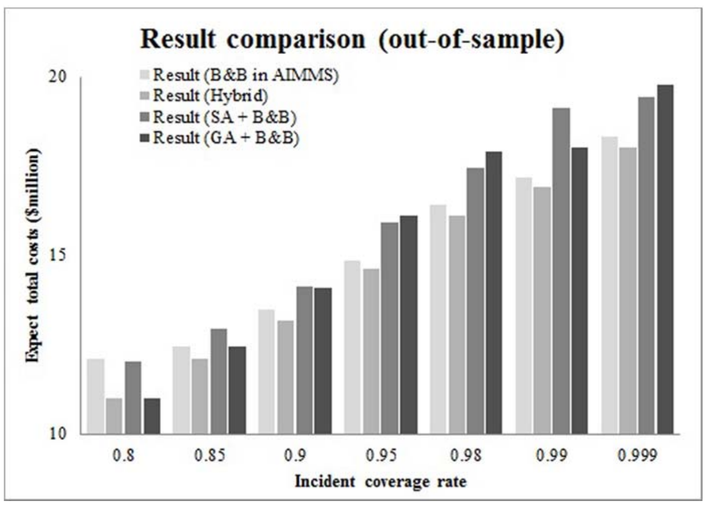


Figure 7.6 Results comparison in terms of the out-of-sample analysis across different incident coverage rates.

Although the differences between the expected values of the hybrid and the default method (i.e., the standard Branch and Bound algorithm in AIMMS) generally became smaller when the coverage rate increased, the results for the hybrid method in the out-of-sample analysis are closer to the solutions obtained from the optimization problem than those from the default method.

This indicates that the hybrid method is more reliable and its solutions are more robust in the face of untested uncertainty scenarios. This makes sense because the hybrid method can find optimal solutions for the sub-problems in a fairly easy way, while the default method can only find approximations of optimal solutions.

Table 7.1 Summary of results of the out-of-sample analysis ($CoV = 0.9$) ($million).

Alternative	Mean	STD	STD error	P5	P95	Median
Branch and Bound	13.48	1.07	0.034	11.86	15.31	13.39
Hybrid	13.19	0.91	0.029	11.80	14.73	13.15
SA	14.13	1.00	0.032	12.44	15.84	14.10
GA	14.08	0.98	0.031	12.56	15.80	13.58
Best?	Hybrid	Hybrid	Hybrid	Hybrid	Hybrid	Hybrid

Table 7.1 summarizes the results of the out-of-sample analysis for all methods when the coverage rate was 90%. The smallest value in columns 2 to 5 indicates the best solution. As can be seen, the hybrid method outperformed the other three methods in all significant parameters. The difference between the mean values for the Hybrid and Branch & Bound algorithms is statistically significant in terms of its p-value ($0.002983 < 0.05$). Although the results of the standard B&B method were the second best overall, its standard deviation was the largest. This may be because the default method did not solve the sub-problems to optimality, as compared to the other heuristics. Again, the

results show that the hybrid method yielded more reliable solutions than other methods.

Overall, the analysis of this chapter indicates that the proposed hybrid heuristic performs the best across all four alternatives. The standard B&B algorithm could solve the design problem, but the quality of the solution heavily depends on the upper bound of the solving time as well as the requirement to the relative gap. However, the standard B&B method may not be efficient and effective when the scale of the problem is fairly large (e.g., a design problem with more than 15 districts). From an engineering standpoint, it may be necessary to have a heuristic to solve the design problem in an acceptable amount of time, while the optimal solution is approximately close to the global optima (i.e., relative gap is acceptably small). The proposed hybrid heuristic is demonstrated to be feasible in finding acceptable and reasonable solutions, and to be effective in solving the design problem faster than the standard B&B algorithm. The form of the solution obtained by applying the hybrid heuristic is the same as the one obtained by the standard B&B algorithm. The quality of the solution is also better than that of the standard B&B method when the incident coverage rate is required to be higher than or equal to 95%. On the one hand, the hybrid method may not be better than the default B&B method in a small scale problem because the hybrid algorithm is based on heuristic, which could not guarantee to find the global optima. The solving time for a small scale problem by using the default B&B

method is acceptable in both experiment and practice. This hybrid method, on the other hand, may be helpful in solving large scale problem and thus make the proposed mathematical framework be practical. Moreover, the proposed hybrid algorithm is a globally convergent heuristic. The inner algorithm based on B&B is arguably considered as globally convergent, which has been proved many times. Although the outer algorithm based on aOFAT requires a "good" BFS to find the optimal solution efficiently, it eventually converges for an arbitrary BFS as the algorithm stops by the same predetermined requirement. In other words, the proposed hybrid algorithm is also globally convergent. It is possible that this hybrid algorithm may not consider interactions and thus may miss the optimal solution. If the computational time is sufficient, the factorial approach could be a good alternative as it explicitly considers interactions. Adaptive OFAT is used in this thesis because computational time is always an issue in a practical context. The proposed algorithm could offer a good enough solution in a fairly small amount of time.

Chapter 8 FINDINGS AND DISCUSSION

"One must sustain one's effort when a task is nearing completion." – Liu Xiang (c. 77 – 6 BC, Strategies of the Warring States)

This chapter discusses the findings obtained from the numerical analyses in Chapter 5 to Chapter 7. The discussion here is not just a simple repetition of that in the above chapters. Instead, Sections 8.1 to 8.3 focus on answering the research questions introduced in Chapter 3 and determining whether they were well addressed, as well as discussing the potential of the results to impact architecture and decision-making at the highest level of policy-making for more general infrastructure systems. The discussion in this section also offers recommendations for system planners, managers, and academic researchers on how to understand and apply the approach proposed in this thesis to design flexible infrastructure systems in practice. Section 8.4 explicitly describes the potential of the proposed heuristic framework for dealing with more general, larger-scale problems in practice.

8.1 Area 1: Design of Flexible Infrastructure Systems

The research questions proposed in Section 3.1 are combined, reorganized, and restated as follows:

"What is the best design for an infrastructure system to make it adaptable to the changing environment, and to improve the anticipated performance over

a system's life cycle under uncertainty? How can one design a flexible system that can be easily applied to the real world, and evaluate the corresponding value of flexibility as compared to a benchmark design (i.e., typical design without flexibility)?"

The results in Chapter 5 and Chapter 6 indicate that the proposed novel design approach incorporating the concept of strategic-level flexibility may produce better design alternatives for infrastructure systems, as compared to typical robust designs (i.e. stochastically optimal although rigid capacity designs). The results show that flexible designs outperformed robust designs in both case studies in terms of expected performance. The two case studies also demonstrated that the flexible design outperformed the competitors in terms of significant statistical indicators (e.g., STD, P5, P95), representing its ability to adapt to various realizations of uncertainty scenarios, protecting from downside conditions, and enabling a system to capitalize on upside opportunities. Flexible designs take advantage of change rather than resisting it, and thus have a better performance in different environments, according to the results of the sensitivity analyses in the above chapters.

The methodology discussed in Chapter 4 can help system designers and planners in thinking about how to develop a flexible design that can be used in practice. Typical DP-based ROA approaches require the decision-maker to determine the current stage and state in a decision tree based on the past and existing state of the main uncertainty drivers, predict the states of these drivers in the future, and then perform backward induction to determine the optimal policy at any given time. Thorough knowledge and understanding of dynamic programming is required to understand and carefully apply these operations to find, and also exercise the solution in operations. Further challenges include using these approaches when multiple uncertainty drivers are considered and determining the stochastically optimal initial configurations and the best decision rules for implementation within an acceptable amount of time.

The proposed approach based on decision rules can handle these issues, as demonstrated in the case studies. Compared to the solution obtained via typical ROA approaches, the output vector of this approach is easy to understand and thus to implement, even without advanced knowledge of mathematical modeling or simulation. The recommended solution consists of an initial configuration and a set of decision rules. System planners can deploy capacity at the beginning of the life cycle, as suggested, and then manage the system by following the decision rules based on the realization of uncertainty scenarios - they do not necessarily require advanced training in the mathematical model or the mechanisms of behind the approach. All that the system planners have to do is collect past and current information (e.g., demand) to generate samples of uncertainty scenarios to make the solution more reliable. The planners are then able to design an infrastructure system incorporating favorable flexibility strategies by modifying the design variables and constraints accordingly. The decision rules are intuitive to use and do not require high-level knowledge of mathematical programming and/or simulation. In addition, the modeling framework discussed in Chapter 4 contains a systematic evaluation method for valuing the flexibility strategies. This method was applied and verified in both case studies, and is feasible for practical use.

8.2 Area 2: Design and Management of EMS Systems

"Can one develop an EMS system incorporating flexibility so that it can stochastically dominate existing rigid designs (i.e., the benchmark) in terms of KPIs, such as total costs and/or incident coverage rate? If so, is it always worth developing a flexible EMS system to gain a better anticipated performance under uncertainty in the long term?"

