Robust Design of Supply Network Subject to Disruptions by Considering Congestion Effects

Alireza Ebrahim Nejad

A Thesis

In the Department

of

Mechanical, Industrial and Aerospace Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Industrial Engineering) at

Concordia University

Montreal, Quebec, Canada

July 2019

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Alireza Ebrahim Nejad**

Entitled: **Robust Design of Supply Network Subject to Disruptions by Considering Congestion Effects**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Industrial Engineering)

complies with the regulation of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

Dr. Abdel R. Sebak — Chair

Dr. Walid Klibi — External Examiner

Dr. Navneet Vidyarthi — External to Program

Dr. Ali Akgunduz — Examiner

Dr. Masoumeh Kazemi Zanjani — Examiner

Dr. Onur Kuzgunkaya — Supervisor

Approved by
Dr. Ali Dolatabadi, Graduate Program Director

Monday, September 09, 2019

Dr. Amir Asif , Dean, Faculty of Engineering and Computer Science

Abstract

Robust Design of Supply Network Subject to Disruptions by Considering Congestion Effects

Alireza Ebrahim Nejad
Concordia University, 2019

This thesis is focused on the supply chain disruptions and it reviews cost-efficient risk mitigation strategies to sustain supply chain functionality when disruptions occur. In particular, we study the robust design of supply flow subject to minor operational risks and major disruptions. The contingent sourcing along with strategic stock is incorporated as risk management strategies. We consider a firm with two suppliers where the main supplier is cost-effective but prone to disruptions and the back-up supplier is reliable but expensive. The back-up supplier can scale up its capacity according to a speed related to its configuration in order to supply the required flow of material when the main supplier disrupts. When minor disruption occurs, the strategic stock can cover the losses. The design problem considered is to determine optimal strategic stock level and response speed of volume-flexible back-up supplier.

The back-up supplier might not provide the required supply level instantaneously due to non-steady production state and congestion during the response time. Therefore, there could be material shortages if the actual level of available capacity during the response time is ignored. The first chapter includes the incorporation of the clearing function into a contingency capacity planning model in order to represent the impact of congestion. The appropriate response speed is selected through a decision tree analysis considering different attitudes of the decision maker towards risk. The results show that considering congestion impact is especially critical for risk-neutral decision makers. The second chapter considers the randomness associated with the available capacity through a two-stage robust optimization model. The results show improvement in the quality of optimal solution by considering the randomness. The objective in the third chapter is to find an equitable solution which has an efficient performance with respect to all plausible scenarios. Therefore, the Ordered Weighted Averaging aggregation operator is incorporated in the objective function of a MIP robust model. In order to address the computational complexity associated with large set of scenarios, a novel clustering based scenario reduction model based on location covering model is proposed. The results show that the proposed methodology provide an accurate reduced scenario set within relatively short computational time.

Acknowledgement

I would like to express my special thanks to my supervisor, Dr. Onur Kuzgunkaya, for his guidance, all the insightful comments, and full time enrollment. He was always willing to help me, and I would not have been able to finish this journey without his support. I am also very thankful to the members of my PhD committee, Dr. Akgunduz, Dr. Kazemi, Dr. Klibi and Dr. Vidyarthi for their valuable comments and feedback throughout my PhD journey.

To my colleagues at Concordia University especially Dr. Iman Niroomand, I am grateful to have worked with you and to being a part of your community.

Special thanks to my lovely family for the encouragement which helped me in the completion of this thesis. Words cannot explain how grateful I am to my parents, Abbas and Ziba, my sisters, Mina, Lida, Nazanin, my dear brother in law, Mr. Nasser Ajabi for all his supports, and my nephews Bardia and Arshia.

My beloved and supportive best friend, Parya, thank you for being by my side and all the support and love you gave me which helped me in the completion of this thesis.

Table of Contents

List of Figures ... viii
List of Tables ... ix
Chapter 1: Introduction ... 1
1.1 Supply Disruptions ... 1
1.2 Robust Supply Chain Network Design ... 3
1.3 Supply Mitigation Strategies ... 5
1.4 Scope and Objectives ... 9
1.5 Thesis Outline ... 10
Chapter 2: Literature Review ... 11
2.1 Contingent Sourcing ... 11
2.1.1 Response Time ... 13
2.1.2 Response Time Characteristics ... 15
2.1.3 The Impacts of Congestion on Throughput during the Response Time ... 15
2.2 Robustness ... 16
2.2.1 Solution Robustness ... 16
2.2.2 Solution Robustness under Deep Uncertainty ... 17
2.3 Scenario Reduction ... 18
2.4 Conclusion ... 21
Chapter 3: Responsive contingency planning in supply risk management by considering congestion effects ... 23
3.1 Solution methodology ... 27
3.1.1 Capacity planning model ... 29
3.1.2 Representation of disruptions in the MIP model ... 40
3.1.3 Decision tree analysis ... 41
3.2 Numerical Experiments ... 44
3.2.1 The impact of congestion on contingency strategy performance ... 46

3.2.2 The value of incorporating congestion in the selection of response speed 49
3.2.3 Discussion of results 58
Chapter 4: On the value of response time characteristics in robust design of supply flow......................... 60
4.1 Solution Methodology 62
4.1.1 Robust Stochastic Optimization Model 65
4.1.2 Scenario generation... 70
4.2 Numerical Experiments... 73
4.2.1 The impact of considering the available capacity randomness 75
4.2.2 The importance of incorporating congestion 80
4.2.3 Discussion of results 82
Chapter 5: A Clustering based Scenario Reduction Approach to Design Robust Supply Network under Operational Risks and Disruptions 84
5.1 Solution Methodology 86
5.1.1 Clustering based MIP Scenario Reduction (CBSR) Model 87
5.1.2 Capacity Planning Model... 91
5.1.3 Robust SCND Model 96
5.2 Numerical Experiments... 97
5.2.1 Investigating the performance of CBSR model 99
5.2.2 Robust Formulation Assessment... 104
5.2.3 Discussion of Results... 111
Chapter 6: Conclusion... 113
Future Research Directions... 115
References... 117
Appendix A... 128
Appendix B 131
Appendix C 134

List of Figures

Figure 3.1 The supply chain network configuration 24
Figure 3.2 An example of a contingency plan execution. 25
Figure 3.3 Solution methodology. 28
Figure 3.4 The throughput curve during response time. 35
Figure 3.5 Clearing function. 38
Figure 3.6 The decision tree for the optimal selection of the response speed (e.g., $T=8$). 42
Figure 3.7 The impact of considering congestion effects on service level. 49
Figure 3.8 Sensitivity analysis of the expected costs of the supply chain. 50
Figure 3.9 Optimal response speed, Risk-neutral. 51
Figure 3.10 Optimal response speed, Risk-averse, Load-independent models. 53
Figure 3.11 Optimal response speed, Risk-averse, Load-dependent models 54
Figure 3.12 Optimal response speed, Mean risk, Load-independent models, γ = 97.5%. 56
Figure 3.13 Optimal response speed, Mean risk, Load-dependent models, γ = 97.5%. 56

Figure 4.1 A snapshot of the scenario tree 72
Figure 4.2 Robust design of the supply flow, Experiment 1 76
Figure 4.3 Robust design of the supply flow, Experiment 2 77
Figure 4.4 Robust design of the supply flow, Experiment 3 78
Figure 4.5 Robust design of supply flow, Experiment 4. 78
Figure 4.6 Robust design of the supply flow, *RO VS LIRO*, Experiment 2 80

Figure 5.1 Solution Methodology 87
Figure 5.2 Cumulative distribution of cost of optimal solutions, small scenario set (162 scenarios) 101
Figure 5.3 Cumulative distribution of cost of optimal solutions, medium scenario set (556 scenarios) 101
Figure 5.4 Cumulative distribution of cost of optimal solutions, large scenario set (1571 scenarios) 102
Figure 5.5 Safety Stock Selections of Robust Formulations. 108

Figure A.1 The clearing function approximated by lines, N=3. 128
Figure A.2 The clearing function approximated by planes, V =3. 129

Figure B.1 The no-disruption scenario. 131
Figure B.2 A disruption scenario with time of occurrence = m, length = n. 132
Figure B. 3 A disruption scenario with time of occurrence = m, $n+m=T+1$. 133

List of Tables

Table 2.1 Summary of the reviewed supply chain design problems with contingent sourcing 13

Table 3.1 List of Notations 29
Table 3.2 List of decision variables 30
Table 3.3 Demand scenario... 44
Table 3.4 Supplier’s costs parameters $/Unit 44
Table 3.5 Excess and shortage costs $/Unit 45
Table 3.6 RMS capacity levels with respect to its structure pcs/Period 45
Table 3.7 The coefficients of capacity boundaries corresponding to different speeds 46
Table 3.8 Capacity planning in load-independent versus load-dependent models 47
Table 3.9 The expected service level of the supply chain (%)... 52
Table 3.10 Improvements in the expected service level of the supply chain with load-dependent models (%), Risk-averse 55
Table 3.11 Improvements in the expected service level of the supply chain with load-dependent models (%), Mean-risk 57

Table 4.1 List of notations 62
Table 4.2 List of decision variables 63
Table 4.3 Cost parameters... 73
Table 4.4 Fractions of the added capacity available in different scenarios................................ 74
Table 4.5 Experiments based on the probabilities of available capacity levels during response time........ 75
Table 4.6 % Reduction in the objective function by RO solution 79
Table 4.7 Improvement in total costs by considering congestion effects 81

Table 5.1 List of Decision Variables in Clustering based MIP Scenario Reduction Model....................... 89
Table 5.2 List of Input Parameters in Clustering based MIP Scenario Reduction Model 89
Table 5.3 List of Notations and Decision Variables 91
Table 5.4 List of Cost Parameters.. 98
Table 5.5 Fraction of the added available capacity available during recovery time - $U_{j,i}$ 98
Table 5.6 Statistics on scenario reduction results 102
Table 5.7 Performance of Robust Formulations with Respect to Robustness Measures 109

Table C.1 Transition probabilities 134

Chapter 1

Introduction

1.1 Supply Disruptions

The number of natural and manmade catastrophes which disrupts the supply chain performance has increased dramatically in the last two decades. The strikes at two of General Motors parts plants in 1998 resulted in closure of 100 other plants, 26 assembly plants and shortage of cars in dealers for several months even after the strikes (Snyder et al. 2016). Ford stopped production in five plants because of the air traffic suspension after the terrorist attacks on September 11th 2001 (Tang 2007). The longshoremen strike at the LA docks in 2002 significantly impacted the availability of raw materials and products which were supplied from China and sold or consumed in United States (Vakharia and Yenipazarli, 2009). The manufacturing plants in Northeastern of United States lost their production capacity for several days due to the Blackout in August 2003 (Gonge et al. 2013). Such incidences represent the vulnerability of supply chains to rare but high profile disruptions.

The supply disruptions could have a drastic impact on supply chains missing protection against them. The Japan tsunami in 2011 interrupted Japanese automotive production, as well as automotive production companies all over the world dependent on Japanese suppliers; Toyota, Nissan and Honda closed their plants in Japan and General Motors suspended production in its assembly plant in USA due to the shortage of parts (Ghadge et al. 2011). Thailand's floods in 2011 lead to significant interruptions of the global computer hardware supply chain (Wai and Wongsurawat, 2012). Ericsson lost 400 million euros after a random lightning bolt struck its semiconductors supplier firm in New Mexico in 2000 (Tang 2007). Land Rover Discovery model's production line was shutdown for a period of nine months as a results of its exclusive chassis supplier's bankruptcy which also leads to the loss of 1500 jobs (Revilla and Saenz, 2017). Furthermore, the interruptions in supply flow may also occur by operational risks (Rezapour et al. 2018). The machine breakdowns in supplier's firm, slow shipments, customs delays, and quality defects are common types of operational risks associated with supply flow which can interrupt the production at manufacturing/assembly plant (Lakovou et al. 2010). Therefore, the vulnerability of supply chains to supply major disruptions and operational risks motivates academicians and practitioners in identifying appropriate risk management strategies which sustain the supply chain performance when such events happen.

Within the last two decades, there have been some examples where companies implement proactive or reactive risk management strategies in order to recover from disruptions. After the Hurricane Katrina in 2005, the Home Depot was able to satisfy the customers demand and kept its stores running thanks to emergency supplies which were stocked-up after lessons learned from past hurricanes. The Walmart started to stock-up its distribution centers and planned for alternate distribution network with Hurricane Katrina approaching. Therefore, after the hurricane struck, Walmart was able to quickly start delivering customers' orders and recover from disruption (Schmitt and Snyder, 2012). The proactive planning which both of these companies implement enable them to recover from Hurricane Katrina quickly.

The Nokia, a cell-phone manufacturer executes a different strategy to deal with shortage of supply which occurs due to a fire at its semiconductor chip supplier plant, Philips in New Mexico in 2000. When Nokia realized that Philips could not quickly recover from disruption, it switched its sourcing strategy to alternate suppliers which replace Philips. This proactive back-up supplier selection strategy enabled Nokia to recover from the loss of their primary supplier. However, Ericsson, the other customer of Philips who did not pre-plan for back-up supply options was impacted and it was exposed to market share losses for several months (Sheffi 2005).

The examples mentioned above present the value of pre-planning which enable the supply chain to quickly recover from disruptions. However, in order to recover from both operational risks and disruptions, cost-efficient risk management strategies should be employed in the supply chain structure based on the expected intensity and length of disruptions. These include strategic decisions which have long-term implications on supply chain cost and performance and they should be considered in the design stage of supply chain network (Klibi and Martel 2012).

1.2 Robust Supply Chain Network Design

The strategic decisions which are commonly considered in the supply chain network design include location and capacity levels of production and distribution facilities, selection of suppliers and third party contractors for logistic, warehousing, distribution and the location and level of strategic buffers such as safety stock (Klibi and Martel 2012). These strategic level design decisions are identified here and now but they impact the supply chain overall performance which includes operational costs, service level and revenue for several years.

The supply chain network is exposed to day to day activities such as procurement, production, warehousing, distribution, transportation and demand management which create the material flow across the network. Today’s supply chains are globalized and dispersed across the world due to business requirement which also make them vulnerable to uncertainties as described in section 1.1 (Baghalian et al. 2013). Therefore, in order to ensure the efficiency of the supply chain network under all circumstances in future, the factor of uncertainty should be considered in design stage of the supply chain.