The results presented in Chapter 5 answer this research question in a numerical way. A novel design was proposed based on the new design approach by incorporating strategic-level flexibility and implemented via decision rules based on the realization of uncertainty scenarios. The objective in the mathematical model was to minimize the expected total cost. This objective makes sense because budget is always a concern for the issue of capacity deployment. Besides cost, the incident coverage rate (also referred to as fleet size) is another concern for system planners. One may determine the objective and corresponding constraints, and then develop a mathematical model following the framework for different purposes.

The benchmark design is a rigid one that deploys all capacity at once at the beginning of a system's life cycle. This design is usually applied for short-term planning, as in the literature, and is not always the best solution for long term planning. The reason that this design is considered here is for the purpose of comparison. One may want to know how good the flexible design is compared to a typical robust design. In order to make the analysis more reliable, a long term robust design is considered as an alternative. This robust design is an extended, upgraded version of a short term design proposed by Beraldi and Bruni (2009). This long term design allows for capacity deployment over time and space based on a fixed plan. It does not have the ability to adapt itself to various conditions, and thus cannot benefit from upside opportunities. However, it is hard to say if the flexible design dominates stochastically the typical designs. Referring to the results in Section 5.3, the flexible design had a better overall performance when the incident coverage rate was required to be higher than 90%. The flexible design may be favorable to system planners when the requirement for the system fleet size is considerably high.

In addition, improvement to the objective of interest (i.e., value of flexibility) is not fixed and may vary significantly depending on the setting of the input parameters. The most influential parameter for such design problems is the discount rate. According to the results obtained in the above case studies, the higher the discount rate, the less valuable flexibility is in the face of uncertainty. This makes sense because higher discount rates provide more incentives to defer decision-making regarding capacity deployment to later periods, and vice versa. If the discount rate is too small, the flexible design may not be valuable due to the cost premium for enabling flexibility. Although this parameter is usually a fixed number in analytical models, there is a possibility that its value may change over a system's life cycle in the real world. This unexpected change could affect the ultimate expected performance and make the recommended design more or less valuable. One possible way to deal with this issue is to consider the discount rate as an uncertainty driver in the mathematical model. The recommended solution obtained by solving the problem to optimality is then more adaptable to various scenarios, and the value of flexibility is thus more robust.

Another parameter that significantly influences the expected performance is the mean growth rate of long term demand. More specifically, the faster the demand grows, the more valuable flexibility is. This is true because typical robust designs need to deploy more capacity in early periods in order to meet growing demand in later periods, and thus cost more. The flexible design, by contrast, has the ability to deploy capacity later only if it is needed, and thus cost less than typical designs. It should be noted that flexibility may not be favorable if the mean growth rate is relatively low or even zero (i.e., the demand is constant over time). This is because the proposed flexibility strategies, such as phased deployment and capacity expansion, are implemented after losing appropriate coverage of incidents (i.e. demand). When the mean growth rate is too small, the flexible design cannot benefit enough from its adaptable strategies to offset the cost of enabling flexibility and/or the penalty for losing demand. System planners must be careful when considering the flexible design when the demand is relatively stable in the long term. Also, the volatility of long term demand could affect the value of flexibility in a similar way. The flexible design could no longer be the best choice when the long term demand of the system is flat.

Besides the general parameters discussed in the above paragraphs, one may care about special parameters like the coverage radius used in the model. Such special parameters have the potential to make the original design solution infeasible if they are not consistent with the values considered in the mathematical model. For example, the radius of a station is unlikely to be a constant over a system's life cycle due to population growth and changing traffic conditions in the long term. System planners may need to solve the problem more than once based on the new inputs of those parameters in order to find a more suitable solution.

This thesis introduced a novel approach for the design and management of infrastructure systems in the long term. This approach explicitly considers long term uncertainty drivers during the design process and deals pro-actively with such drivers by incorporating the concept of strategic-level flexibility. The approach is demonstrated by case studies in two engineering contexts where it improved significantly the long term expected performance over typical robust design approaches. It was also demonstrated that this approach has the potential to be applied to the issue of capacity deployment for different infrastructure systems over time and space by changing some of the design variables and constraints. Again, the proposed modeling framework is capable of testing different decision rules by modifying the corresponding non-anticipative constraints accordingly.

8.3 Area 3: Design and Management of Nuclear Systems

The research questions for the design and management of nuclear systems are similar to those for EMS systems, and are restated as follows:

"Can one site nuclear power plants flexibly so that these energy systems can have good anticipated performance under uncertainty? What flexibility strategies can one consider in the design and which one most benefits the system? What is the influence of social acceptance on the expected performance of a nuclear system, as well as the decision-making processes?"

The results presented in Chapter 6 indicate that the modeling framework can be used for developing more general infrastructure systems. The flexible design considers three flexibility strategies – phased deployment, capacity expansion, and life extension – to deal pro-actively with uncertainty drivers such as demand for electricity and social acceptance. These strategies were also implemented via decision rules, as was done for EMS systems. The objective was to minimize the expected total cost over the system's life cycle. This objective can be replaced by other monetary objectives, like the levelized cost of electricity, if needed.

The benchmark design for siting nuclear power plants deploys all capacity at once at period 0, similar to the benchmark design for EMS systems. This benchmark represents the planning of such capacity deployment in the real world, as described in the relevant literature. In order to analyze the importance of flexibility strategies, the flexible design was reorganized as two sub-flexible designs where each partially considered a real option strategy. The results indicate that the most flexible design – that which considered all flexibility strategies – outperformed the other three designs, no matter if social acceptance was considered or not. The life extension strategy was more valuable than phased deployment or capacity expansion because there is no cost premium for enabling it (in reality this cost would be very negligible, as compared to the capital costs involved). All of the flexible designs stochastically dominated the benchmark when the discount rate was less than 10%, indicating that flexibility is favorable to this nuclear system in most cases.

It is, however, not always worth investing in the flexible design. The value of flexibility is not obvious unless the discount rate is sufficiently small (i.e., less than or equal to 6%). If the cost premium for enabling flexibility strategies is not smaller than the value of flexibility, there may not be an incentive to site the power plants in a way that exploits flexibility. Thus, system planners need to be careful about the discount rate at the time when they initiate the whole project. To them, flexibly siting nuclear power plants may be of lesser interest than it would be to China or Russia, because nuclear systems in the these countries are mainly operated by the government and thus may have quite small discount rates (e.g., 3-4%).

In addition, the results show that consideration of social acceptance indeed increased the expected total cost. It also decreased the value of flexibility because: 1) flexibility strategies cannot be implemented if social acceptance is unfavorable, and 2) the system may be shut down earlier than expected. For countries that care highly about acceptance, the flexible design may not be the best choice. On the contrary, the novel design approach may be favorable to countries that require faster development and are thus less concerned about social acceptance.

8.4 Proposed Hybrid Heuristic

The hybrid heuristic proposed in this thesis helps find the optimal solution faster than simply using a commercial solver (e.g. standard Branch and Bound algorithm in AIMMS), and it thus makes the design approach more practical. The hybrid method consists of inner and outer search algorithms, iteratively searching for the optimal solutions for sub-problems by a given vector of the initial configuration as well as decision rule parameters. The inner search uses the Branch and Bound algorithm, while the outer search is based on adaptive OFAT. In addition to the default algorithm embedded in the software (e.g., standard Branch and Bound), two other methods using meta-heuristics were applied for comparison. These methods were developed based upon the same two-phase framework as the hybrid method, while the outer search methods were based on SA and/or GA.

The results show that the hybrid method was both efficient and effective in terms of quality of solution and time to best solution. In general, the hybrid method was about five times faster than the default method, while the solving time of the hybrid method was half and one quarter of that of SA and GA, respectively (see Figure 7.5). The form of the solutions obtained by three heuristic methods (the hybrid heuristic plus SA and GA) is the same as the one obtained by using the default Branch and Bound algorithm. On the other hand, the relative gap of the solutions obtained by the hybrid method was fairly small compared to the three other methods, and much smaller than the gap of the default method as well. Moreover, the solution of the hybrid method was the most reliable one in terms of the significant statistical parameters (e.g., mean, standard deviation, etc.). It was demonstrated by the numerical analysis that the hybrid method could make the proposed modeling framework more practical in the real world. It should be noted that the default B&B method is able to solve a problem with relatively small scale (e.g., number of districts is no more than 10) and the time spent for the solving procedure is acceptable.

This B&B algorithm is theoretically guaranteed to find the global optima if the solving time is no longer a restriction. It is therefore best suited for a small scale problem (i.e., less districts), or a problem with less interactional behaviors (e.g., the design problem in Chapter 6). The siting problem considered in Chapter 6 can be well solved by the default method (i.e., gap = 0) in a very small amount of time. When the problem becomes big (e.g., a real large scale design problem), or the interactions becomes complicated between facilities (e.g., emergency stations), the proposed hybrid method is practically useful for finding good or approximate solutions in an acceptable time. The quality of the solution, however, could not be guaranteed because the method is based on heuristic. The quality of the solution could heavily depend on the initial treatment, number of decision variable and constraints, as well as type of the model (e.g., linear or nonlinear).

The efficiency and effectiveness of the hybrid method was affected by the initial treatment for determining the starting point (i.e., a basic feasible solution). A good starting point close to the true optimal solution may help accelerate the entire search process. Currently, this treatment is based on data analysis of historical incident arrival rates at the beginning of the life cycle. The combination of initial points could still be too high if the number of candidate sites is fairly large. In that case, considering a priority mechanism could be an effective way to find better starting points among the tens of thousands of potential sites. That is, a candidate site that could cover more incident calls would have a higher priority to deploy capacity.

Furthermore, meta-heuristics like SA and GA did not perform well in the numerical analysis. Even though the results obtained by solving the optimization problem for SA and GA were better than those obtained by the default method, these two heuristics were not consistently good in terms of TBS and also performed poorly in the out-of-sample analysis. This makes sense because these meta-heuristics are generic methods that need to be specifically developed and optimized. In light of the results for QoS, the methods based on these meta-heuristics must be further improved. One possible way to improve them is by using the same idea of initial treatment to find a good starting point.

Chapter 9 CONCLUSION

"Reach the same goal by different routes." – Classic of Changes (c. 1100)

"Tact is the knack of making a point without making an enemy." – Sir Isaac Newton (1643 – 1727)

This thesis demonstrated that the proposed modeling framework, incorporating the concept of strategic-level flexibility, can be a novel design approach for infrastructure systems with the consideration of uncertainty drivers over the system's life cycle. The thesis answered research questions regarding the methodology and application domains with two engineering case studies. Results of the two case studies demonstrated that this framework was helpful in modeling new flexible designs, and systematically valuing the corresponding flexibility strategies. Results also demonstrated that proposed flexible designs significantly improved anticipated long term system performance over its life cycle in terms of key performance indicators (KPIs). Moreover, comparison results for alternative solving methods demonstrated that the proposed hybrid heuristic based on a two-phase framework significantly improved solutions in term of the quality and the search time, as compared to the default Branch and Bound algorithm as well as meta-heuristics. It also demonstrated that the proposed modeling framework was suitable for use in practice.

An important finding is that the concept of strategic-level flexibility can significantly improve anticipated long term performance for an infrastructure system by dealing pro-actively with uncertainty, as it does for typical (site-specific) engineering systems (e.g. oil platforms, real estate development project, etc.) An infrastructure system is of course complex, and involves massive interactions between elements (e.g., emergency stations and vehicles, power plants) during the system operations. Uncertainty drivers such as long term demand and traffic/power transmission conditions significantly affect the system performance over its life cycle. Strategic-level flexibility like phased deployment and capacity expansion were demonstrated to be useful for dealing with those uncertainty drivers, by deploying capacity in phases over time and space, instead of deploying them all at once. The discounted expected system performance thus benefits from such strategies because they could save money in the early stages but still maintain a high-level system performance overall as expected.

The modeling framework represents an important tool for the design and evaluation of strategic-level flexibility in the urban context. The concept of flexibility is captured by an analytical model – a mixed integer one – in the form of stochastic programming, and the embedded flexibility strategies are analyzed via managerial decision rules. Decision rules are referred to as "IF-THEN-ELSE" statements and thus are intuitive to understand and exercise. The understandability and usability may be favorable for system planners who lack the required knowledge in DP-based ROA methods. Besides, the systematic four-step evaluation procedure was demonstrated to be helpful and applicable in valuing the flexibility under different circumstances for making the final decisions. Note that flexibility may not always be better than typical robust designs, especially when uncertainty drivers are less fluctuating in the long term. The sensitivity analysis within the evaluation procedure provides an opportunity to know better on how the system is influenced by input parameters and the order of influence for them.

The hybrid heuristic was demonstrated to be the most effective and efficient method across four alternative methods by a fair and systematic comparison. The solution gap for the hybrid method is the smallest in general, and the solving time is also the shortest. This showed that the proposed heuristic has the potential to make the modeling framework practical in use, as its solving time is about hours not days or weeks. Compared to the default method as well as meta-heuristics, the solutions obtained from the hybrid method are also more reliable in terms of statistical significance.

9.1 Limitations of Current Approach

There are several shortcomings in this thesis that may fuel opportunities for future work. First of all, the modeling framework is a stochastic one based on a scheme of sample average approximation. It thus inherits weaknesses from the sample average approximation method, i.e., the observed performance differs when out-of-sample scenarios are used, even though the flexible solution is shown to handle such cases much better than for the rigid alternative solutions. The optimal solution obtained by the stochastic model may even be infeasible in some extreme out-of-sample cases if they are not considered in the optimization problem.

Secondly, the flexible design may not always be the best compared to the rigid designs in some realizations of uncertainty scenarios due to the assumption of the managerial decision rules. The definition of decision rules indicates that decision rules are only implemented when the corresponding requirement is satisfied. For example, an emergency station will expand a given number units of capacity if this station missed a certain number of emergency calls in the last strategic period. The implementation of decision rules only depends on the realization of the uncertainty drivers (e.g., incident/electricity demand, social acceptance), and is independent from prior relevant decision-making processes. Deploying more capacity at current periods due to the satisfaction of decision rules may not reap enough benefits compared to investment when long term demand is relatively flat with few consecutive peak points.

The computational time for the optimization problem is always a concern for resource allocation or capacity deployment problems like the EMS systems and energy systems. The size of the computational problem increases significantly in terms of variables and constraints when more candidate sites and scenarios are considered (e.g., $|J| \geq 20$ and $|N| \geq 15$). The model then becomes very difficult to solve by the default exact solution method within an acceptable and practical time. Also, the assumptions of the method for generating uncertain inputs (e.g., INES) in the case studies may be unrealistic. There is however no unique and ultimately accepted approach for generating such data at present in simulations. The outer algorithm based on adaptive OFAT also has a drawback that the interactions may not be considered totally (the chance is still there even it may be fairly small).

9.2 Extending the Current Approach to General Urban Systems

One important contribution of this thesis is the modeling framework. It is demonstrated that this framework can be a novel design approach for determining the optimal siting policy with respect to a capacity deployment problem in the urban context. To apply this approach to a general infrastructure system, one may consider designing the system following the given procedure by the framework.

First, it is always good to investigate past typical designs without consideration of flexibility and find out the best one under the deterministic or stochastic condition. This design could be used as a benchmark for us to evaluate the value of flexibility. Second, system planners need to identify the potential uncertainty drivers. This is important because which flexibility strategy is going to be used in the design will depend on what uncertainty the system is facing. Once uncertainty drivers are determined, approaches like DSM and prompting could be applied for determining the strategic-level flexibility considered in the design.

To evaluate the theoretical value of flexibility, one may need to modify the stochastic model accordingly, based on a specific urban system, including design variables and constraints. Sensitivity analysis can be implemented after obtaining the optimal solution. The most influential parameters will then be determined in orders. If the scale of the design problem is fairly large and the solution looks bad in terms of QoS and TBS, the two-phase framework could be considered as an alternative to develop search heuristic specifically. It is recommended that the inner search method is better based on Branch and Bound method as it can guarantee optimality of the solution. For the outer search, there is no criterion to determine which method is the best. One may test different alternatives to decide which one is the most appropriate for the design problem. In terms of result comparisons in this thesis, adaptive OFAT may be the best at present, but other local search algorithms could be explored.

9.3 Future Research Opportunities

This thesis generates many opportunities for future research in the field of design and management of infrastructure systems. Using this modeling framework enables evaluation and comparison of other flexibility strategies or objectives for general infrastructure systems. An interesting study in siting nuclear power plants could consider switching between different nuclear reactors as a main flexibility strategy to deal with uncertainty of social acceptance. It is assumed that a more advanced but expensive technology can be safer than existing ones in the face of unexpected safety issues, and thus it can keep operation even social acceptance is fairly low. There is clearly a trade-off between using new technology and the old one, depending on the realization of the uncertainty scenarios. The objective of such study could be to minimize the levelized cost of electricity, which is an economic metric for evaluating the performance of an energy system.