The uncertainties in a supply chain network are characterized by price of material, labor, equipment, finished product and yield, lead time of suppliers, production, assembly plants and product demand, exchange rates and etc. In addition to these business-as-usual uncertainties, there are rare catastrophic events which might occur at any stage of supply chain. Such disasters will shut down the entire supply chain for generally a long period of time. Most recent examples of supply chain vulnerability to catastrophic events have been elaborated in section 1.1. Under uncertainty, there might be different levels and quality of information available (Klibi et al.

2010). The partial availability refers to the situation where there is sufficient data available to estimate the likelihood, length and/or intensity of a future event such as demand seasonality or the machine breakdown. On the other hand, there might be lack of any information available to estimate the attributes of a plausible event such as likelihood of an earthquake or a flood. This latter category is called deep uncertainty which has a drastic impact on supply chain performance since it is commonly not considered in the design stage of supply chain network.

The concept of robust supply chain network design has raised a lot of attention in decision making under uncertainty literature. The term robustness represents the efficient flexibility corresponding to a decision which provides many options for the selections to be made in the future (Wong and Rosenhead, 2000). Therefore, considering the flexibility in the strategic design of the supply chain network leads to the definition of a robust supply chain network design as follows. A supply chain network design can be stated as robust within the planning horizon if it has the capability to provide sustainable value creation under all plausible future scenarios which may include business as usual uncertainties with partial level of information available as well as catastrophic events with deep uncertainty (Klibi et al. 2010).

In order to design a robust supply chain network, the responsiveness and resilience mitigation strategies could be incorporated into the supply chain structure. The responsiveness strategies provide resources to protect the supply chain operations against frequent and low impact variations in supply flow, customer demand and production or transportation capacity levels. These strategies are embedded into the supply chain network beforehand. The most common responsiveness strategies include capacity buffers, safety stock pooling, flexible sourcing and subcontracting, overtime, product substitution, shipment rerouting (Tomlin 2006, Klibi et al. 2010, Chopra and Sodhi 2004, Sheffi 2005). The resilience strategies impact the supply chain structure by determining the level of resources in order to avoid disruptions and recover fast when disruptions occur. These strategic policies could also provide the capability for the efficient implementation of responsiveness strategies. The resilience strategies are provided as a result of investing in flexible and redundant network design. The flexibility based strategies are incorporated into the supply chain structure beforehand but they are deployed as needed. Some examples of flexibility based mitigation strategies include production systems with functionality to produce multiple products, partially interchangeable and scalable suppliers. The redundancy based strategies include having extra resources in the supply chain network in order to compensate for the disrupted resource(s). The excess capacity and safety stock are examples of redundancy based strategies. The difference between flexibility and redundancy based strategies is that the flexibility based strategies are determined to have least cost impact since the supply chain only incurs cost upon the deployment of these strategies. However the challenge with flexibility based strategies is the time which is required for the solution strategy to become fully operational. Therefore, the redundancy based strategies are efficient to cover business as usual uncertainties due to their immediate availability. Furthermore, the flexibility based strategies are least costly and most efficient to protect against disruptions which have low likelihood of occurrence but significant impact (Tomlin 2006). The flexibility and redundancy based mitigation strategies which are known to be efficient to protect the supply chain against supply uncertainties are presented next.

1.3 Supply Mitigation Strategies

The impact of supply uncertainties is not limited to the subsequent downstream stage or facility. The interruptions in supply flow impact the order on-time availability as it moves downstream from the impacted stage or site. This behavior represents the existence of reverse bullwhip effect which could be created as result of supply interruptions (Rong et al. 2009). In order to hedge the supply chain’s performance against supply uncertainties, the most common mitigation strategies are as follows (Snyder et al. 2016).

- Safety Stock: this is a redundancy based strategy which is deployed proactively. Furthermore, this extra inventory could be raw material kept in supplier site or finished good inventory piled in manufacturing firm. The safety stock mitigation is an efficient strategy to cover frequent and low impact operational risks. However, it is not sufficient to cover high impact and long disruptions. Furthermore, it is costly to hold inventory for a long time for disruptions that may never occur.

- Multiple Sourcing: this flexibility based strategy requires the firms to source raw materials from multiple suppliers. In case one supplier is disrupted, the firm only loses material flow from disrupted supplier and it still receives supply from non-disrupted suppliers. However, the order quantities received from non-disrupted suppliers do not change after disruption.
- Contingent Sourcing: this strategy is considered as an extension of multiple sourcing where the firm has multiple suppliers. In case one supplier is disrupted, the non-disrupted suppliers ramp up production in order to cover for the disrupted supplier. This capability is raised from other supplier’s volume flexibility however the challenge is in making the substitute supply available within a short response time (Tomlin and Wang, 2010).
- Acceptance: there are some cases where the cost of mitigation strategies exceeds the benefit associated with them. In such situations, the firm simply accepts the risk of disruptions and the resulting financial consequences.
- Demand Substitution: it might be possible to shift the demand to another available product when one product is out of stock because of a disruption. However, this strategy significantly depends on product’s market such competitor’s status, the phase in product life cycle in which disruption occurs.

In order to design a robust supply chain network with capability to sustain its functionality under operational risks and disruptions, the mix of strategic stock along with contingent sourcing is considered as an efficient strategy (Hopp and Yin 2006, Kouvelis and Li 2012). In this setting, the strategic stock can be utilized to cover operational risks and the contingent sourcing can be used as an effective reactive approach to cover major disruptions (Tomlin 2006, Schmitt 2011). However, the effectiveness of contingent sourcing depends on making the product available within a short response time (Tomlin and Wang, 2010). The response time is defined as the time when the firm responds to a supply disruption by placing an emergency capacity increase order with the backup supplier plus the time required for the backup supplier to provide the required capacity order (Tomlin 2006).

The response time is a crucial characteristic of contingent sourcing since only a fraction of the required capacity might be available within this period (Matta et al. 2007). In addition to this, shifting the demand to the back-up resource during the response time would create congestion that increases the lead time in that facility. This congestion is created as a result of randomness associated with parts arrival and back-up supplier production rate. Ignoring these facts in the supply chain planning stage leads to the overestimation of the available backup capacity, resulting in creating product shortage within the response time. This may also degrades the robustness of the supply chain.

The reduction in the response time can be achieved by making investment in scalable equipment which can quickly ramp up their capacities in small increments, whereas a supplier that is relying on dedicated equipment to reduce production cost will have a long response time (Putnik et al. 2013). While improving the response time can be similar to reducing the mean time to repair (MTTR) (Hopp and Iravani, 2012), it is also critical that the backup supplier provides an appropriate level of capacity during the response time. This is critical mainly due to the loss of market share during this period, creating significant long-term implications for the firm (Hendricks and Singhal, 2005). The amount of the available capacity during the response time depends on response speed defined as the speed of the backup supplier to reach the desired capacity level (Niroomand et al. 2012). The layout configuration of the backup supplier is one of the main factors identifying the response speed level. Therefore, a strategy to improve the available capacity within the response time can be achieved through the backup supplier's investment in layout configuration.

The strategic stock could also be used at the beginning of a major disruption and during the response time to keep the supply chain running until the back-up material can be received (Schmitt 2011). However, the required level of strategic stock depends on available capacity of the back-up supplier during the response time. Furthermore, the available capacity of back-up supplier during the response time depends on level of response speed. Therefore, the level of strategic stock and response speed of back-up supplier are key strategic level supply chain design decisions that may significantly impact the operation costs in the future. In order to design a supply chain network with robustness against supply uncertainties, the optimal level of strategic stock and response speed of back-up supplier should be identified in the design stage. Furthermore, an accurate estimate of the backup supplier's capacity during the response time and specifically the impact of congestion over capacity should be considered in design stage in order to prevent potential product shortages in future. These requirements shape the objectives of this thesis and they are summarized in following section.

1.4 Scope and Objectives

The main scope of this thesis is to present the decision makers with a tool to design a robust supply chain subject to operational risks and disruptions of supply network. The list of the specific contributions is summarized as follows.

- To present an approach to determine the available capacity of back-up supplier during the response time by considering the congestion impact.
- To determine the response speed of back-up supplier and the level of strategic stock in order to achieve a robust supply chain network subject to supply uncertainties and random capacity levels during response time when partial information is available.
- To determine the response speed of back-up supplier and the level of strategic stock in order to achieve a robust supply chain network subject to supply uncertainties and random capacity levels during response time when there is deep uncertainty about the future business environment.

1.5 Thesis Outline

This manuscript is organized as follows. Chapter 2 reviews the relevant literature in Supply Chain Network Design (SCND), Risk Mitigation Strategies, Robustness, and Scenario Reduction. Chapter 3 presents the modeling assumptions and the mathematical formulation of a deterministic mixed-integer programming (MIP) based capacity planning model which also include the non-linear clearing function in order to represent the congestion impact.

Chapter 4 provides a stochastic optimization MIP based capacity planning model in which the response speed level of back-up supplier and the level of strategic stock are first stage decision variables and the major and minor disruptions along with level of capacity available during the response time are considered as random parameters. The scenario tree approach to generate original scenario set is also presented.

Chapter 5 considers the situation where there is a deep uncertainty about random parameters and it presents a robust optimization MIP based capacity planning model. Furthermore, the Ordered Weighted Averaging (OWA) Aggregator operator is incorporated in the objective function of the robust model in order to achieve a fair solution with respect to all plausible scenarios. The computational complexity of this problem is reduced by presenting a novel clustering based MIP scenario reduction model. This model also includes the gradual coverage function of facility location problems in order to improve the computational time. The thesis ends with conclusions and future research directions in Chapter 6.

Chapter 2

Literature Review

In this Chapter we review the relevant literature to this thesis. More specifically, in Section 2.1 we review the concept and application of contingent sourcing as a risk mitigation strategy. Furthermore, we discuss the requirement to consider the operational characteristics of contingent sourcing in the design stage of supply chain. In Section 2.2, we explain the concept of robust supply chain design and the application of solution robustness as a benchmark to measure the robustness of a supply chain design. We also review the methodologies to compute solution robustness with respect to different levels of uncertainty. Since the supply chain design problems with deep level of uncertainty are computationally intractable, we review the literature on scenario reduction methodologies in Section 2.3. This chapter ends with a summary of gaps which exist in the reviewed literature.

2.1 Contingent Sourcing

Sethi and Sethi (1990) define flexibility as the capability of changing in order to deal with a changing environment. They categorize flexibility into two groups called mix flexibility and volume flexibility. The mix flexibility is defined as the capability to produce multiple products. The mix flexibility has been incorporated into the design of manufacturing systems and it inspires the development of Flexible Manufacturing Systems (FMS). The examples of FMS are computer numerically automated-CNC machines which can produce a variety of products with low change over time (Mehrabi et al. 2000). On the other hand, the volume flexibility is defined as the ability to alter the production capacity of a manufacturing process in order to meet the demand requirements (Hallgren and Olhager, 2009). The idea of volume flexibility leads to the development of Reconfigurable Manufacturing Systems (RMS) which has a modular configuration enabling them to change their production capacity by adding or removing modules (Koren et al. 1999). An industrial application of RMS is presented in Deif and ElMaraghy (2006) in electronics industry. Other examples in metal machining and assembly systems can be seen in Koren et al. (1999).

The concept of volume flexibility could be part of supply chain risk mitigation strategies to deal with supply uncertainties. In this setting, the supply network includes multiple suppliers where the suppliers have volume flexibility as a built in technology. In the case of disruption in any of suppliers, the non-disrupted suppliers will act as back-up and they increase their production capacity to cover for the disrupted supplier(s). This strategy is also known as contingent sourcing (Snyder et al. 2016).

There exists a growing body of literature which incorporates the contingent sourcing along with strategic stock in order to mitigate the impact of supply disruptions. Kouvelis and Li (2012) evaluate the value of safety stock, safety lead time and emergency back-up in managing uncertain supply lead-time. The emergency back-up is assumed to consist of price fluctuating suppliers in which the required capacity is available instantaneously. They conclude that the effectiveness of emergency back-up increases in randomness associated with original order lead-time. Qi (2013) studies a supply chain with one retailer and two suppliers which include an unreliable primary and a reliable but more expensive back-up. The back-up supplier is assumed to be able to provide the requested capacity immediately, similar to Kouvelis and Li (2012). Qi (2013) identifies the optimal waiting time of the retailer before switching to the back-up supplier in the case of primary supplier breakdown. Although these papers assume the instant availability of supply flow as the firm places the capacity increase order with back-up supplier, the Table 2.1 presents other works which consider the back-up supply flow to become available after a delay.

Table 2.1 Summary of the reviewed supply chain design problems with contingent sourcing

Author(s)	Strategic Design Decision(s)	Objective(s)	Back-up Availability
Kouvelis and Li (2012)	Safety stock, Safety lead time, Time and Size of back-up order	Total cost minimization	Immediate back-up
Qi (2013)	Cap waiting time, Safety stock	Total cost minimization	Immediate back-up
Bundschuh et al (2003)	Suppliers selection	Improving reliability and robustness	After response time
Bilsel and Ravindran (2011)	Primary and back-up suppliers selection and order allocation	Total cost and total lead-time minimization, Maximizing total quality of products	After response time
Fang et al (2012)	Optimal sourcing strategy	Total cost minimization	After response time
Hopp and Yin (2006)	Safety stock and back-up capacity locations	Total Cost minimization	After response time
Schmitt and Singh (2012)	Safety stock location, Back-up response type	Total cost minimization, Target service level	After response time

2.1.1 Response Time

Bundschuh et al. (2003) present different models for strategic design of robust and reliable supply chain. In the robust model, the supply chain could have contingency supply after a pre-determined lead time when disruptions occur. This is provided through extra supply of remaining suppliers in addition to their regular contractual supply. Bilsel and Ravindran (2011) develop a multi-objective stochastic supplier selection and order allocation model with randomness in demand, supplier's capacity and costs. They assume that the back-up supplier might require a positive lead time before supplying the required service level. Furthermore, Bilsel and Ravindran (2011) demonstrate the value of the solutions achieved by stochastic model compared to the deterministic counterpart in their problem. Fang et al. (2012) propose a dynamic programming formulation to select the optimal sourcing strategy for different risk profile settings. They assume the back-up supplier to have unlimited capacity which would be available after response time. The authors present the companies optimal sourcing strategy between dual and contingent sourcing where the latter is promising if the back-up supplier has short lead-time.