It should be noted that mixed integer stochastic programming is not the only modeling method to capture the concept of strategic-level flexibility from a mathematical perspective. Another research avenue could be to use other methods to develop models based on the framework, for example, using constraint programming. It is observed that the procedure for establishing non-anticipative constraints regarding decision rules can be challenging if the logical relationship is complicated. Constraints programming is an alternative programming paradigm supported by various solvers. Using it with logical statements can be easily done in a just a few lines of code. This method however cannot guarantee the optimality of the solution. It would be feasible to apply it when the original problem is a small scaled one, or separating the original problem into multiple sub-problems, then solving them using exact search methods (like the Branch and Bound). Also, it is possible to consider costs as an uncertainty driver in the modeling process. For many reasons such as technological improvement and/or supply shortage, costs may decrease or increase accordingly over time. Considering costs as a random variable could make the model more realistic. One possible way to capture this random variable is to use the so-called "learning effect". This concept could be used to represent the economic relationship between costs and experience. Typically, costs decrease along with an increase in experience doing a particular thing. The relationship could be described by several main functions, such as the exponential growth or the power law. This learning effect can be applied to the installation of new stations/power plants and capacity expansion.

One more research opportunity is to develop a decision support system for training engineers using the solution obtained from the approach proposed in this thesis. It could be the case that the solution performs very well in the mathematical model and out-of-sample analysis, but performs badly in reality when used by system operators. Regardless of the assumptions made in the model, the timing for implementing the flexibility strategies significantly affects the overall performance. Serious gaming and simulation games can then create environments for system planners and/or engineers to design and operate the system under uncertainty by given a set of best decision rules (Cardin *et al.*, 2014; Cardin *et al.*, 2015b; Ligtvoet & Herder, 2012). This platform provides engineers an opportunity to familiarize themselves with the flexibility strategies and decision rules. They can operate the system in the platform based on either their past experience and/or recommended strategies, and compare results of system performances. The platform also provides a chance to realize how uncertainty could affect the system performance, and thus the potential of strategic-level flexibility to the design and management of infrastructure systems.

Bibliography

Abdelhamid, M. B., Aloui, C., & Chaton, C. (2009). A Real Options Approach to Investing in the First Nuclear Power Plant under Cost Uncertainty: Comparison with Natural Gas Power Plant for the Tunisian Case. *International Journal of Oil, Gas, and Coal Technology, 2*(1), 44-57.

Aboueljinane, L., Sahin, E., & Jemai, Z. (2013). A Review on Simulation Models Applied to Emergency Medical Service Operations. *Computers & Industrial Engineering, 66*(4), 734-750.

Abrecht, P., Arungu-Olende, S., Francis, J. M., de Gaspar, D., Nashed, W., Nwosu, B. C. E., Rose, D. J., & Shinn, R. L. (1977). *Public Acceptance of Nuclear Power - Some Ethical Issues*. Paper presented at the International Conference on Nuclear Power and Its Fuel Cycle, Saizburg, Austria.

Abudeif, A. M., Abdel Moneim, A. A., & Farrag, A. F. (2015). Multicriteria Decision Analysis based on Analytic Hierarchy Process in GIS Environment for Siting Nuclear Power Plant in Egypt. *Annals of Nuclear Energy, 75*, 682-692.

AIMMS. (2014) (Version 4.2). The Netherlands: Paragon Decision Technology B.V.

Ajak, A. D., & Topal, E. (2015). Real Option in Action: An Example of Flexible Decision Making at A Mine Operational Level. *Resources Policy, 45*, 109-120.

Aliyu, A. S., Evangeliou, N., Mousseau, T. A., Wu, J., & Ramli, A. T. (2015). An Overview of Current Knowledge Concerning the Health and Environmental Consequences of the Fukushima Daiichi Nuclear Power Plant (FDNPP) Accident. *Environment International, 85*, 213-228.

Alsalloum, O. I., & Rand, G. K. (2006). Extensions to Emergency Vehicle Location Models. *Computers & Operations Research, 33*(9), 2725-2743.

Aly, A. A., & White, J. A. (1978). Probabilistic Formulation of the Emergency Service Location Problem. *Journal of the Operational Research*

Society, 29(12), 1167-1179.

Araz, C., Selim, H., & Ozkarahan, I. (2007). A Fuzzy Multi-objective Covering-based Vehicle Location Model for Emergency Services. *Computers & Operations Research, 34*(3), 705-726.

Aytug, H., Khouja, M., & Vergara, F. E. (2003). Use of Genetic Algorithms to Solve Production and Operations Management Problems: A Review. *International Journal of Production Research, 41*(17), 3955-4009.

Aytug, H., & Saydam, C. (2002). Solving large-scale maximum expected covering location problems by genetic algorithms: A comparative study. *European Journal of Operational Research, 141*(3), 480-494.

Babajide, A., de Neufville, R., & Cardin, M.-A. (2009). Integrated Method for Designing Valuable Flexibility in Oil Development Projects. *SPE Projects, Facilities, and Construction, 4*, 3-12.

Badri, M. A., Mortagy, A. K., & Alsayed, C. A. (1998). A Multi-objective Model for Locating Fire Stations. *European Journal of Operational Research, 110*(2), 243-260.

Ball, M. O., & Lin, F. L. (1993). A Reliability Model Applied to Emergency Service Vehicle Location. *Operations Research, 41*(1), 18-36.

Baron, O., Berman, O., Kim, S., & Krass, D. (2009). Ensuring Feasibility in Location Problems with Stochastic Demands and Congestion. *IIE Transactions, 41*(5), 467-481.

Başar, A., Çatay, B., & Ünlüyurt, T. (2009). A backup double covering model and tabu search solution approach for locating emergency medical stations.

Başar, A., Çatay, B., & Ünlüyurt, T. (2011). A Multi-period Double Coverage Approach for Locating the Emergency Medical Service Stations in Istanbul. *Journal of the Operational Research Society, 62*(4), 627-637.

Başar, A., Çatay, B., & Ünlüyurt, T. (2012). A Taxonomy for Emergency Service Station Location Problem. *Optimization Letters, 6*(6), 1147-1160.

Batta, R., Dolan, J. M., & Krishnamurthy, N. N. (1989). The Maximal Expected Covering Location Problem: Revisited. *Transportation*

Science, 23(4), 277-287.

Bayes, T., & Price, M. (1763). An Essay towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of Royal Society of London, 53*(0), 370-418.

Beraldi, P., & Bruni, M. (2009). A Probabilistic Model Applied to Emergency Service Vehicle Location. *European Journal of Operational Research, 196*(1), 323-331.

Beraldi, P., Bruni, M., & Conforti, D. (2004). Designing Robust Emergency Medical Service via Stochastic Programming. *European Journal of Operational Research, 158*(1), 183-193.

Berman, O., Hajizadeh, I., & Krass, D. (2013). The Maximum Covering Problem with Travel Time Uncertainty. *IIE Transactions, 43*(1), 81-96.

Berman, O., & Krass, D. (2002). The generalized maximal covering location problem. *Computers & Operations Research, 29*(6), 563-581.

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy, 81*(3), 637-654.

Brennan, M. J., & Schwartz, E. S. (1985). Evaluating Natural Resource Investments. *The Journal of Business, 58*(2), 135-157.

Brotcorne, L., Laporte, G., & Semet, F. (2003). Ambulance Location and Relocation Models. *European Journal of Operational Research, 147*(3), 451-463.

Buurman, J., Zhang, S., & Babovic, V. (2009). Reducing Risk Through Real Options in Systems Design: The Case of Architecting a Maritime
Domain Protection System. *Risk Analysis, 29*(3), 366-379.

Cardin, M.-A. (2007). *Facing Reality: Design and Management of Flexible Engineering Systems.* (Master of Science Thesis in Technology and Policy), Massachusetts Institute of Technology, Cambridge, MA, United States.

Cardin, M.-A. (2011). *Quantitative Performance-based Evaluation of a Procedure for Flexible Design Concept Generation.* (Doctoral Dissertation in Engineering Systems), Massachusetts Institute of Technology, Cambridge, MA, United States.

Cardin, M.-A. (2014). Enabling Flexibility in Engineering Systems: A Taxonomy of Procedures and a Design Framework. *Journal of Mechanical Design, 136*(1).

Cardin, M.-A., & de Neufville, R. (2009). *A Direct Interaction Approach to Identify Real Options 'In' Large-Scale Infrastructure Systems*. Paper presented at the Real Options Conference, Braga (Portugal), Santiago (Spain).

Cardin, M.-A., & de Neufville, R. (2013). *Design Catalogues: An Efficient Search Approach for Improved Flexibility in Engineering Systems Design*. Paper presented at the 23rd Symposium of the International Council on Systems Engineering, Philadelphia, PA.

Cardin, M.-A., de Neufville, R., & Geltner, D. M. (2015a). Design Catalogs: A Systematic Approach to Design and Value Flexibility in Engineering Systems. *Systems Engineering*, Accepted.

Cardin, M.-A., de Neufville, R., & Kazakidis, V. (2008). Process to Improve Expected Value of Mining Operations. *Mining Technology : IMM Transactions section A, 117*(2), 65-70.

Cardin, M.-A., & Hu, J. (2016). Analyzing the Tradeoffs Between Economies of Scale, Time-Value of Money, and Flexibility in Design Under Uncertainty: Study of Centralized vs. Decentralized Waste-to-Energy Systems. *Journal of Mechanical Design, 138*(1), 011401.

Cardin, M.-A., Jiang, Y., Yue, H. K.-H., & Fu, H. (2014). *Training Design and Management of Flexible Engineering Systems: An Empirical Study Using Simulation Games*. IEEE Transactions on Systems, Man, and Cybernetics: Systems.

Cardin, M.-A., Jiang, Y., Yue, H. K. H., & Fu, H. (2015b). Training Design and Management of Flexible Engineering Systems: An Empirical Study Using Simulation Games. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on, 45*(9), 1268-1280.

Cardin, M.-A., Kolfschoten, G., Frey, D. D., de Neufville, R., de Weck, O. L., & Geltner, D. (2013a). Empirical Evaluation of Procedures to Generate Flexibility in Engineering Systems and Improve Lifecycle

Performance. *Research in Engineering Design, 24*(3), 277-295.

Cardin, M.-A., Ranjbar Bourani, M., & de Neufville, R. (2015c). Improving the Lifecycle Performance of Engineering Projects with Flexible Strategies: Example of On-Shore LNG Production Design. *Systems Engineering, 18*(3), 253–268.

Cardin, M.-A., Steer, S. J., Nuttall, W. J., Parks, G. T., Gonçalves, L. V. N., & de Neufville, R. (2012). Minimizing the Economic Cost and Risk to Accelerator-Driven Subcritical Reactor Technology. Part 2: The Case of Designing for Flexibility. *Nuclear Engineering and Design, 243*, 120-134.

Cardin, M.-A., Yue, H. K.-H., Fu, H., Tang, L. C., Jiang, Y., Zhang, S., & Huang, B. (2013b). *Simulation Gaming to Study Design and Management Decision-Making in Flexible Engineering Systems.* Paper presented at the IEEE International Conference on System, Man, and Cybernetics, Manchester, United Kingdom.

Cavender, B. A. (2011). *A Review of the Methods of Economic Analysis of Nuclear Power Plants.* (Master of Science Thesis in Engineering), The University of Texas at Austin, United States.

Chanta, S., Mayorga, M. E., & McLay, L. A. (2011). Improving Emergency Service in Rural Areas: A Bi-objective Covering Location Model for EMS Systems. *Annals of Operations Research, 221*(1), 133-159.

Chapman, S., & White, J. (1974). *Probabilistic formulations of emergency service facilities location problems.* Paper presented at the ORSA/TIMS Conference, San Juan, Puerto Rico.

Charnes, A., & Cooper, W. W. (1959). Chance-Constrainted Programming. *Management Science, 6*(1), 73-79.

Chow, J. Y. J., & Regan, A. C. (2011). Network-based Real Option Models. *Transportation Research Part B: Methodological, 45*(4), 682-695.

Church, R., & ReVelle, C. (1974). The Maximal Covering Location Problem. *Papers in Regional Science: The Journal of the RSAI, 32*(1), 101-118.

Copeland, T. E., & Antikarov, V. (2001). *Real Options: A Practitioner's Guide*. New York, NY: Thomson Texere.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of financial Economics, 7*(3), 229-263.

Daskin, M. S. (1983). A Maximum Expected Covering Location Model: Formulation, Properties and Heuristic Solution. *Transportation Science, 17*(1), 48-70.

Daskin, M. S., Hopp, W. J., & Medina, B. (1992). Forecast Horizons and Dynamic Facility Location Planning. *Annals of Operations Research, 40*(1), 125-151.

de Neufville, R. (2010). ESD.71: Engineering Systems Analysis for Design. Cambridge, MA, United States: Massachusetts Institute of Technology.

de Neufville, R., & Scholtes, S. (2011). *Flexibility in Engineering Design*. Cambrige, MA, United States: MIT Press.

de Neufville, R., Scholtes, S., & Wang, T. (2006). Real Options by Spreadsheet: Parking Garage Case Example. *Journal of Infrastructure Systems, 12*(2), 107-111.

de Weck, O., de Neufville, R., & Chaize, M. (2004). Staged Deployment of Communications Satellite Constellations in Low Earth Orbit. *Journal of Aerospace Computing, Information, and Communication, 1*(3), 119-136.

de Weck, O., & Eckert, C. (2007). *A Classification of Uncertainty for Early Product and System Design*. Massachusetts Institute of Technology. Cambridge, MA, United States.

Deng, Y., Cardin, M.-A., Babovic, V., Santhanakrishnan, D., Schmitter, P., & Meshgi, A. (2013). Valuing Flexibilities in the Design of Urban Water Management Systems. *Water Research, 47*(20), 7162-7174.

Dias, M. A. G., & Teixeira, J. P. (2003). *Continuous-time option games: review of models and extensions. Part 1: Duopoly under uncertainty.* Paper presented at the Proceedings of the 2003 International Real Options Conference, George Town University, USA.

Diaz, B. A., & Rodriguez, F. (1997). A Simple Search Heuristic for the MCLP: Application to the Location of Ambulance Bases in a Rural Region. *Omega, 25*(2), 181-187.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton, New Jersey, USA: Princeton University Press.

DTI. (2007). *The Future of Nuclear Power in a Low Carbon Economy*. Retrieved from London, United Kingdom:

Du, Y., & Parsons, J. E. (2009). *Update on the Cost of Nuclear Power*. Retrieved from

Du, Y., & Parsons, J. E. (2012). *Capacity Factor Risk At Nuclear Power Plants*. Retrieved from

Eaton, D. J., Daskin, M. S., Simmons, D., Bulloch, B., & Jansma, G. (1985). Determining Emergency Medical Service Vehicle Deployment in Austin, Texas. *Interfaces, 15*(1), 96-108.

Eckert, C., de Weck, O., Keller, R., & Clarkson, P. J. (2009). *Engineering Change: Drivers, Sources, and Approaches in Industry*. Paper presented at the ICED 09 - the 17th International Conference on Engineering Design, Vol 4: Product and Systems Design, Palo Alto, CA, USA.

Eckhause, J., & Herold, J. (2014). Using Real Options to Determine Optimal Funding Strategies for CO2 Capture, Transport and Storage Projects in the European Union. *Energy Policy, 66*, 115-134.

Enerdata. (2015). Global Energy Statistical Yearbook - Electricity domestic consumption. Retrieved August 8, 2015, from https://yearbook.enerdata.net/electricity-domestic-consumption-data-by-region.html

Engel, A., & Browning, T. R. (2008). Designing Systems for Adaptability by Means of Architecture Options. *Systems Engineering, 11*(2), 125-146.

Eppinger, S. D., & Browning, T. R. (2012). *Design Structure Matrix Methods and Applications*. Cambrige, MA, United States: MIT Press.

Erol, İ., Sencer, S., Özmen, A., & Searcy, C. (2014). Fuzzy MCDM Framework for Locating a Nuclear Power Plant in Turkey. *Energy Policy, 67*, 186-197.

Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., & Goh, M. (2012). Covering Problems in Facility Location: A Review. *Computers &*

Industrial Engineering, 62(1), 368-407.

Farahani, R. Z., Drezner, Z., & Asgari, N. (2009). Single Facility Location and Relocation Problem with Time Dependent Weights and Discrete Planning Horizon. *Annals of Operations Research, 167*(1), 353-368.

Faria, A. J., Hutchinson, D., Wellington, W. J., & Gold, S. (2008). Developments in Business Gaming: A Review of the Past 40 Years. *Simulation & gaming, 40*(4), 464-487.

Ferreira, N., Kar, J., & Trigeorgis, L. (2009). Option Games: The Key to Competing in Capital-Intensive Industries. *Harvard Business Review, 87*(3), 101-107.

Finger, S., & Dixon, J. R. (1989a). A Review of Research in Mechanical Engineering Design. Part I: Descriptive, Prescriptive, and Computer-Based Models of Design Processes. *Research in Engineering Design, 1*(1), 51-67.

Finger, S., & Dixon, J. R. (1989b). A Review of Research in Mechanical Engineering Design. Part II: Representations, Analysis, and Design for the Life Cycle. *Research in Engineering Design, 1*(2), 121-137.

Fossa, C. E., Raines, R. A., Gunsch, G. H., & Temple, M. A. (1998). *An Overview of the IRIDIUM (R) Low Earth Orbit (LEO) Satellite System.* Paper presented at the Aerospace and Electronics Conference, 1998. NAECON 1998. Proceedings of the IEEE 1998 National.

Frey, D. D., & Wang, H. (2006). Adaptive One-Factor-at-a-Time Experimentation and Expected Value of Improvement. *Technometrics, 48*(3), 418-431.

Fricke, E., & Schulz, A. P. (2005). Design for Changeability (DfC): Principles to Enable Changes in Systems Throughout Their Entire Lifecycle. *Systems Engineering, 8*(4), 342-359.