Hopp and Yin (2006) try to find the optimal placement of the inventory and/or back-up capacity to protect a multi-echelon supply network in the case of catastrophic failures. They conclude that the inventory or back-up capacity should be provided at most in one node along each path to the customer. The location of inventory is strongly affected by the response time since it can improve the product availability during this period. Schmitt and Singh (2012) determine the location of the safety stock and response type of the back-up resource in a multi-echelon supply chain where disruptions could occur at any stage. They assume the back-up supplier to have limited capacity and provide the disrupted capacity partially after a certain period called response time. The results show that finished goods inventory increase service level significantly. In addition to this, it is better to have quick and small response as the probability of the upstream disruptions increase. While, the cited papers above assume the supply capacity to be entirely available after the response time, there are a few papers in literature which study the strategies to reduce the response time.

Schmitt (2011) studies the optimal selection of the response speed of the back-up resource in a multi-echelon supply chain where disruption might happen at any stage in order to protect a predetermined service level under all plausible future scenarios. Although Schmitt (2011) assumes that the firm can make investments in equipment to improve the response speed but she does not explain those investments explicitly.

Wang and Koren (2012) identify machine configuration as a parameter which affects the response speed of manufacturing systems. In a serial configuration, the response speed is slow since the added capacity can only become available after completing the capacity installation and ramp-up phases of all stages. On the other hand, the pure parallel configuration provides faster response speed level because each machine could go under the aforementioned phases independently. However, this may come at the higher investment cost since each machine should be capable of performing all the steps in order to create a parallel configuration. The differences in the manufacturing system configuration indicate that the configuration selection of the back-up resource affects directly the response speed and it should be considered in the supply chain design stage where the trade-off is between cost and response speed.

2.1.2 Response Time Characteristics

The back-up resource may provide some portion of the supply order during the response time. Klibi and Martel (2012) consider the partial availability of the capacity of a depot during the recovery period. They propose a discrete stepwise function to represent the gradual capacity recovery of the disrupted depot based on the intensity of the disruption and the time to recovery. Niroomand et al. (2012) illustrate the partial availability of the capacity within the response time in a strategic capacity planning model. The authors consider a two-echelon supply chain where the production stage includes a dedicated manufacturing system (DMS) and a reconfigurable manufacturing system (RMS) as a volume-flexible backup resource achieved through reconfiguration. The reconfiguration process refers to capacity installation and ramp up/down phases. The model incorporates a partial availability of the RMS capacity during the ramp-up phase to better represent the modular structure of the RMS. However, these models ignore a critical aspect in the level of material flow which is originated in the back-up resource during the response time.

2.1.3 The Impacts of Congestion on Throughput during the Response Time

In a situation where the main resource is disrupted, its demand would be transferred to the backup resource under a contingency strategy. This may create an overload of demand at the backup resource due to the randomness in production capacity during the response time. This randomness is a result of frequent occurrences of the system breakdown, rework, scrap and low skill of the operator to work with new configuration during the response time (Matta et al. 2007). As a result of this overload, queues will build up, degrading performance due to the congestion. Ignoring this fact may result in creating products shortages during the response time along with negative financial impacts (Pahl et al. 2007).

In order to consider the impact of congestion on the system throughput, Kim and Uzsoy (2008) employ the clearing function in a multi-work center capacity expansion problem. The clearing function is initially introduced by Karmarkar (1989) and presents the expected throughput of a resource over a planning period as a function of the expected Work in Process (WIP). The majority of studies that use clearing functions are in the production planning field. There are a few studies in risk management which consider the impact of congestion. Vidyarthi et al. (2009) propose a stochastic capacity planning model for a two-echelon supply chain that includes distribution centers and customers under random demand arrivals. The model's objective is to minimize the lead time and capacity expansion costs, while the relationship between the lead time and congestion is captured through queuing models. Even though Vidyarthi et al. (2009) represents the congestion effect in a capacity planning model under the common problems of matching supply and demand, this phenomenon so far is ignored in the literature that focuses on the management of major disruptions (Hopp and Yin 2006, Schmitt 2011).

2.2 Robustness

Recently, the concept of robustness has drawn a lot of attention in the literature with focus on decision making under uncertainty. The robustness concept could have different meanings based on the decision making context in which is it applied to (Roy 2010). While the model robustness measures the solution feasibility, the solution robustness measures the performance of solution with respect to the optimal solution of each scenario (Mulvey et al. 1995). This thesis is focused on solution robustness or more specifically supply chain network design robustness. In the supply chain management literature, the term robustness is defined as the extent to which the supply chain is able to carry its functions for a variety of plausible scenarios (Snyder and Daskin, 2006). Furthermore, a supply chain design is identified as robust if it has capability to sustain value creation under operational risks and major disruptions in future (Klibi et al. 2010, Wieland, 2013).

2.2.1 Solution Robustness

In order to evaluate the robustness of a supply chain design with respect to future disruption occurrences, the concept of solution robustness has been applied as a performance measure in several papers. A solution is called robust if it remains close to optimal for any occurrence of scenarios (Sahebjamnia et al. 2018). Baghalian et al (2013) identify the optimal location of facilities subject to supply side disruptions by minimizing the trade-off between expected cost of supply chain and solution robustness. They compute the solution robustness based on the difference between the cost of each scenario and expected cost of all scenarios. This approach has been frequently applied in supply chain network design problems in order to compute the solution robustness (Sadghiani et al. 2015, Rouzhen and Wang 2016, Nooraie and Parast 2016, Joonrak et al. 2018).

In the same context of problems, there have been other methodologies incorporated into the problem's formulation in order to measure solution robustness. Snyder and Daskin (2006) p -robust formulation minimizes the expected cost of the supply chain while bounding the relative regret in each scenario to be lower than the constant p. They compute the regret of a solution in a given scenario as the difference between the cost of the solution in that scenario and the cost of the optimal solution for that scenario. Sawik (2014) develop a combinatorial stochastic optimization formulation in order to identify robust solutions in a supplier selection and demand allocation problem subject to supplier disruptions. He tries to minimize the ordered weighted averaging aggregation of the expected value and the expected worst-case value of the objective function in order to obtain an equitably efficient solution. Such a solution is expected to equitably optimize the performance of a supply chain with respect to all plausible scenarios as well as in the worst-case scenario.

2.2.2 Solution Robustness under Deep Uncertainty

All approaches cited above could be applied into the supply chain network design problems in order to achieve solution robustness when the probabilities of scenarios are available. However, there is an open challenge for the case in which the scenario probabilities are unavailable called deep uncertainty (Klibi et al. 2010). Terrorist attacks, epidemics, geo-political instability, extreme weather events due to the climate change and related natural catastrophes are typical examples which are typically rare and hard to predict (Heckmann et al.2015). The Minimization of maximum cost or regret (absolute or relative) is one of the most common approaches to achieve solution robustness when scenario probabilities are not available (Kouvelis and Yu, 2013). However it is known to be too pessimistic because of considering only the worst case scenario.

In order to compute solutions with less level of conservatism, Roy (2010) presents a new robustness formulation called bw-robustness which not only provides a solution that guarantees an objective value of at least w across all scenarios but also maximize the probability of reaching a target value of b ($b > w$). This approach holds great appeal for managers due to its simplicity however the results are limited to the range provided by two boundary values of w and b. Kalai et al. (2012) propose another robustness approach called lexicographic α-robustness which minimizes not only the maximum cost but also the second largest cost, the third one and so on with respect to a given threshold called α. This methodology is considered as a combination of Minmax and p-robust formulation described earlier. Furthermore, the lexicographic α-robustness formulation compensates for the conservatism of Minmax formulation by reordering the performance vector e.g. cost from the worst to the best and identify the robust solution such that the reordered performance vector is close to a given threshold. Although the lexicographic α-robustness approach is known to provide fair solutions with respect to objectives considered, but it has equivalent computational complexity to Minmax formulation especially for problems with large number of scenarios (Kalai et al. 2012).

2.3 Scenario Reduction

In order to reduce the computational complexity in stochastic programming and robust optimization associated with large number of scenarios within the original scenario set, one solution is to develop a reduced scenario set of the original set by selecting a few representative scenarios. This approach is called scenario reduction in literature. There are different scenario reduction techniques in literature including backward reduction and forward selection heuristics developed by Dupacova et al (2003), k-means clustering algorithm (Sutiene et al. 2010) and the probabilistic distance based reduction methodologies (Zeballos et al. 2014, Li et al. 2014).

The forward selection heuristic determines scenarios that will not be eliminated in a recursive manner. The scenarios with the minimum sum of the distances to the unselected scenarios are preserved. The termination condition is a specified number of scenarios that has to be preserved. The backward reduction heuristics has an inverse mechanism with the objective set to identify scenarios that have to be deleted. Both of these approaches have been frequently applied in supply chain network design literature (Govindan and Fattahi 2017, Esmaeili et al. 2016, Hamta et al. 2017). Furthermore, Dupacova et al. (2003) states that the reduced scenario sets determined by forward selection heuristic are slightly better with respect to accuracy however the computation requires higher CPU time.

Heitsch and Romisch (2003) propose new versions of forward selection and backward reduction algorithms presented by Dupacova et al. (2003). The major differences include considering all deleted scenarios into each backward step of backward reduction algorithm simultaneously and also assigning identical weights to each scenario in the objective function of optimization model in order to simplify the forward selection processes. The new algorithms are called fast forward selection and simultaneous backward reduction. When comparing accuracy, Heitsch and Romisch (2003) results show that fast forward selection algorithm has best performance. Furthermore, the simultaneous backward reduction algorithm also provides more accurate solutions compared to backward reduction algorithm of Dupacova et al. (2003) but at the expense of higher computational times.

Sutiene et al. (2010) develop a new clustering approach called k -means clustering which group data points into clusters such that each data point is in the cluster whose mean is closest. Khatami et al. (2015) utilize this methodology to reduce the size of the scenario set prior to applying Benders' decomposition to solve their closed-loop supply chain network design problem. Crainic et al. (2014) use k -means clustering to create multi-scenario sub-problems. Applying a progressive hedging-based meta-heuristic to solve sub-problems, the results show that the quality of solutions is improved compared to the case where heuristic is applied to single-scenario sub-problem. Furthermore, the time complexity is proved to be linear with respect to the number of scenarios.

Zeballos et al. (2014) apply a reduction algorithm based on the probability distance metric to their multi-period, multi-product, closed-loop supply chain (CLSC) design problem which is subject to uncertain levels in the amount of raw material supplies and customer demand. The probability distance is a function of scenario probabilities and the distances between scenario values. Therefore, the reduction algorithm deletes scenarios when they are close or have small probabilities. Finally, a sub-set of the original scenario set is achieved which include preserved scenarios with new probabilities. The preserved scenarios represent the deleted scenarios and their new probabilities are the summation of their probabilities in the original scenario set plus the probabilities of scenarios which are represented by them. Furthermore, the reduction algorithm Zeballos et al. (2014) applied to their CLSC problem can be found in the library SCENRED of GAMS. Their results show the importance of using a reduction algorithm to decrease the size of the problem, considering several outcomes at each time period for each uncertain parameter.

Li et al. (2014) propose a new scenario reduction approach which minimizes not only the probabilistic distance between the distributions of the original scenario set and the reduced distribution of selected scenarios but also the difference between the best, worst and expected performance. To the best of our knowledge, this approach is the only MIP optimization based scenario reduction methodology available in literature. Li et al. (2014) results show that their approach has a better performance compared to GAMS scenario reduction routine SCENRED2. However, this method is constrained by the size of the problem such that it cannot compute the reduced scenario set for problems with large number of scenarios in an efficient manner. In order to address this limitation, Li and Floudas (2016) develop a sequential scenario reduction framework for problems with multiple uncertain parameters. First, the scenario set is decomposed into multiple subsets where each subset is created based on a single uncertain parameter. Next, the single stage scenario reduction approach proposed by Li et al. (2014) is applied to each subset. Finally, the selected scenarios correspond to each subset are included in the reduced scenario set. Li and Floudas (2016) results verify the efficiency of the proposed decomposition based approach in solving large scale problems generated from multiple uncertain parameters. However, this approach is not applicable to large scale scenario sets developed based on a single uncertain parameter since such problems are not decomposable.

2.4 Conclusion

In the first part of literature review, we focus on the works which apply contingent sourcing as part of their risk mitigation strategies to deal with supply disruptions. Considering the response time as a crucial parameter in the successful implementation of contingent sourcing, we review the literature on strategies to reduce the response time. The investment in the back-up supplier configuration helps to increase the response speed however the efficiency of this strategy depends on the consideration of available capacity levels during response time. There are works in the literature such as Klibi and Martel (2012), Niroomand et al. (2012) which focus on estimation of the partial capacity available during the response time but the impact of the congestion created as result of randomness in production rate of back-up supplier over system throughput is ignored so far. This may result in overestimating the back-up supplier capacity during the response time and therefore creating product shortages. We resolve this issue in Chapter 3 by incorporating the clearing functions into the contingency capacity planning model.

The selections of the optimal level of back-up supplier's response speed and strategic stock are our design problems; therefore we focus on the robust design of supply chain in the second part of our literature review. Furthermore, we model the randomness associated with available capacity during the response time in a robust optimization model in Chapter 4. Former studies such as Bundshuh et al. (2003), Bilsel and Ravindran (2011), Fang et al. (2012) and Schmitt and Singh (2012) ignore this fact in their analysis. This may result in an inaccurate representation of the production capacity during the response time.

The concept of solution robustness could be applied to measure the robustness of a supply chain design solution. Therefore, our objective is to achieve solution robustness in identifying our design decisions. In order to achieve solution robustness, there are two streams of research which consider different levels of information that might be available about uncertain parameters. The first stream assumes there is enough information available to estimate the probabilities of disruptions. The common approaches to achieve solution robustness are minimization of the trade-off between expected cost and solution robustness (Baghalian et al. 2013) and p -robust formulation (Snyder and Daskin, 2006). In Chapter 4, we assume that scenario probabilities are available and we apply an approach similar to Baghalian et al. (2013) to achieve solution robustness.