GAIA. (2015). Serious Games. Retrieved March 12, 2015, from http://cs.gmu.edu/~gaia/SeriousGames/index.html

Gendreau, M., Laporte, G., & Semet, F. (1997). Solving an Ambulance Location Model by Tabu Search. *Location science, 5*(2), 75-88.

Gendreau, M., Laporte, G., & Semet, F. (2001). A Dynamic Model and

Parallel Tabu Search Heuristic for Real-time Ambulance Relocation. *Parallel computing, 27*(12), 1641-1653.

Ghaderi, A., & Jabalameli, M. S. (2013). Modeling the Budget-constrained Dynamic Uncapacitated Facility Location–network Design Problem and Solving it via Two Efficient Heuristics: A Case Study of Health Care. *Mathematical and Computer Modelling, 57*(3–4), 382-400.

Giffin, M., de Weck, O., Bounova, G., Keller, R., Eckert, C., & Clarkson, P. J. (2009). Change Propagation Analysis in Complex Technical Systems. *Journal of Mechanical Design, 131*(8), 1-14.

Glasstone, S., & Sesonske, A. (1994). *Nuclear Reactor Engineering: Reactors Systems Engineering* (Vol. 2): Springer, 4th edition.

Glover, F. (1990). Tabu Search: A Tutorial. *Interfaces, 20*(4), 74-94.

Golay, M. W. (2001). *On Social Acceptance of Nuclear Power*. Retrieved from Workshop on Nuclear Energy Technologies: A Policy Framework for Micro-Nuclear Technology, James A. Baker III Institute for Public Policy, Rice University, Houston TX:

Goldberg, J. B. (2004). Operations Research Models for the Deployment of Emergency Services Vehicles. *EMS Management Journal, 1*, 20-39.

Gollier, C., Proult, D., Thais, F., & Walgenwitz, G. (2005). Choice of Nuclear Power Investments under Price Uncertainty: Valuing Modularity. *Energy Economics, 27*(4), 667-685.

Gorbachev, M. (1996) *The Battle of Chernobyl/Interviewer: T. Johnson*. Discovery Channel.

Grimston, M. (2006). Nuclear Energy. In T. Jamasb, W. J. Nuttall, & M. G. Pollitt (Eds.), *Future Electricity Technologies and Systems*: Cambridge University Press.

Grimston, M., Nuttall, W. J., & Vaughan, G. (2014). The Siting of UK Nuclear Reactors. *Journal of Radiological Protection, 34*(2), 1-24.

Guma, A., Pearson, J., Wittels, K., Neufville, R. d., & Geltner, D. (2009). Vertical Phasing as a Corporate Real Estate Strategy and Development Option. *Journal of Corporate Real Estate, 11*(3), 144-157.

Gunawardane, G. (1982). Dynamic Versions of Set Covering Type Public

Facility Location Problems. *European Journal of Operational Research, 10*(2), 190-195.

Halpern, J. Y. (2003). *Reasoning About Uncertainty*. Cambridge, MA: MIT Press.

Hassan, R., & de Neufville, R. (2006). *Design of Engineering Systems Under Uncertainty Via Real Options and Heuristic Optimization*. Paper presented at the Real Options Conference, New York, NY.

Hazelrigg, G. A. (1998). A Framework for Decision-Based Engineering Design. *Journal of Mechanical Design, 120*(4), 653-658.

He, G., Mol, A. P. J., Zhang, L., & Liu, Y. (2013). Public Participation and Trust in Nuclear Power Development in China. *Renewable and Sustainable Energy Reviews, 23*, 1-11.

Helmer, O. (1967). *Analysis of the future: The Delphi method*. Retrieved from

Hesseldahl, A. (2001). The Return Of Iridium. Retrieved March 5, 2015, from http://www.forbes.com/2001/11/30/1130tentech.html

Hogan, K., & ReVelle, C. (1986). Concepts and applications of backup coverage. *Management Science, 32*(11), 1434-1444.

Howard, R. A. (1966). *Decision Analysis: Applied Decision Theory*: Stanford Research Institute.

Hu, J. (2013). *Flexible Engineering System Design with Multiple Exogenous Uncertainties and Change Propagation.* (Doctorial Dissertation in Industrial and Systems Engineering), National University of Singapore, Singapore.

Hu, J., & Poh, K. L. (2010). *Pareto Set-Based Concept Modeling and Selection for Large-Scale Systems*. Paper presented at the 4th Asia-Pacific Conference on Systems Engineering, Keelung, Taiwan.

Huang, L., Zhou, Y., Han, Y., Hammitt, J. K., Bi, J., & Liu, Y. (2013). Effect of the Fukushima Nuclear Accident on the Risk Perception of Residents near a Nuclear Power Plant in China. *Risk Analysis, 110*(49), 19742-19747.

Hultman, N. E., Koomey, J. G., & Kammen, D. M. (2007). What Histroy Can Teach Us About the Future Costs of U.S. Nuclear Power.

Environmental Science and Technology, 41(7), 2088-2093.

IAEA. (2008). *Nuclear Technology Review*. Retrieved from

Institute of Electrical and Electronics Engineers. (1990). IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. New York, NY.

International Atomic Energy Agency. (2015). INES - The International Nuclear and Radiological Event Scale. Retrieved May 20, 2015, from http://www-ns.iaea.org/tech-areas/emergency/ines.asp

Jacoby, H. D., & Loucks, D. P. (1972). The Combined Use of Optimization and Simulation Models in River Basin Planning. *Water Resource Research, 8*(6), 1401-1414.

Jain, S., Roelofs, F., & Oosterlee, C. W. (2013). Valuing Modular Nuclear Power Plants in Finite Time Decision Horizon. *Energy Economics, 36*, 625-636.

Jia, H., Ordóñez, F., & Dessouky, M. (2007a). A Modeling Framework for Facility Location of Medical Services for Large-scale Emergencies. *IIE Transactions, 39*(1), 41-55.

Jia, H., Ordóñez, F., & Dessouky, M. M. (2007b). Solution Approaches for Facility Location of Medical Supplies for Large-scale Emergencies. *Computers & Industrial Engineering, 52*(2), 257-276.

Joskow, P. L. (2006). *The Future of Nuclear Power in the United States: Economic and Regulatory Challenges*. Retrieved from MIT Center for Energy and Environmental Policy Research:

Joskow, P. L., & Parsons, J. E. (2012). The Future of Nuclear Power After Fukushima. *Economics of Energy and Environmental Policy, 1*(2), 99-113.

Jugulum, R., & Frey, D. D. (2007). Toward a Taxonomy of Concept Designs for Improved Robustness. *Journal of Engineering Design, 18*(2), 139-156.

Kalligeros, K. (2006). *Platforms and Real Options in Large-Scale Engineering Systems.* (Doctoral Dissertation in Engineering Systems), Massachusetts Institute of Technology, Cambridge, MA, United States.

Kamrad, B., & Ritchken, P. (1991). Multinomial Approximating Models for Options with k State Variables. *Management Science, 37*(12), 1640-1652. doi:doi:10.1287/mnsc.37.12.1640

Keese, D. A., Seepersad, C. C., & Wood, K. L. (2009). Product Flexibility Measurement With Enhanced Change Modes and Effects Analysis (CMEA). *International Journal of Mass Customisation, 3*(2), 115-145.

Kelly, S. (1998). A binomial lattice approach for valuing a mining property IPO. *The Quarterly Review of Economics and Finance, 38*(3, Part 2), 693-709.

Kessides, I. N. (2010). Nuclear Power: Understanding the Economic Risks and Uncertainties. *Energy Policy, 38*(8), 3849-3864.

Khansa, L., & Liginlal, D. (2009). Valuing the flexibility of investing in security process innovations. *European Journal of Operational Research, 192*(1), 216-235.

Kidd, S. W. (2013). Nuclear Power - Economics and Public Acceptance. *Energy Strategy Reviews*, 277-281.

Kiriyama, E., & Suzuki, A. (2004). Use of Real Options in Nuclear Power Plant Valuation in the Presence of Uncertainty with CO2 Emission Credit. *Journal of Nuclear Science and Technology, 41*(7), 756-764.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science, 220*, 671-680.

Kojo, M., & Richardson, P. (2014). The Use of Community Benefits Approaches in the Siting of Nuclear Waste Management Facilities. *Energy Strategy Reviews, 4*, 34-42.

Land, A. H., & Doig, A. G. (1960). An Automatic Method of Solving Discrete Programming Problems. *Econometrica, 28*(3), 497-520.

Larson, R. C. (1974). A Hypercube Queuing Model for Facility Location and Redistricting in Urban Emergency Services. *Computers & Operations Research, 1*(1), 67-95.