The second stream represents the situation where it is not possible to estimate the probabilities of disruptions. In this case, the lexicographic α-robustness formulation proposed by Kalai et al. (2012) could be applied in order to achieve fair solutions with respect to all plausible scenarios. However, there is computational complexity associated with this approach especially for problems with large number of scenarios.

The scenario reduction is known as an efficient approach to reduce the computational complexity in stochastic programming and robust optimization when scenario sets are large. The review of literature on the most well-known scenario reduction techniques reveals that the MIP optimization models calculate the reduced sets with better solution quality compared to heuristic based techniques (Li et al. 2014) however they cannot compute the reduced set for problems with large number of scenarios in a reasonable time. Therefore, we propose a novel clustering based MIP optimization scenario reduction model which includes the gradual coverage function in order to improve the computational time in Chapter 5. The computed reduced sets are then used in a robust optimization model which includes an Ordered Weighted Averaging aggregator operator in its objective function in order to achieve fair solutions when scenarios probabilities are not available.

Chapter 3

Responsive contingency planning in supply risk management by considering congestion effects

In this Chapter, we focus on contingent sourcing as a cost-effective risk management strategy to deal with major supply disruptions. In order to improve the supply chain responsiveness, our objective is to determine the appropriate response speed level of the volume-flexible backup supplier. To this end, we develop a decision-making tool which considers the operational characteristics of contingent sourcing such as response time and congestion impacts in order to make an accurate decision. We evaluate the impact of the different failure and recovery probabilities over the selection process. Furthermore, we investigate whether it is important to consider the congestion effects in the supply chain strategic level design decisions.

We consider a single product supply chain that includes a warehouse with dual sourcing as presented in Figure 3.1. The main supplier is cost-effective as a result of dedicated facilities (DMS) but prone to disruptions. It could be up or down completely for an integer number of periods within the planning horizon. Similar to the Nokia and Chiquita's suppliers (Tomlin 2006), we assume that there is a backup supplier located in a low-risk region that is available when the main supplier is disrupted. The backup supplier has volume-flexible production facilities where it can scale up its capacity according to a speed related to its configuration. Furthermore, the production cost of the backup supplier P_r is higher than the main supplier's production cost P_d due to its scalability.

The supply chain is analyzed in a long-term planning horizon T (multiple years). This assumption is made because of the fact that the planning horizon should be longer than the recovery period of any disruption scenario (Tomlin 2006, Schmitt 2011). Furthermore, each period t represents a quarter. The product demand is deterministic and follows the classical lifecycle pattern which includes introduction, growth, maturity and decline phases (Rink and Swan, 1979). It is assumed that the product demand is not affected by the disruption since the main supplier is not located in the demand region. Demand in any period, D_t, must be met by the main and backup supplier. If the demand is not met within its period, it is considered as lost, represented by s_t.

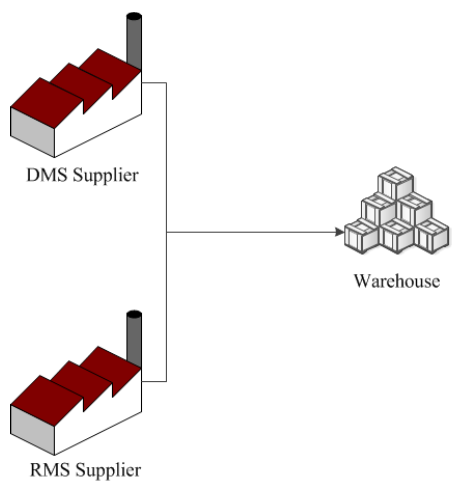


Figure 3.1 The supply chain network configuration

As illustrated in Figure 3.2, the main supplier provides the required supply up to a maximum level of C_d during normal periods. The raw material $\rho_{d,t}$ would be released into the DMS facility at the beginning of the period, which results in the production throughput of $x_{d,t}$. Due to the queuing effects, there would exist a work in process inventory, $\omega_{d,t}$. Furthermore, the inventory level at the end of the period is represented by $v_{d,t}$.

If the main supplier which is equipped with DMS fails due to a major disruption, the scalability of the back-up supplier with RMS is being used to supply the required flow of material. Therefore, the backup supplier increases its capacity to meet the warehouse demand. The time and the magnitude Δ_t^+ of these changes are decided in a multi-period contingency capacity planning model with respect to the trade-off between shortage cost S, RMS reconfiguration cost R, RMS excess capacity cost E_r and RMS production cost P_r, as described in section 3.1.1.

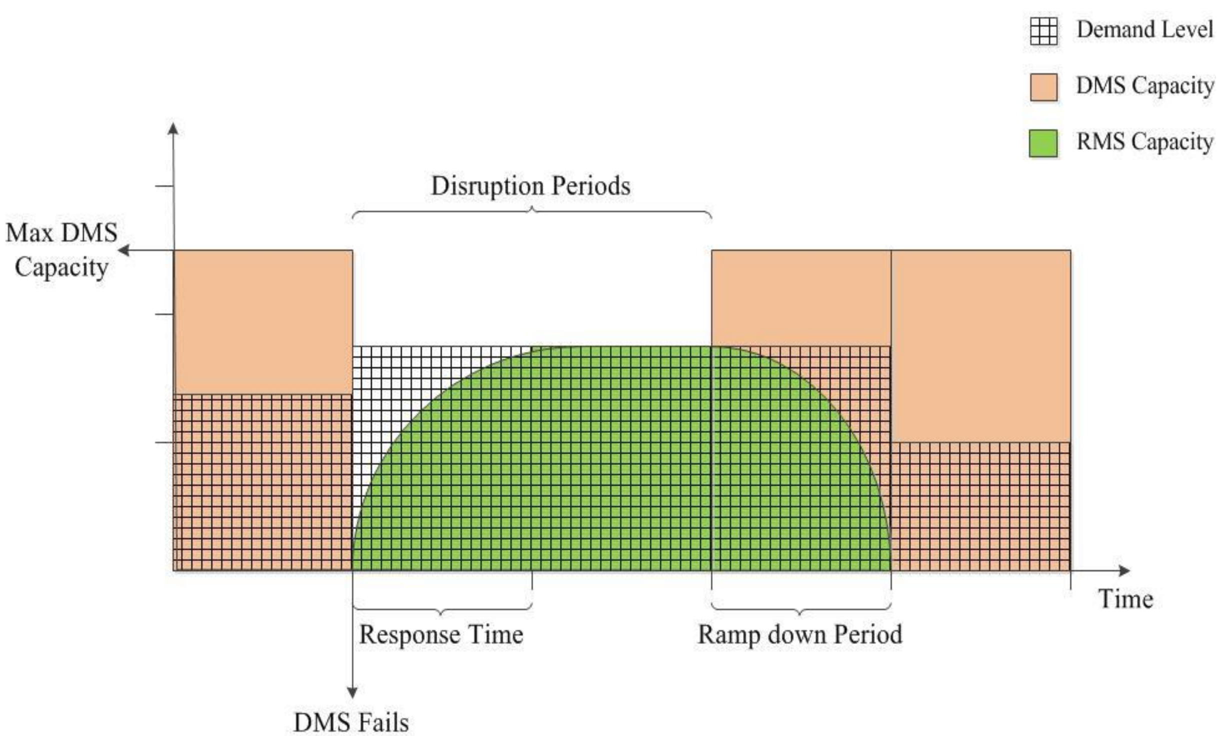


Figure 3.2 An example of a contingency plan execution.

The RMS consists of machines having a base structure on which modules can be added to increase the capacity. Each module can increase the capacity with discrete steps C, where smaller increments mean better scalability. Furthermore, there is an upper limit f on the number of modules that can be attached to a base due to machine or space limitations (Jordan and Grave, 1995). The target capacity ξ_t would be gradually achieved within the response time due to the reconfiguration process (Koren et al. 1999). Therefore a fraction of the target capacity represented by τ_t is available during the response period. Since the backup supplier is not fully capable of producing at the required rate during the response time, shifting the demand to the backup supplier creates an overflow of demand, resulting in congestion. This congestion would decrease the throughput during the response time due to the increase in the lead time.

In order to implement a responsive contingency planning, operational characteristics such as response time should be considered in the design stage (Tomlin et al. 2010). Otherwise, the supply chain may incur shortages due to overestimation of the production capacity. For this purpose, we model the available capacity during the response time and the impact of the congestion over the system's throughput, described in sections 3.1.1.1 and 3.1.1.2 respectively.

As indicated in Figure 3.2, the backup facility reaches the desired level at the end of the response time. Furthermore, the backup supplier ramps down to its initial capacity at the end of the disruption, when the main supplier resumes supplying the product to the warehouse. The amount of the available capacity during the response time depends on the response speed such that a faster response speed provides more capacity within the response time. Furthermore, the response speed depends on the RMS layout configuration (Hale and Moberg, 2005).

A parallel configuration leads to a faster response speed compared to a serial configuration. A configuration with mostly parallel machines will increase the available capacity during the response time at the expense of increased reconfiguration cost. On the other hand, the RMS capacity within the response time is important, since the supply chain incurs shortage costs if the available capacity level during this period is lower than the required capacity. As a result of these factors, RMS layout configuration should be determined at the design stage of the supply chain in order to minimize the expected costs. Note that the selected configurations would remain fixed during the planning horizon while the capacity might change upon the realization of the different disruption scenarios.

In this Chapter, we identify a scenario as the status of the main supplier within the planning horizon. It could be active or inactive due to the disruption occurrences. With respect to the frequency and the length of disruptions within a specified planning horizon, several scenarios can be identified. Since the focus is on the rare catastrophic events, we determine the scenarios where disruptions occur once within the planning horizon. Therefore, disruption scenarios are generated with respect to the time of occurrence m and the length of the disruption n.

In the following section, we present a two-stage solution methodology to determine the appropriate response speed of the backup supplier. For a given response speed, the contingency plans corresponding to the disruption scenarios are generated in the first stage. This is repeated for various speed levels. In the second stage, the appropriate speed level is selected in a decision tree analysis.

3.1 Solution methodology

In order to find the optimal response speed of the backup supplier, a solution methodology based on mixed integer programming and decision tree analysis is proposed as illustrated in Figure 3.3. We first develop a mixed integer programming (MIP)-based multi-period capacity planning model for a deterministic demand within the planning horizon. Afterwards, the capacity plan is subjected to a set of possible DMS disruption scenarios, where each scenario's probability of occurrence is calculated using discrete Markov chain distribution.

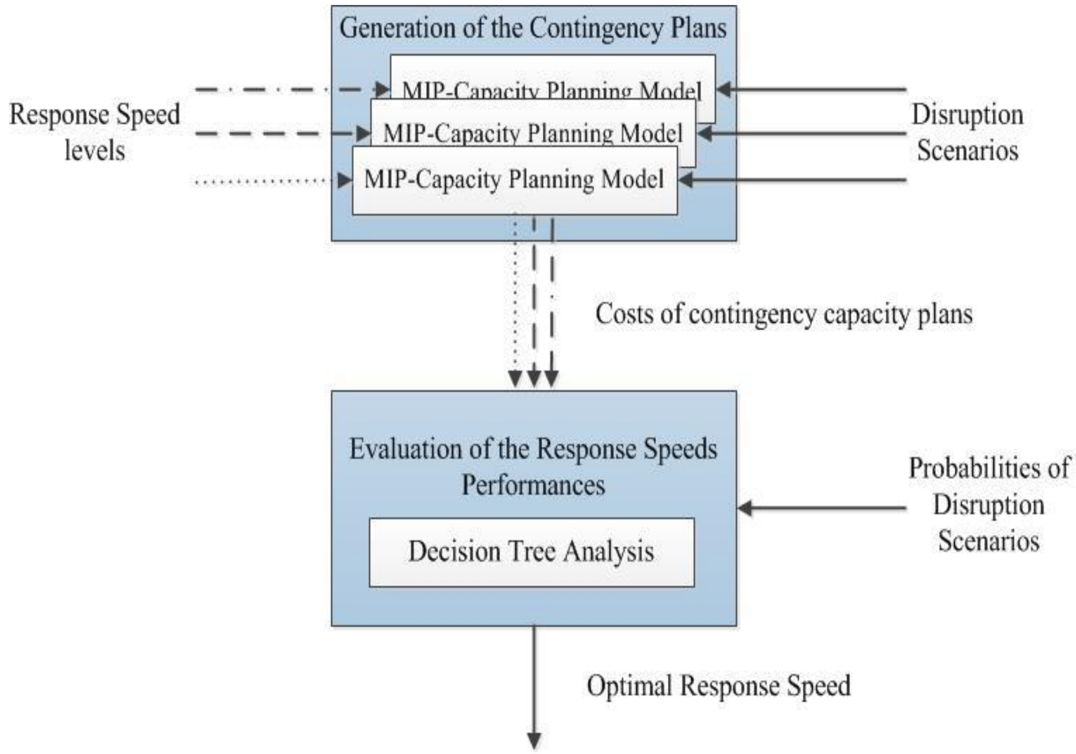


Figure 3.3 Solution methodology.

Each disruption scenario is inserted to the MIP model to represent the capacity disruptions to the DMS facility, which in turn will trigger the need for the RMS to ramp up its supply capacity to meet the demand. We analyze three different response speed levels; each corresponds to a certain level of available capacity during the response time. For each speed level, the MIP model generates the contingency capacity plans and their resulting costs corresponding to different disruption scenarios. The costs of the contingency capacity plans as well as the probabilities of disruption scenarios are then incorporated in a decision tree. For a given failure and recovery probability, the optimal response speed under all plausible future scenarios is selected through this decision tree analysis.

Since the selection of the response speed can depend on the attitude of the decision maker towards risk, we can determine the optimal policy under risk-neutral and risk-averse conditions. The risk-neutral decision maker selects the optimal response speed with the objective of minimizing the expected cost under all plausible scenarios, while the risk-averse decision maker would like to minimize the expected cost of worst-case scenarios. The worst-case scenarios are identified as the scenarios which have the highest operational cost according to the objective function.

We first introduce the MIP capacity planning model in section 3.1.1. More detailed explanations are provided for the constraints representing the available capacity during the response time and congestion effects in sections 3.1.1.1, and 3.1.1.2. The modifications required to generate the contingency plans are given in 3.1.2. Finally, the decision tree analysis is presented to provide details on the second phase of the proposed methodology.

3.1.1 Capacity planning model

The first step of the proposed methodology consists of the mixed integer programming model to determine the capacity, production, inventory and WIP levels of DMS and RMS suppliers for a predetermined planning horizon. The list of notations and decision variables is shown Table 3.1 and

Table 3.2.

Table 3.1 List of Notations

	Indices
t	Current time
d	DMS supplier
r	RMS supplier
i	Number of added or removed modules
j	RMS nominal capacity level
	Input parameters
$T = \{1, 2, ..., T'\}$	Planning horizon consisting of T' periods
M	A big number
D_t	Demand at time t
P_d	Production cost of DMS

P_r	Production cost of RMS
R	Reconfiguration cost
S	Shortage cost
E_d	Excess capacity cost of DMS
E_r	Excess capacity cost of RMS
H_d	Finished good holding cost of DMS
H_r	Finished good holding cost of RMS
W_d	Work in Process holding cost of DMS
W_r	Work in Process holding cost of RMS
M_d	Release material cost of DMS
M_r	Release material cost of RMS
C_d	Maximum DMS capacity
C	RMS Module capacity
U_i	The coefficient of the upper limit for actual RMS capacity change
L_i	The coefficient of the lower limit for actual RMS capacity change
$I = \{0,1,2,...,f\}$	Set of number of modules that could be added or removed with maximum of f
$J = \{C,2C,...,fC\}$	Set of RMS nominal capacity levels

Table 3.2 List of decision variables

Decision variables	
$x_{d,t}$	DMS production at time t
$x_{r,t}$	RMS production at time t

Δ_t^+	RMS added capacity at time t
Δ_t^-	RMS removed capacity at t
s_t	Lost demand at time t
$\pi_{d,t}$	DMS satisfied demand at time t
$\pi_{r,t}$	RMS satisfied demand at time t
$\varepsilon_{d,t}$	DMS excess capacity at time t
$\varepsilon_{r,t}$	RMS excess capacity at time t
$v_{d,t}$	DMS finished good inventory at time t
$v_{r,t}$	RMS finished good inventory at time t
$\omega_{d,t}$	DMS Work in Process at time t
$\omega_{r,t}$	RMS Work in Process at time t
$\rho_{d,t}$	Amount of raw material released to DMS at time t
$\rho_{r,t}$	Amount of raw material released to RMS at time t
τ_t	RMS actual capacity at time t
ξ_t	RMS nominal capacity at time t
u_t	RMS actual capacity added at time t
l_t	RMS actual capacity removed at time t
$y_{i,t}$	1 if i modules are added to RMS; 0 otherwise
$k_{j,t}$	1 if nominal capacity j is reached; 0 otherwise
b_t	1 if demand loss exists; 0 otherwise
q_t	1 if there is capacity addition; 0 otherwise

The objective function (1) includes the production cost, the system reconfiguration cost, the lost demand and the excess capacity costs, the holding cost of the finished good inventory, the work in process holding cost and the raw material purchasing cost.

$$\begin{aligned} \textit{Minimize } Z = \sum_{t \in T} \{ & P_d x_{d,t} + P_r x_{r,t} + R(\Delta_t^+ + \Delta_t^-) + S s_t + E_d \varepsilon_{d,t} + E_r \varepsilon_{r,t} + H_d v_{d,t} + H_r v_{r,t} \\ & + W_d \omega_{d,t} + W_r \omega_{r,t} + M_d \rho_{d,t} + M_r \rho_{r,t} \} \end{aligned} \quad (1)$$

After the demand is realized for a period, it could be satisfied through the inventory or the current RMS and DMS production (2), (3). The unsatisfied demand is lost and it is not carried over to the next period (4). We assume that it is not possible to have both demand loss and the inventory at the end of any period (5), (6). The work in process inventories consist of the jobs in the queue or under operation. Constraints (7) and (8) represent the balance equations between the raw material release, production quantity and WIP level for each period.

Constraints:

$$v_{d,t} = v_{d,t-1} + x_{d,t} - \pi_{d,t} \qquad \forall\, t \in T \qquad (2)$$

$$v_{r,t} = v_{r,t-1} + x_{r,t} - \pi_{r,t} \qquad \forall\, t \in T \qquad (3)$$

$$\pi_{d,t} + \pi_{r,t} + s_t = D_t \qquad \forall\, t \in T \qquad (4)$$

$$s_t \le M(1 - b_t) \qquad \forall\, t \in T \qquad (5)$$

$$v_{d,t} + v_{r,t} \le M b_t \qquad \forall\, t \in T \qquad (6)$$

$$\omega_{d,t} = \omega_{d,t-1} + \rho_{d,t} - x_{d,t} \qquad \forall\, t \in T \qquad (7)$$

$$\omega_{r,t} = \omega_{r,t-1} + \rho_{r,t} - x_{r,t} \qquad \forall\, t \in T \qquad (8)$$

The production of the DMS and RMS are limited to the available capacity of each system, and any unutilized capacity is considered as excess capacity (9), (10). The maximum workload in any period is bounded by the available capacity during that period, since the utilization of a resource cannot exceed 100% (11), (12).

$$x_{d,t} + \varepsilon_{d,t} = C_d \qquad \forall\, t \in T \qquad (9)$$

$$x_{r,t} + \varepsilon_{r,t} = \tau_t \qquad \forall\, t \in T \quad (10)$$

$$\omega_{d,t-1} + \rho_{d,t} \le C_d \qquad \forall\, t \in T \quad (11)$$

$$\omega_{r,t-1} + \rho_{r,t} \le \tau_t \qquad \forall\, t \in T \quad (12)$$

Since RMS capacity can be changed to respond to DMS disruptions, the response time and the capacity changes should be represented in the MIP model. The following section describes how these two transitions are represented.

3.1.1.1 The impact of response time on RMS capacity

In order to have an appropriate estimation of the available capacity of the RMS during the reconfiguration process, we assume that only a portion of the added capacity is available during the response time. Therefore, during this period we deal with two characteristics of capacity: the nominal capacity and the actual capacity. The nominal capacity determines the amount of capacity that the system is set to reach for the following period (13).

$$\xi_t = \xi_{t-1} + \Delta_t^+ - \Delta_t^- \qquad \forall\, t \in T \quad (13)$$

$$\tau_t = \xi_{t-1} + u_t - l_t \qquad \forall\, t \in T \quad (14)$$

$$\xi_t \le fC \qquad \forall\, t \in T \quad (15)$$

The actual capacity will be less than the nominal capacity since some portion of the nominal capacity is lost during the ramp-up period. In a ramp-down period, the actual capacity will be slightly higher than the target level. Therefore, the system is not able to reach the nominal capacity instantly. The actual capacity represents the amount of capacity that is available during the response time (14). During the reconfiguration process, we assume that RMS capacity can be added or removed by changing the modules of the system. The maximum number of modules that could be added to a system determines the maximum RMS capacity (15). Since adding or removing of the modules requires a new setup, we assume that adding or removing each module incurs a reconfiguration cost.

Within the recovery period, Klibi and Martel (2012) represent the gradual capacity recovery of the disrupted depot. In contrast, we assume that there is no supply from the main supplier. Furthermore, we assume that the RMS provides the capacity gradually during the response time. The available capacity of the RMS is modeled as a fraction of nominal capacity through the constraint set (16) to (17).

$$u_t + l_t \leq U_i\left(\Delta_t^+ + \Delta_t^-\right) + M(1 - y_{i,t}) \qquad \forall\, i \in I,\ \forall\, t \in T \qquad (16)$$

$$u_t + l_t \geq L_i\left(\Delta_t^+ + \Delta_t^-\right) - M(1 - y_{i,t}) \qquad \forall\, i \in I,\ \forall\, t \in T \qquad (17)$$

For each capacity change, we assume that L_i to U_i of the added capacity is available during the response time, where $U_i, L_i \in [0,1]$. The impact of the RMS response speed is illustrated through the RMS's available capacity during the response time, such that there is more capacity availability within the response time as the response speed increases. This is indicated through the coefficients of the upper bound U_i and lower bound L_i in constraints (16), (17). The value of these coefficients would be higher for faster speed levels. This can be explained by observing the evolution of the throughput curve *TH(t)* during the response time. As shown in Figure 3.4, the actual capacity during the ramp-up period is measured by the area under the throughput curve. Accordingly, the maximum capacity that can be added is equal to the area determined by $TH(\Delta_t^+) \times t = \Delta_t^+$. The evolution of this throughput curve depends on the layout configuration type (Niroomand et al. 2012). A serial configuration will demonstrate a slower evolution in the beginning of the response time and will reach the desired throughput only at the end of a period when all the stages' reconfiguration is complete. On the other hand, a parallel configuration's throughput curve can improve faster thanks to better scalability (Wang and Koren, 2012). This will result in improved capacity availability for a parallel configuration.

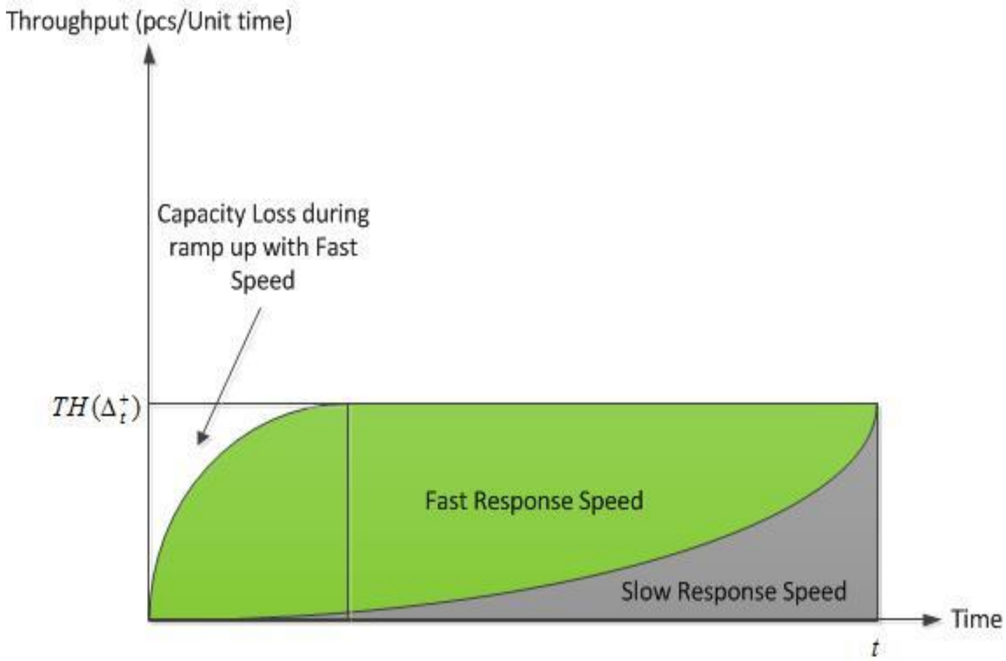


Figure 3.4 The throughput curve during response time.

We assume that the nominal capacity can be changed in predetermined module sizes, which are identified in constraint (18). In any period where there is a reconfiguration process, there could only be a capacity addition or removal (19), (20). These constraints prevent the selection of both capacity addition and removal simultaneously. The response time is also determined through a set of binary variables, in the case of no reconfiguration process; no capacity could be added or removed (21) - (23).

$$\Delta_t^+ + \Delta_t^- = \sum_{i=1}^{f} iCy_{i,t} \qquad \forall\, t \in T \qquad (18)$$

$$\Delta_t^+ + u_t \le Mq_t \qquad \forall\, t \in T \qquad (19)$$

$$\Delta_t^- + l_t \le M(1-q_t) \qquad \forall\, t \in T \qquad (20)$$

$$\sum_{i=0}^{f} y_{i,t} = 1 \qquad \forall\, t \in T \qquad (21)$$

$$\Delta_t^+ + u_t \le M\sum_{i=1}^{f} y_{i,t} \qquad \forall\, t \in T \qquad (22)$$

$$\Delta_t^- + l_t \leq M\sum_{i=1}^{f} y_{i,t} \qquad \forall\ t\ \in\ T \qquad (23)$$

The consideration of supply disruptions at the main supplier can result in an overflow of the demand towards the backup supplier. This overflow results in an accumulation of WIP despite the ramp-up at the backup supplier. There are two drawbacks associated with this situation. First, the congestion created by this overflow will decrease the throughput. Second, the decreased throughput will then result in lost demand. Furthermore, the effects of congestion are also important in order to properly assess the actual capacity of the suppliers, especially in a contingency strategy.

3.1.1.2 The impact of the congestion on throughput

Under the contingent rerouting strategy, the disruption in the supply chain can lead to significant losses if the selected backup supplier is not responsive enough. These losses can have negative implications in the short term as well as in the long term. In the short term, there would be some lost sales if customers are not filled promptly. In the long term, the incurred losses may result in losing market share to competitors and end up being far more severe than short-term losses (Farahani et al. 2013). In our context, the amount of these losses depends on the scalability of the backup supplier. Despite the scalability level of the backup supplier, rerouting the total demand at the beginning of the response period will create an overflow when the backup capacity is not 100% available. As a result of this overflow, the congestion effects such as queue build up and increased lead time will be observed. Therefore, the amount of the backup supply during the response time would be limited.

The consideration of congestion effects can allow better representation of the available capacity during the response period. This representation will allow determining the appropriate response speed level of the backup supplier. In order to incorporate the effects of congestion, we present the suppliers as a single server system with Poisson arrivals and general service time distribution (M/G/1 system). The relationship developed using this model allows developing the clearing function, defining the relationship between the workload and throughput in steady state

(Missbauer 2002). Based on this clearing function, the expected system throughput $E(x_t)$ in any period is a function of the expected work load $E(\omega_{t-1}+\rho_t)$, available capacity τ and the mean and the variance of the processing time:

$$E(\mathrm{x}_t)=\frac{1}{2}\left[\tau+k+E(\omega_{t-1}+\rho_t)-\sqrt{\tau^2+2\tau k+k^2-2\tau E(\omega_{t-1}+\rho_t)+2kE(\omega_{t-1}+\rho_t)+E(\omega_{t-1}+\rho_t)^2}\right] \quad (24)$$

Where k is defined based on the mean $\frac{1}{\mu}$ and the variance $\frac{1}{\sigma^2}$ of the processing time as follows:

$$k=\frac{\mu\sigma^2}{2}+\frac{1}{2\mu} \quad (25)$$

The clearing function in (24) is concave and nonlinear (Missbauer 2002). An outer approximation approach has been used to generate a set of lines in order to linearize the MIP model. In order to minimize the error between the actual curve and the approximated lines, the location and the number of tangent points are determined by the subtractive clustering method (Chiu 1994). This allows identifying the number of lines that will approximate the clearing function. The detailed explanation of the method is presented in Appendix A.