Lavelle, M. (Producer). (2012). One Year after Fukushima, Japan Faces Shortage of Energy, Trust. Retrieved Access Date, from http://news.nationalgeographic.com/news/energy/2012/03/120309-japa n-fukushima-anniversary-energy-shortage/

Ligtvoet, A., & Herder, P. M. (2012). Simulation and Gaming for Understanding the Complexity of Cooperation in Industrial Networks. In O. Hammami, D. Krob, & J.-L. Voirin (Eds.), *Complex Systems Design & Management* (pp. 81-92): Springer Berlin Heidelberg.

Locatelli, G., Boarin, S., Pellegrino, F., & Ricotti, M. E. (2015). Load Following with Small Modular Reactors (SMR): A Real Options Analysis. *Energy, 80*, 41-54.

Louberge, H., Villeneuve, S., & Chesney, M. (2002). Long-term Risk Management of Nuclear Waste: A Real Options Approach. *Journal of Economic Dynamics and Control, 27*(1), 157-180.

Marianov, V., & ReVelle, C. (1992a). The Capacitated Standard Response Fire Protection Siting Problem: Deterministic and Probabilistic Models. *Annals of Operations Research, 40*(1), 303-322.

Marianov, V., & ReVelle, C. (1992b). A Probabilistic Fire-protection Siting Model with Joint Vehicle Reliability Requirements. *Papers in regional science, 71*(3), 217-241.

Marianov, V., & ReVelle, C. (1994). The Queuing Probabilistic Location Set Covering Problem and Some Extensions. *Socio-Economic Planning Sciences, 28*(3), 167-178.

Marianov, V., & ReVelle, C. (1996). The Queueing Maximal Availability Location Problem: A Model for the Siting of Emergency Vehicles. *European Journal of Operational Research, 93*(1), 110-120.

Marianov, V., & Serra, D. (2001). Hierarchical Location-Allocation Models for Congested Systems. *European Journal of Operational Research, 135*(1), 195-208.

Marianov, V., & Serra, D. (2002). Location–Allocation of Multiple-Server Service Centers with Constrained Queues or Waiting Times. *Annals of Operations Research, 111*(1-4), 35-50. doi:10.1023/A:1020989316737

Marreco, J. d. M., & Carpio, L. G. T. (2006). Flexibility Valuation in the Brazilian Power System: A Real Options Approach. *Energy Policy, 34*(18), 3749-3756.

Martinez-Cesena, E. A., Mutale, J., & Rivas-Davalos, F. (2013). Real Options Theory Applied to Electricity Generation Projects: A Review. *Renewable and Sustainable Energy Reviews, 19*, 573-581.

McLay, L. A. (2009). A Maximum Expected Covering Location Model with Two Types of Servers. *IIE Transactions, 41*(8), 730-741.

McLay, L. A., & Mayorga, M. E. (2011). Evaluating the Impact of Performance Goals on Dispatching Decisions in Emergency Medical Service. *IIE Transactions on Healthcare Systems Engineering, 1*(3), 185-196.

McLay, L. A., & Mayorga, M. E. (2013a). A Dispatching Model for Server-to-Customer Systems That Balances Efficiency and Equity. *Manufacturing & Service Operations Management, 15*(2), 205-220.

McLay, L. A., & Mayorga, M. E. (2013b). A Model for Optimally Dispatching Ambulances to Emergency Calls with Classification Errors in Patient Priorities. *IIE Transactions, 45*(1), 1-24.

Melese, Y. G., Heijnen, P. W., Stikkelman, R. M., & Herder, P. M. (2015). Exploring for Real Options During CCS Networks Conceptual Design to Mitigate Effects of Path-dependency and Lock-in. *International Journal of Greenhouse Gas Control, 42*, 16-25.

Mishra, S. (2012). Social Acceptance of Nuclear Power in India. *Air Power Journal, 7*(3), 55-82.

Morgan, M. G., & Henrion, M. (1992). *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, UK: Cambridge University Press.

Myers, S. C. (1984). Finance Theory and Financial Strategy. *Interfaces, 14*(1), 126-137.

Myerson, R. B. (1997). *Game Theory: Analysis of Conflict*: Harvard University Press.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences, 36*(1), 48-49.

Noyan, N. (2010). Alternate Risk Measures for Emergency Medical Service System Design. *Annals of Operations Research, 181*(1), 559-589.

OECD/NEA. (2003). *Nuclear Electricity Generation: What Are the External Costs?* Retrieved from

Olewnik, A., Brauen, T., Ferguson, S., & Lewis, K. (2003). A Framework for Flexible Systems and Its Implementation in Multiattribute Decision Making. *Journal of Mechanical Design, 126*(3), 412-419.

Olewnik, A., & Lewis, K. (2006). A Decision Support Framework for Flexible System Design. *Journal of Engineering Design, 17*(1), 75-97.

Ong, M. E., Ng, F. S., Overton, J., Yap, S., Andresen, D., Yong, D. K., Lim, S. H., & Anantharaman, V. (2009). Geographic-time distribution of ambulance calls in Singapore: utility of geographic information system in ambulance deployment (CARE 3). *Ann Acad Med Singapore, 38*(3), 184-191.

Owen, S. H., & Daskin, M. S. (1998). Strategic Facility Location: A Review. *European Journal of Operational Research, 111*(3), 423-447.

Pindyck, R. S. (2000). Irreversibilities and the Timing of Environmental Policy. *Resource and Energy Economics, 22*(3), 233-259.

Pirkul, H., & Schilling, D. A. (1988). The Siting of Emergency Service Facilities with Workload Capacities and Backup Service. *Management Science, 34*(7), 896-908.

Pirkul, H., & Schilling, D. A. (1991). The Maximal Covering Location Problem with Capacities on Total Workload. *Management Science, 37*(2), 233-248.

Pouret, L., Buttery, N., & Nuttall, W. J. (2009). Is Nuclear Power Inflexible? *Nuclear Future, 5*(6), 333-341 and 343-344.

Power Reactor Information System (IAEA). (2015). PRIS - Home. Retrieved March 20, 2015, from http://www.iaea.org/pris/

Pringles, R., Olsina, F., & Garcés, F. (2015). Real Option Valuation of Power Transmission Investments by Stochastic Simulation. *Energy Economics, 47*, 215-226.

Pumain, D. (Ed.) (2006). *Hierarchy in Natural and Social Sciences*: Springer Netherlands.

Qureshi, A., Murphy, J. T., Kuchinsky, B., Seepersad, C. C., Wood, K. L., &

Jensen, D. D. (2006). *Principles of Product Flexibility.* Paper presented at the ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Philadelphia, Pennsylvania, USA.

Rajagopalan, H. K., Saydam, C., & Xiao, J. (2008). A Multiperiod Set Covering Location Model for Dynamic Redeployment of Ambulances. *Computers & Operations Research, 35*(3), 814-826.

Rajagopalan, H. K., Vergara, F. E., Saydam, C., & Xiao, J. (2007). Developing Effective Meta-heuristics for a Probabilistic Location Model via Experimental Design. *European Journal of Operational Research, 177*(1), 83-101.

Rajan, P. K. P., Wie, M. V., Campbell, M. I., Wood, K. L., & Otto, K. N. (2005). An Empirical Foundation for Product Flexibility. *Design Studies, 26*(4), 405-438.

Rekik, M., Ruiz, A., Renaud, J., Berkoune, D., & Paquet, S. (2013). A Decision Support System for Hunmanitarian Network Design and Distribution Operations *Humanitarian and Relief Logistics* (pp. 1-20): Springer New York.

ReVelle, C., & Hogan, K. (1989a). The Maximum Availability Location Problem. *Transportation Science, 23*(3), 192-200.

Revelle, C., & Hogan, K. (1989b). The Maximum Reliability Location Problem and α-reliable p-center Problem: Derivatives of the Probabilistic Location Set Covering Problem. *Annals of Operations Research, 18*(1), 155-173.

Rose, D. H. (1985). Nuclear Power. *Bulletin of the Atomic Scientists*.

Ross, A. M. (2006). *Managing Unarticulated Value: Chageability in Multi-attribute Tradespace Exploration.* (Doctoral Dissertation in Engineering Systems), Massachusetts Institute of Technology, Cambridge, MA, United States.

Ross, S. M. (1995). *Stochastic Processes*. Canada: Wiley.

Ross, S. M. (2014). *Introduction to Probability Models*. Amsterdam: Elsevier.

Rothwell, G. (1998). Comparing Asian Nuclear Power Plant Performance.

Pacific and Asian Journal of Energy, 8(1), 51-64.

Rothwell, G. (2006). A Real Options Approach to Evaluating New Nuclear Power Plants. *The Energy Journal, 27*(1), 37-53.

Rothwell, G. (2009). The Value of Being Able to Start Construction of A Nuclear Power Plant in Chile by 2020. *International Journal of Nuclear Governance, 2*(4), 323-336.