Based on this linearization, the constraints generated for each supplier can be defined as follows. The production by the DMS supplier could not be more than the expected throughput, which is estimated by its clearing function (26). This estimation is based on a predetermined service rate. The set of lines N represents the clearing function where A_η is the slope and B_η is the constant value of the line η .

$$x_{d,t}\le A_\eta(\omega_{d,t-1}+\rho_{d,t})+B_\eta \qquad \forall\ \eta\in N,\ \forall\ t\in T \quad (26)$$

Since the RMS has varying capacity levels within the planning horizon, a set of binary variables are presented in (18), (27), (28) to activate the clearing function corresponding to each level.

$$\xi_t = \sum_{j=C}^{fC} j k_{j,t} \qquad \forall\, t \in T \qquad (27)$$

$$\sum_{j=C}^{fC} k_{j,t} = 1 \qquad \forall\, t \in T \qquad (28)$$

Similar to (26), the production of the RMS supplier in the periods with fixed capacity level j is limited by the expected throughput (29). In this constraint, the set of lines N_j is employed to replace the clearing function.

$$x_{r,t} \le A_\eta(\omega_{r,t-1} + \rho_{r,t}) + B_\eta + M(2 - y_{0,t} - k_{j,t}) \qquad \forall\, j \in J,\ \forall\, \eta \in N_j,\ \forall\, t \in T \qquad (29)$$

If there is a change in the capacity level of the RMS supplier, the throughput during the reconfiguration period is a function of both the workload and actual capacity, as indicated in Figure 3.5.

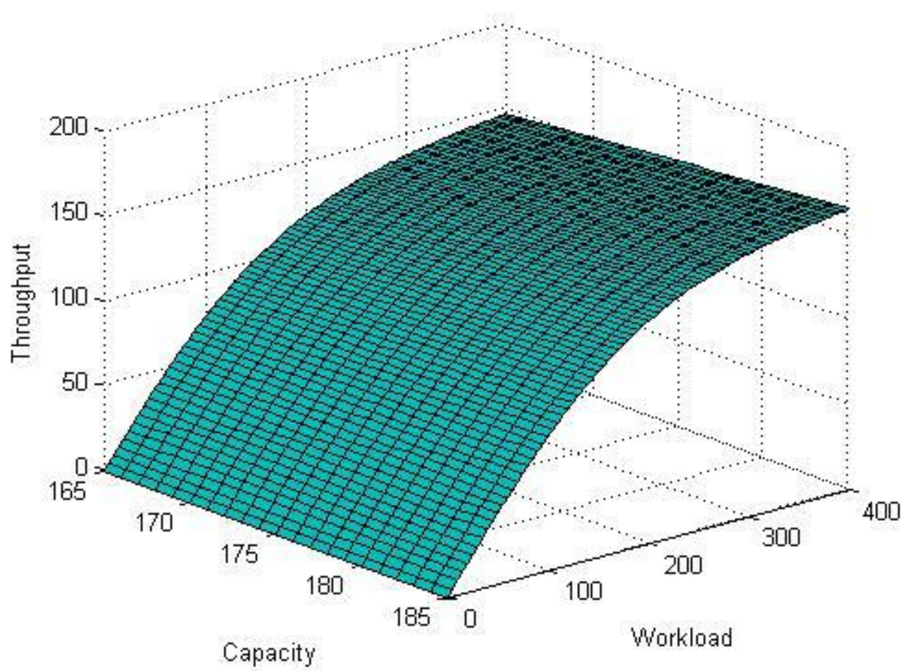


Figure 3.5 Clearing function.

In order to incorporate this clearing function in the MIP model, it should be converted into a set of hyper planes. Based on this conversion, the production quantity of the RMS supplier within the response time once it adds i modules to reach to the capacity level of j could not exceed the expected throughput (30).

For any capacity increment scenario i, j , the expected throughput is approximated through the set of hyper planes $V^{+}_{i,j}$. In this set, the parameters $A_{\upsilon}, B_{\upsilon}$ are the slopes with respect to the workload and capacity respectively and G_{υ} is the constant value of the hyper plane υ.

$$x_{r,t} \le A_{\upsilon}(\omega_{r,t-1} + \rho_{r,t}) + B_{\upsilon}\tau_t - G_{\upsilon} + M(2 - y_{i,t} - k_{j,t}) \quad \forall\, i \in I,\ \forall\, j \in J,\ \forall\, \upsilon \in V^{+}_{i,j},\ \forall\, t \in T \qquad (30)$$

The above constraint incorporates the set of hyper planes $V^{+}_{i,j}$ as an upper bound over the production quantity of the RMS when it increases its capacity level. In order to apply this constraint for capacity reduction case, the set of hyper planes $V^{-}_{i,j}$ is presented in (31).

$$x_{r,t} \le A_{\upsilon}(\omega_{r,t-1} + \rho_{r,t}) + B_{\upsilon}\tau_t - G_{\upsilon} + M(2 - y_{i,t} - k_{j,t}) \quad \forall\, i \in I,\ \forall\, j \in J,\ \forall\, \upsilon \in V^{-}_{i,j},\ \forall\, t \in T \qquad (31)$$

The non-negativity constraints are presented as follows:

$$\begin{array}{l} x_{d,t} \ge 0,\ x_{r,t} \ge 0,\ \Delta^{+}_{t} \ge 0,\ \Delta^{-}_{t} \ge 0,\ s_t \ge 0,\ \pi_{d,t} \ge 0,\ \pi_{r,t} \ge 0,\ \varepsilon_{d,t} \ge 0,\ \varepsilon_{r,t} \ge 0, \\ \nu_{d,t} \ge 0,\ \nu_{r,t} \ge 0,\ \omega_{d,t} \ge 0,\ \omega_{r,t} \ge 0,\ \rho_{d,t} \ge 0,\ \rho_{r,t} \ge 0,\ \tau_t \ge 0,\ \xi_t \ge 0,\ u_t \ge 0, \\ l_t \ge 0,\ b_t \in \{0,1\},\ q_t \in \{0,1\} \end{array} \quad \forall\, t \in T \qquad (32)$$

$$y_{i,t} \in \{0,1\} \quad \forall\, i \in I,\ \forall\, t \in T \qquad (33)$$

$$k_{j,t} \in \{0,1\} \quad \forall\, j \in I,\ \forall\, t \in T \qquad (34)$$

While the presented MIP model generates the capacity plan under the normal operational condition of the DMS, additional variables and constraints allow integrating the disruption scenarios and generating the performance information for each scenario. In the case of DMS failure, the contingency capacity plan is generated by modifying the MIP according to the steps explained in the following section.

3.1.2 Representation of disruptions in the MIP model

In generating the contingency plan for each disruption scenario, the following sets of changes are incorporated in the MIP model. First, the binary variable O_t is incorporated to indicate the DMS supplier failure.

$$O_t = \begin{cases} 1 & \text{If DMS is available} \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, t \in T \qquad (35)$$

Once the DMS supplier is disrupted, the demand could be satisfied through current inventory and/or RMS production (36). For a given disruption scenario there should be no production and material release to the DMS supplier when it is disrupted (37), (38). The DMS WIP level during the disrupted periods remains equal to the WIP level of the last period before the disruption (39).

$$v_{d,t} = v_{d,t-1} + O_t x_{d,t} - \pi_{d,t} \qquad \forall\, t \in T \qquad (36)$$

$$x_{d,t} + \varepsilon_{d,t} = O_t\, \mathrm{C}_d \qquad \forall\, t \in T \qquad (37)$$

$$\rho_{d,t} \le MO_t \qquad \forall\, t \in T \qquad (38)$$

$$\omega_{d,t} = \omega_{d,t-1} + O_t(\rho_{d,t} - x_{d,t}) \qquad \forall\, t \in T \qquad (39)$$

In order to avoid the MIP model building inventory in advance of the disruption period m, the inventory levels of the DMS and RMS and the capacity plan of the RMS for the periods before the disruption are set to the values obtained in the initial capacity planning model. These values are presented in constraints (40)-(42) by $I_{d,t}$, $I_{r,t}$, C_t accordingly. This allows generating a contingency plan, which requires the use of the backup supplier for the disrupted periods.

$$v_{d,t} = I_{d,t} \qquad \forall\, t \in \{1,\dots,m-1\} \qquad (40)$$

$$v_{r,t} = I_{r,t} \qquad \forall\, t \in \{1,\dots,m-1\} \qquad (41)$$

$$y_{i,t} = C_t \qquad \forall\, t \in \{1,\dots,m-1\} \qquad (42)$$

These changes allow generating the supply chain contingency plan under the failure of the DMS supplier. In the disruption periods, the MIP model should trigger a capacity change in the backup supplier according to a preselected response speed.

3.1.3 Decision tree analysis

Due to the stochastic nature of the disruptions, an optimal response speed can be identified based on the expected supply chain costs and the attitude of the decision maker towards risk. In identifying the optimal response speed for the backup supplier, a decision tree analysis is conducted. The decision tree is a well-known technique in the field of decision analysis under risk (Berger et al. 2004).

As indicated in Figure 3.6, the square node is the decision options and the circle nodes represent the chance events. The decision options regarding the response speeds RS_k include Fast speed RS_1, Medium speed RS_2 and Slow speed RS_3. For a disruption scenario (m,n) belonging to the set of all plausible future scenarios S with the probability of $P_{(m,n)}$, each of these response speeds associated with the backup supplier corresponds to a total cost $Z_{(RS_k,m,n)}$ as a result of the contingency plan generated in the MIP model. This allows computing the expected cost corresponding to each decision through the following formula:

$$E(RS_k) = \sum_{(m,n)\in S} P_{(m,n)} Z_{(RS_k,m,n)} \qquad \forall k \in \{1,2,3\} \qquad (43)$$

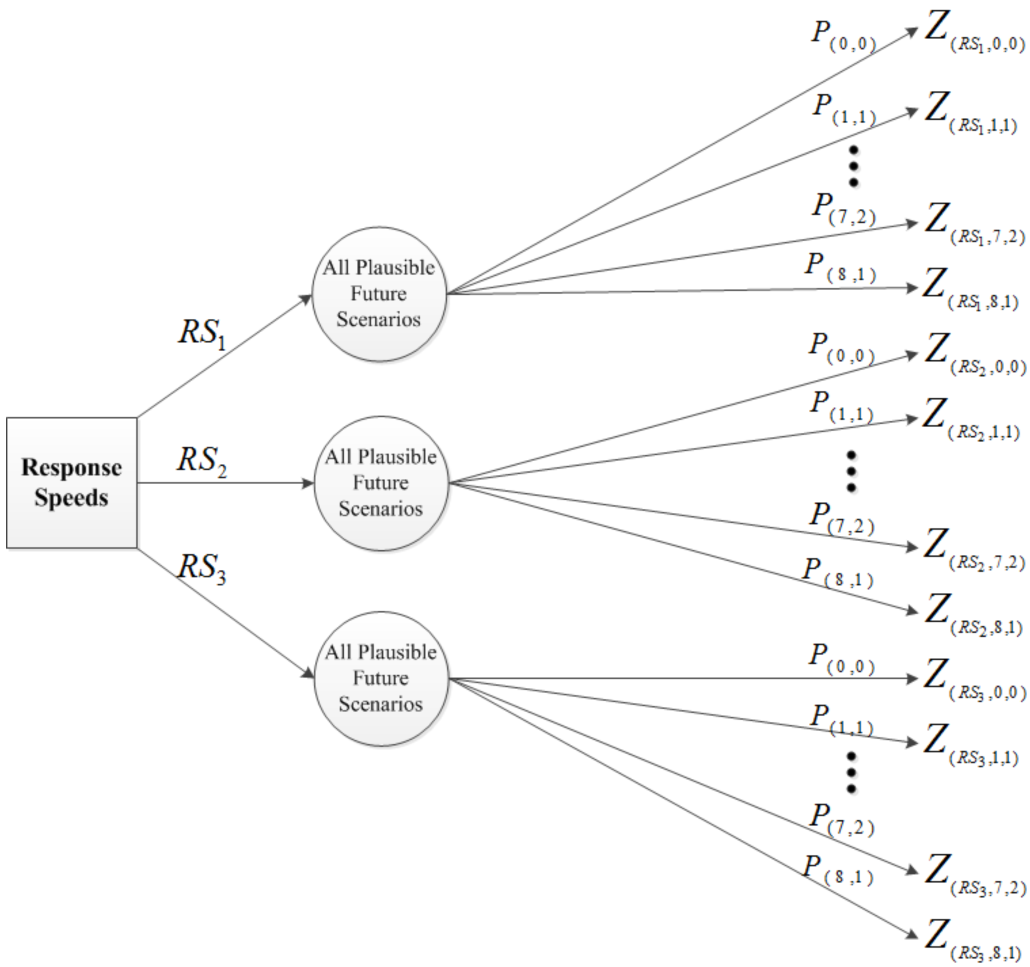


Figure 3.6 The decision tree for the optimal selection of the response speed (e.g., $T=8$).

In order to identify the probability of each disruption scenario, the Markov discrete time distribution is incorporated. The parameter α represents the probability of a disrupted period following a non-disrupted period (failure probability). The parameter β defines the probability of a non-disrupted period following a disrupted period (recovery probability). Based on these assumptions, the probability of a disruption at time m with the length of n is computed through the following formulas.

$$P_{No\ disruption}(m,n) = (1-\alpha)^T \tag{44}$$

$$P_{Disruption}(m,n) = \alpha\beta(1-\alpha)^{m-1}(1-\beta)^{n-1} \qquad n \in \{1,...,T-m\}, \forall m \leq T \tag{45}$$

$$P_{Disruption}(m,n)=\alpha(1-\alpha)^{m-1}(1-\beta)^{n-1} \qquad n=T-m+1, \forall m \leq T \tag{46}$$

The details regarding the derivation of these formulas can be found in Appendix B. The expected cost criteria can be used to identify the optimal response speed where the decision maker is risk-neutral. The response speed, which minimizes the expected cost of the supply chain under all plausible future scenarios, is the optimal speed. If the decision maker is risk-averse then the criteria of Conditional Value at Risk $CVaR$ can be utilized.

Once the decision maker is risk-averse, the risk of high losses as a result of disruptions is controlled by the confidence level γ. This means that there is a target cost of portfolio VaR such that the costs for γ percent of the scenarios would be less than or equal to VaR. The remaining $1-\gamma$ percent of scenarios are the worst-case scenarios. A risk-aversive decision maker minimizes the expected cost of the worst-case scenarios defined as Conditional Value at Risk $CVaR$ (Sawik 2012):

$$CVaR_{RS_k} = VaR_{RS_k} + (1-\gamma)^{-1} \sum_{(u.v)\in S} P_{(u,v)} \Delta_{(u,v)} \tag{47}$$

Where (u,v) are the scenarios that cost more than VaR and $\Delta_{(u,v)}$ are defined as follows:

$$\Delta_{(u,v)} = Z_{(RS_k,u,v)} - VaR_{RS_k} \tag{48}$$

The optimal response speed could also be identified with the objective of minimizing the expected supply chain costs under all plausible future scenarios as well as the expected supply chain costs when subjected to worst-case disruptions. In this case, a coefficient λ represents the weight of these objectives in the decision-making stage. For that purpose, the following formula is employed.

$$Min\ \lambda E(RS_k) + (1-\lambda) CVaR_{RS_k} \tag{49}$$

The decision tree analysis allows identifying the appropriate response speed of the backup supplier with the objective of minimizing the expected supply chain costs with respect to the attitude of the decision maker towards risk. The selection of the optimal policy depends heavily on the probability of failure and recovery of the main supplier. The value of these probabilities can depend on several parameters, such as the hazard exposure level of the geographical zone in which the facility is located, as well as the ability to return to the operational condition once it is disrupted (Klibi et al. 2010). Section 3.2 presents an example to investigate the optimal selection of the response speed with respect to different failures and recovery probabilities of the main supplier.

3.2 Numerical Experiments

This section presents an example in order to illustrate the proposed methodology. We consider the supply chain associated with a product whose lifecycle lasts for eight periods. The demand level follows a classical pattern over the lifecycle of the product, as indicated in Table 3.3.

Table 3.3 Demand scenario

t	1	2	3	4	5	6	7	8
D_t	406	530	580	629	629	486	384	303

The MIP model is used to generate the capacity plan by determining the capacity, production, raw material, inventory and WIP levels within the planning horizon. In identifying the capacity plan, the following assumptions have been made regarding cost and capacity-related input data, as indicated in Table 3.4.

While the raw material purchasing cost is the same for both suppliers, the production cost of the RMS is higher than that of the DMS (Tomlin 2006). Therefore the WIP and the finished good inventory holding cost of the RMS are higher than for the DMS. There are three different RMS reconfiguration costs corresponding to three different response speed levels. The reconfiguration costs increase as the response speed levels are improved.

Table 3.4 Supplier’s costs parameters $/Unit

Cost parameter	M_d	M_r	P_d	P_r	W_d	W_r	H_d	H_r	$R(Slow)$	$R(Medium)$	$R(Fast)$
Value	20	20	20	100	10	15	35	40	40	60	100

The RMS excess capacity costs and the product's shortage costs are presented in Table 3.5. The product shortage costs are defined with respect to the demand pattern, such that there is a higher shortage cost in the introduction and growth periods compared to the maturity periods of the lifecycle. On the other hand, the decline periods have a lower shortage cost compared to the maturity phase.

Table 3.5 Excess and shortage costs $/Unit

t	1	2	3	4	5	6	7	8
E_r	20	25	30	35	40	50	60	70
S	460	430	410	390	370	350	330	310

The supplier with the DMS facility has a fixed capacity of 500 while the RMS-equipped supplier can vary its capacity according to the steps indicated in Table 3.6. The initial configuration of the RMS is a base which provides 100 units of capacity. It can raise its capacity level by adding modules according to Table 3.6.

Table 3.6 RMS capacity levels with respect to its structure pcs/Period

i	0	1 Module	2 Modules	3 Modules
j	100	200	300	400

The coefficients of the upper and lower bounds in constraints (16), (17), which represent the amount of the available capacity during the response time, are identified in Table 3.7. According to scalability characteristics explained in 3.1.1.1, we consider higher values for faster response speed configurations. Since adding higher capacity would result in a longer reconfiguration process, the percentage of available capacity decreases as the number of added modules increases (Niroomand et al. 2012).

Table 3.7 The coefficients of capacity boundaries corresponding to different speeds

Response speed	Slow			Medium			Fast		
Number of added modules (i)	1	2	3	1	2	3	1	2	3
U_i	0.75	0.5	0.4	0.85	0.65	0.5	0.95	0.85	0.7
L_i	0.5	0.4	0.2	0.65	0.5	0.3	0.85	0.7	0.55

Based on the stated assumptions regarding the supply chain and input data, the following experiments are conducted. First, we present the benefit of considering the effects of congestion in evaluating the performance of a contingency strategy. Second, an optimal contingency strategy is assessed by identifying the response speed of the RMS within a range of failure and recovery probabilities. The selection of optimal response speeds are then evaluated based on the attitude of the decision maker towards risk. Specifically, we look at the value of considering congestion in the selection of response speed at various levels of tolerance to risk. We quantify this impact through the difference in service levels.

The MIP model to generate the regular capacity plans as well as the contingency plans has been implemented in ILOG CPLEX version 12.5. By setting the desired optimality gap to 0.0001, the results of the contingency plans have been obtained with an average optimality gap of 0.0015 at an average computation time of 4.5 seconds.

3.2.1 The impact of congestion on contingency strategy performance

In this section, the impact of considering the congestion is evaluated by observing the supply chain service level under two conditions. First, the MIP model is used to determine the capacity plan, the production quantities, WIP and inventory levels corresponding to DMS and RMS suppliers without considering the impact of congestion (load-independent model). For this purpose, the constraints (26), (29), (30), (31) representing the clearing functions are removed. This implies that any amount of release to a production system will be exactly produced which is an overestimation of capacity.

Afterwards, the clearing functions present the actual production quantity of the DMS and RMS based on the capacity and WIP levels which have been determined in the load-independent model. The actual demand losses are then computed by the difference between demand and production quantities obtained using the clearing function. This will give the actual service level which is determined as a fraction of satisfied demand over the total demand within the planning horizon.

Second, in order to compute the supply chain service level once the congestion impact is considered (load-dependent model), the proposed MIP model including clearing functions is executed. To illustrate the results, we present the case for a DMS disruption scenario at time 3 with a length of 3 periods, in Table 3.8.

The production quantities $x_{d,t}$, $x_{r,t}$ in the load-independent model are overestimated as a result of ignoring the congestion (e.g., $x_{d,2} = 500$ while the actual production value is 432). Once the impact of congestion is considered (load-dependent model), the MIP model would increase the RMS capacity ξ_t to a higher level compared to the load-independent model to cover the shortages (e.g., $\xi_2 = 100$ in the load-independent model versus $\xi_2 = 200$ in the load-dependent model). As a result of this, the service level of the load-dependent model would be higher than its load-independent counterpart. This behavior is observed in all plausible scenarios.

Table 3.8 Capacity planning in load-independent versus load-dependent models

Decision Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	MIP results in load-independent capacity plan (no congestion effects)							
$x_{d,t}$	474	500	0	0	0	486	384	303
ξ_t	100	100	400	400	400	100	100	100
τ_t	100	100	250	400	400	250	100	100
$x_{r,t}$	0	0	250	400	400	0	0	0
	Actual production levels due to congestion effects							

ACT $x_{d,t}$	419	432	0	0	0	424	358	290
ACT $x_{r,t}$	0	0	186	333	333	0	0	0
s_t	55	68	356	296	296	62	26	13
Service Level	0.70							
	MIP results with load-dependent capacity plan (with congestion effects)							
$x_{d,t}$	431	405	0	0	0	425	379	302
ξ_t	100	200	400	400	400	100	100	100
τ_t	100	184	330	400	400	250	100	100
$x_{r,t}$	0	100	262	333	333	61	5	1
s_t	0	0	318	296	296	0	0	0
Service Level	0.77							

Figure 3.7 represents the service level of the load-independent model and load-dependent model for disruptions which might occur at time 3 at varying lengths, with an RMS at medium response speed.

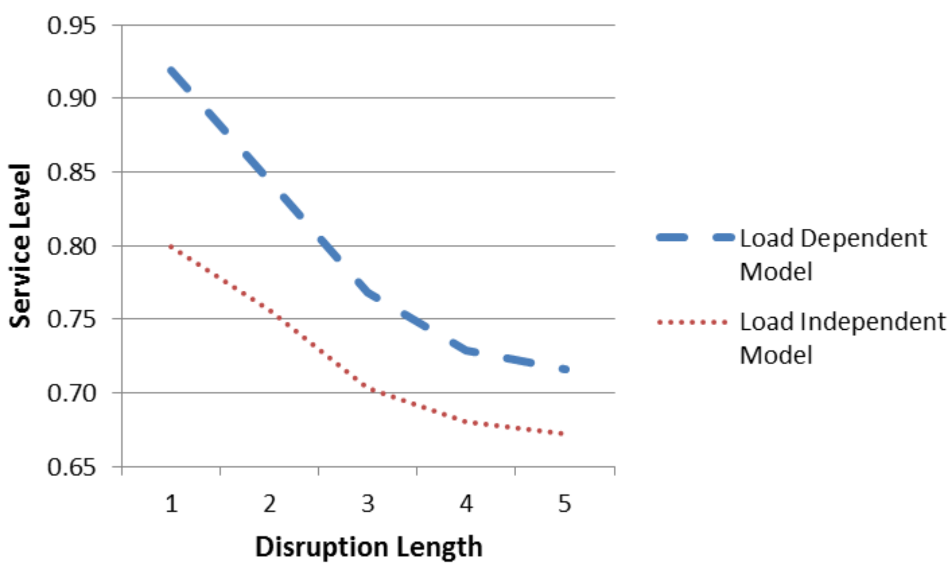


Figure 3.7 The impact of considering congestion effects on service level.

The results show that considering the congestion effects in developing a contingency plan for a supply chain improves the service level. This is due to the fact that the backup supplier capacity can be planned accordingly. As indicated in Figure 3.7, the service level decreases in both cases as the length of the disruption increases. This happens due to losing the main supplier, which has a higher capacity level compared to the backup supplier. As the disruption length increases we observe that the service level difference between two cases decreases. This is due to the fact that the backup supplier can't replace the main supplier over long periods of disruption.

The results presented so far show improvement in the service level of the supply chain by considering congestion. Furthermore, the supply chain responsiveness to major disruptions improves once the appropriate response speed of the RMS is determined at the design stage of the supply chain configuration. While considering congestion improves the accuracy of the decision-making process, it also increases the complexity of the solution methodology. So the questions are what is the impact of not considering congestion? Under which risk tolerance levels is it worth considering the impacts of congestion? The following section answers these questions through the use of the proposed methodology.

3.2.2 The value of incorporating congestion in the selection of response speed

In identifying an appropriate response speed in a contingency strategy, the trade-off between the investment cost of responsiveness and lost sales can be considered as main determinants. The outcome from this trade-off can be different depending on the attitude of the decision maker towards risk.

For the given example, the response speed of the RMS is determined through the decision tree analysis which is explained in section 3.1.3. This selection is based on the outcomes of the MIP model which excludes the clearing functions. Moreover, it is done with respect to three different attitudes of the decision maker toward risk. The impact of the different failure and recovery probabilities on the decision-making process is evaluated by a sensitivity analysis. The failure probabilities range between 0 and 0.2 with increments of 0.05, while the recovery probabilities are within 0 and 0.5 with increments of 0.1. With the objective of evaluating the benefit of considering congestion, the aforementioned process is repeated by adding the constraints related to the clearing function to the MIP model.

3.2.2.1 Risk-neutral decision maker

In a risk-neutral behavior, the optimal response speed is selected by comparing the expected cost of the supply chain under all plausible future scenarios. The probability of occurrence for each scenario is determined by the failure and recovery probabilities of the DMS supplier. As illustrated in Figure 3.8, the expected cost of the supply chain grows as the failure probability increases and/or the recovery probability decreases.

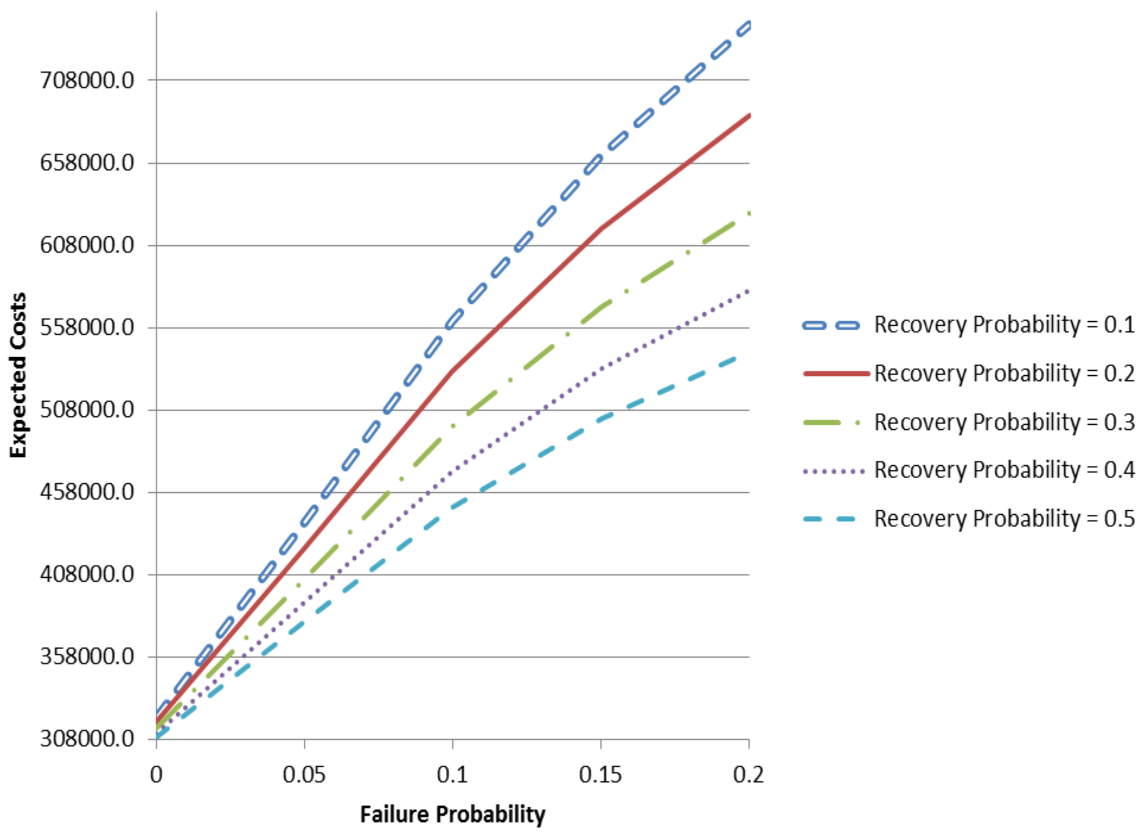


Figure 3.8 Sensitivity analysis of the expected costs of the supply chain.