Rothwell, G. (2010). International Light Water Nuclear Fuel Fabrication Supply: Are Fabrication Services Assured? *Energy Economics, 32*(3), 538-544.

Saleh, J. H., Mark, G., & Jordan, N. C. (2009). Flexibility: A Multi-Disciplinary Literature Review and a Research Agenda for Designing Flexible Engineering Systems. *Journal of Engineering Design, 20*(3), 307-323.

Savage, S. (2002). The Flaw of Averages. *Harvard Business Review, 4*.

Savas, E. (1969). Simulation and Cost-effectiveness Analysis of New York's emergency Ambulance Service. *Management Science, 15*(12), 608-627.

Saydam, C., Repede, J. F., & Burwell, T. (1994). Accurate Estimation of Expected Coverage: A Comparative Study. *Socio-Economic Planning Sciences, 28*(2), 113-120.

Schilling, D. A. (1980). Dynamic Location Modeling for Public-sector Facilities: A Multicriteria Approach. *Decision Sciences, 11*(4), 714-724.

Schilling, D. A., Elzinga, D. J., Cohon, J., Church, R., & ReVelle, C. (1979). The TEAM/FLEET Models for Simultaneous Facility and Equipment Siting. *Transportation Science, 13*(2), 163-175.

Schmid, V., & Doerner, K. F. (2010). Ambulance Location and Relocation Problems with Time-dependent Travel Times. *European Journal of Operational Research, 207*(3), 1293-1303.

Serra, D. (1996). The Coherent Covering Location Problem. *Papers in Regional Science: The Journal of the RSAI, 75*(1), 79-101.

Siddiqui, A., & Fleten, S.-E. (2010). How to Proceed with Competing

Alternative Energy Technologies: A Real Options Analysis. *Energy Economics, 32*(4), 817-830.

Silva, F., & Serra, D. (2008). Locating Emergency Services with Different Priorities: The Priority Queuing Covering Location Problem. *Journal of the Operational Research Society, 59*(9), 1229-1238.

Simon, H. A. (1977). *The New Science of Management Decision*: Prentice-Hall.

Smit, H. (2001). Acquisition Strategies as Option Games. *Journal of Applied Corporate Finance, 14*(2), 79-89.

Smit, H., & Trigeorgis, L. (2009). Valuing Infrastructure Investment: An Option Games Approach. *California Management Review, 51*(2), 82-104.

Song, Y., Kim, D., & Han, D. (2013). Risk Communication in South Korea: Social Acceptance of Nuclear Power Plants (NPPs). *Public Relations Review, 39*(1), 55-56.

Sorensen, P., & Church, R. (2010). Integrating expected coverage and local reliability for emergency medical services location problems. *Socio-Economic Planning Sciences, 44*(1), 8-18.

Steer, S. J., Cardin, M.-A., Nuttall, W. J., Parks, G. T., & Gonçalves, L. V. N. (2012). Minimising the Economic Cost and Risk to Accelerator-Driven Subcritical Reactor Technology: The Case of Designing for Flexibility: Part 1. *Nuclear Engineering and Design, 243*, 135-147. doi:10.1016/j.nucengdes.2011.11.027

Steer, S. J., Nuttall, W. J., Parks, G. T., & Gonçalves, L. V. N. (2011). Predicting the Cost of Unplanned Shutdowns of Power Stations: An Accelerator-Driven Subcritical Reactor Case Study. *Electric Power Systems Research, 81*(8), 1662-1671.

Steward, D. V. (1981). The Design Structure System: A Method for Managing the Design of Complex Systems. *IEEE Transactions on Engineering Management, 28*(3), 71-74.

Strauss, A., & Corbin, J. M. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*: SAGE

Publications.

Suh, N. P. (1990). *The Principles of Design*: Oxford University Press, USA.

Sung, I., & Lee, T. (2012, 9-12 Dec. 2012). *Modeling Requirements for an Emergency Medical Service System Design Evaluator.* Paper presented at the Simulation Conference (WSC), Proceedings of the 2012 Winter.

Syam, S. S., & Côté, M. J. (2010). A location–allocation model for service providers with application to not-for-profit health care organizations. *Omega, 38*(3), 157-166.

Taguchi, G. (1987). *The System of Experimental Design Engineering Methods to Optimize Quality and Minimize Cost* (Vol. 1 and 2). Dearborn, MI, United States: American Supplier Institute.

Thomas, S. (2005). *The Economics of Nuclear Power*. Heinrich Boll Shiftung, Nuclear Issues Paper No. 5.

Tomiyama, T. (2006). *A Classification of Design Theories and Methodologies.* Paper presented at the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Philadelphia, PA, United States.

Tomiyama, T., Gu, P., Jin, Y., Lutters, D., Kind, C., & Kimura, F. (2009). Design methodologies: Industrial and educational applications. *CIRP Annals - Manufacturing Technology, 58*(2), 543-565.

Tomiyama, T., & Yoshikawa, H. (1987). Extended General Design Theory. In H. Yoshikawa & E. A. Warman (Eds.), *Design Theory for CAD* (pp. 95-130). Amsterdam, North-Holland, Netherlands.

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The Location of Emergency Service Facilities. *Operations Research, 19*(6), 1363-1373.

Trigeorgis, L. (1996). *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. Cambrige, MA.: MIT Press.

Tseng, C.-L., & Graydon, B. (2002). Short-Term Generation Asset Valuation: A Real Options Approach. *Operations Research, 50*(2), 297-310. doi:10.2307/3088497

Urich, C., & Rauch, W. (2014). Exploring Critical Pathways for Urban Water Management to Identify Robust Strategies under Deep Uncertainties.

Water Research, 66, 374-389.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*: Princeton University Press, Princeton.

Wakatsuki, Y. (Producer). (2014). New Radioactive Water Leak at Japan's Fukushima Daiichi Plant. Retrieved Access Date, from http://edition.cnn.com/2014/02/19/world/asia/japan-fukushima-daiichi-water-leak/

Walsh, J. (2007). Hydropower. In P. Robbins (Ed.), *Encyclopedia of Environment and Society*: SAGE Publications, Inc.

Wang, H. (2015). *Applying Heuristic Search in Optimizing Flexible Design and Management Decisions in Emergency Medical Service System.* (Bachelor of Engineering thesis in Industrial and Systems Engineering), National University of Singapore, Singapore.

Wang, T. (2005). *Real Options "in" Projects and Systems design - Identification of Options and Solution for Path Dependency.* (Doctoral Dissertation in Engineering Systems), Massachusetts Institute of Technology, Cambridge, MA, United States.

Wang, T., & de Neufville, R. (2005). *Real Options "in" Projects.* Paper presented at the 9th Real Options Annual International Conference, Paris, France.

Wang, X., Cai, Y., & Dai, C. (2014). Evaluating China's Biomass Power Production Investment Based on A Policy Benefit Real Options Model. *Energy, 73*, 751-761.

Webb, G. A. M., Anderson, R. W., & Gaffney, M. J. S. (2006). Classification of Events with An Off-site Radiological Impact at the Sellafield Site between 1950 and 2000, Using the International Nuclear Event Scale. *Journal of Radiological Protection, 26*, 33-49.

World Nuclear Assocation. (2015). Decommissioning Nuclear Facilities. Retrieved March 19, 2015, from http://www.world-nuclear.org/info/Nuclear-Fuel-Cycle/Nuclear-Wastes/Decommissioning-Nuclear-Facilities/

World Nuclear News. (2015). Event scale revised for further clarity.

Retrieved May 20, 2015, from http://www.world-nuclear-news.org/RS_Event_scale_revised_for_further_clarity_0510081.html

Yoshikawa, H. (1981). General Design Theory and a CAD System. In T. Sata & E. A. Warman (Eds.), *Man-Machine Communication in CAD/CAM* (pp. 35-58). Amsterdam, North-Holland, Netherlands.

Yoshikawa, H., & Uehara, K. (1985). Design Theory for CAD/CAM Integration. *CIRP Annals - Manufacturing Technology, 34*(1), 173-178.

Yue, Y., Marla, L., & Krishnan, R. (2012). *An Efficient Simulation-based Approach to Ambulance Fleet Allocation and Dynamic Redeployment.* Paper presented at the National Conference on Artificial Intelligence (AAAI).

Zarandi, M. H. F., Davari, S., & Sisakht, S. A. H. (2013). The Large-scale Dynamic Maximal Covering Location Problem. *Mathematical and Computer Modelling, 57*(3–4), 710-719.

Zhou, Y., Li, Y. P., & Huang, G. H. (2015). A Robust Possibilistic Mixed-integer Programming Method for Planning Municipal Electric Power Systems. *International Journal of Electrical Power & Energy Systems, 73*, 757-772.

Zhu, L. (2012). A Simulation Based Real Options Approach for The Investment Evaluation of Nuclear Power. *Computers & Industrial Engineering, 63*(3), 585-593.