The probability of the scenarios with disruption increases as the failure probability increases. Since these scenarios have higher cost compared to the scenario without disruption, the expected cost of the supply chain increases. On the other hand, the probability of the scenarios with long disruptions increases when the recovery probability decreases. These scenarios have higher cost compared to the scenarios with short disruptions. Therefore the expected cost of the supply chain increases.

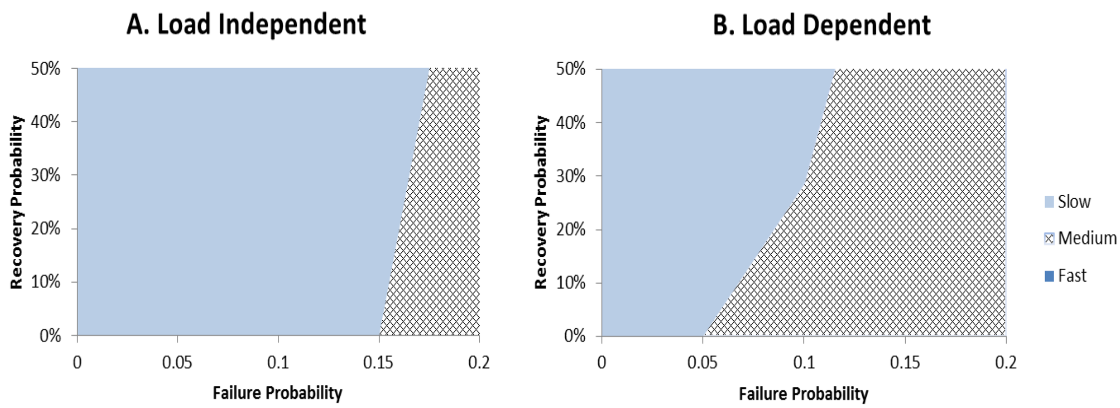


Figure 3.9 Optimal response speed, Risk-neutral.

Figure 3.9 represents the optimal selection of response speed for various failure and recovery probabilities in two conditions. Figure 3.9.A indicates the results once the congestion effects are ignored in the decision-making stage. The slow speed is optimal in most of the failure and recovery probability combinations except the region corresponding to high failure and a long recovery period. Since a higher amount of the capacity is required within the response time for those situations, the medium speed is selected.

The results presented in Figure 3.9.B are achieved by considering the congestion impact. The selected response speeds are at higher levels compared to Figure 3.9.A. These differences result from the need to have higher speed levels to cover the losses that would be created as a result of congestion. In such cases, there are significant improvements in the expected service level of the supply chain, as presented in Table 3.9. For a selected subset of failure and recovery probabilities, the expected service level of the supply chain is presented in Table 3.9.

Table 3.9 The expected service level of the supply chain (%)

	Recovery Probability (β)	Failure Probability (α) 0.01	0.05	0.1	0.15	0.2
Load-Dependent Models	0.2	95.9	90.9	91.4	88.6	85.1
Load-Independent Models		81.3	77.5	73.8	69.8	73.1
Improvement in Service Level		14.6	13.4	17.5*	18.9*	12.0
Load-Dependent Models	0.5	96.2	92.5	88.8	91.1	87.1
Load-Independent Models		81.5	78.6	75.8	72.5	74.4
Improvement in Service Level		14.7	13.9	13.0	18.7*	12.7

*A slower response speed is selected in load-independent model

As shown in Table 3.9, the expected service level of the supply chain is higher when congestion effects are incorporated. In addition, the improvement in service level is significant if an incorrect response speed level is selected when congestion effects are ignored (e.g., β = 0.2, α = 0.1 and α = 0.15).

3.2.2.2 Risk-averse decision maker

In the case where the decision maker is risk-averse, the response speed is selected to minimize the expected cost of the worst-case scenarios according to the level of risk aversion. Figure 3.10 presents the selection of the response speed when the congestion is ignored for a risk-averse decision maker at various tolerance levels.

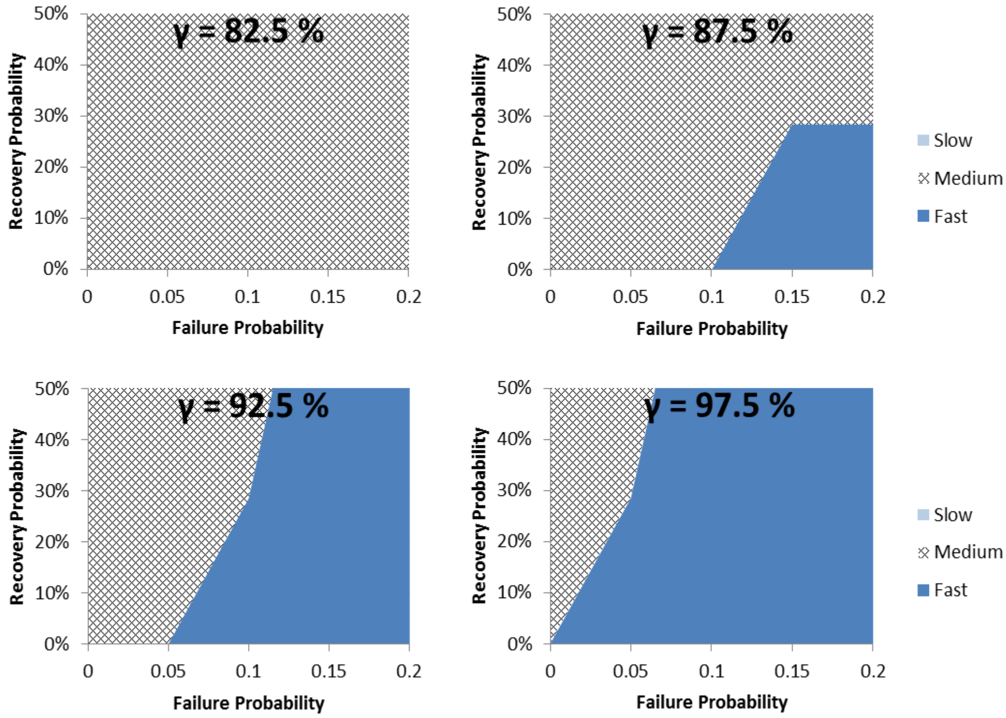


Figure 3.10 Optimal response speed, Risk-averse, Load-independent models.

For the low level of the risk aversion e.g., γ = 82.5%, the slow speed is not optimal since it provides a low capacity within the response time. The medium speed is selected as the optimal speed in all combinations of the failure and recovery probability since there are 17.5% of the worst-case scenarios in the decision-making stage which require a higher amount of the capacity within the response time.

As the decision maker becomes more risk-averse, the focus would be on the smaller portion of the worst-case scenarios, albeit those with higher impacts. Therefore, more capacity within response time is required to minimize the impact of such disruptions. As a result, the need for medium speed is reduced and the tendency to select fast speed increases. The optimal selection of the response speed when we consider the congestion effects is indicated in Figure 3.11.

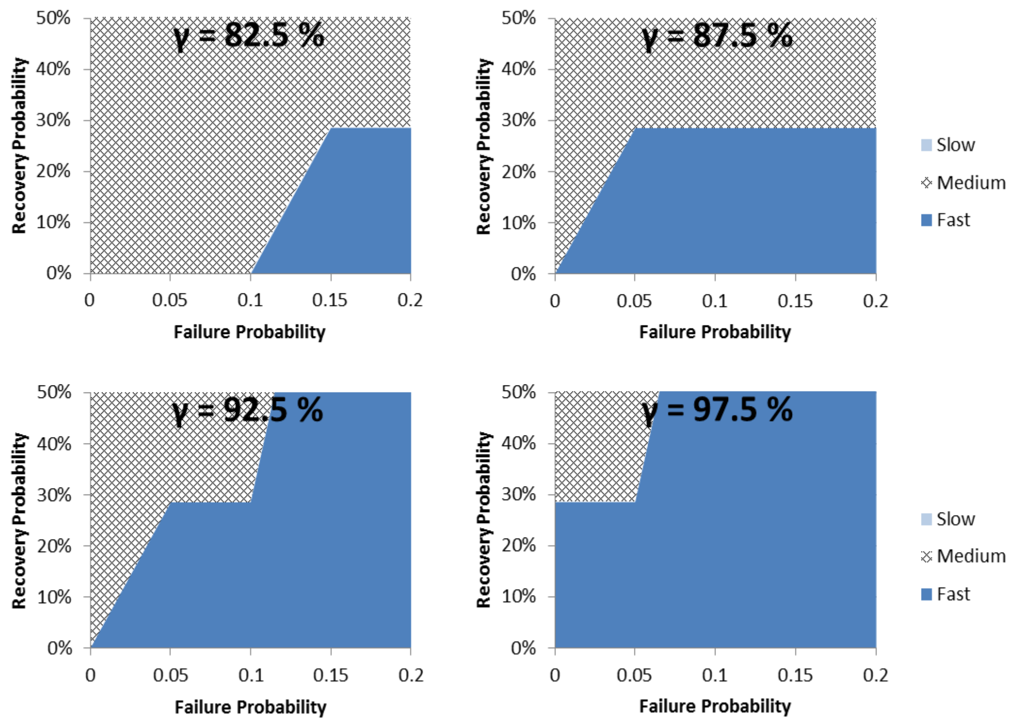


Figure 3.11 Optimal response speed, Risk-averse, Load-dependent models

Incorporating congestion provides the same trend in the selection of response speed: as the decision maker is more risk-averse, faster speeds are selected. However, higher speed levels are selected for various levels of the risk aversion compared to Figure 3.10. These would lead to significant improvements in the expected service level of the supply chain, as illustrated in Table 3.10.

Table 3.10 Improvements in the expected service level of the supply chain with load-dependent models (%), Risk-averse

Recovery Probability (β)	γ	Failure Probability (α)				
		0.01	0.05	0.1	0.15	0.2
0.2	82.5%	14.60	13.30	13.00	21.00*	22.00*
	87.5%	14.60	17.00*	19.00*	14.10	14.00
	92.5%	14.60	16.00*	13.90	14.10	14.00
	97.5%	15.00*	13.30	14.00	14.10	14.00
0.5	82.5%	14.70	14.00	13.90	11.30	12.70
	87.5%	14.70	14.00	13.90	11.30	12.70
	92.5%	14.70	14.00	13.90	14.00	15.00
	97.5%	14.70	14.00	14.00	14.70	15.00

*A slower response speed is selected in load-independent model

For a decision maker with the risk aversion level of 82.5%, the correct selection of response speed by considering the congestion impacts increases the expected service level of the supply chain by 21% and 22%. However, these improvements decrease as the level of the risk aversion increases (e.g., 16% for $\gamma = 92.5\%$). This is due to the fact that as the decision maker becomes more risk-averse, faster speed levels would be selected for high probability of failure and/or low probability of recovery in both load-dependent and load-independent models.

3.2.2.3 Mean-risk approach

The optimal response speed could also be determined by considering the risk neutrality and risk aversion simultaneously. In this case, the weight of the risk neutrality and risk aversion in the decision-making process are determined through λ and $1-\lambda$ respectively in (49). Figure 3.12 indicates the optimal selections of response speed by ignoring the congestion impact for $\gamma = 97.5\%$ with different values of λ.

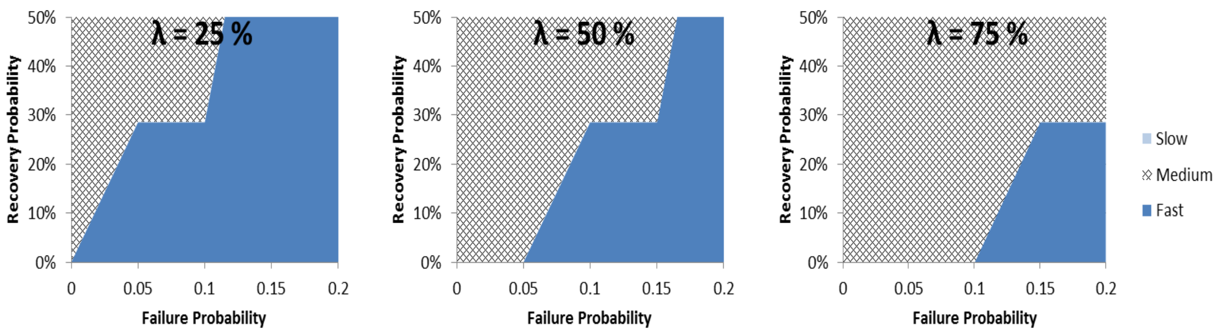


Figure 3.12 Optimal response speed, Mean risk, Load-independent models, γ = 97.5%.

For the low value of λ, the fast response speed has been selected as the optimal speed for the majority of the failure and recovery probability combinations. This is due to relatively high importance of worst-case scenarios and the need for more capacity within the response time. However as the weight of the risk neutrality increases, the lower levels of the response speed are selected. The effect of incorporating the congestion in the decision process of a mean-risk decision maker is presented in Figure 3.13.

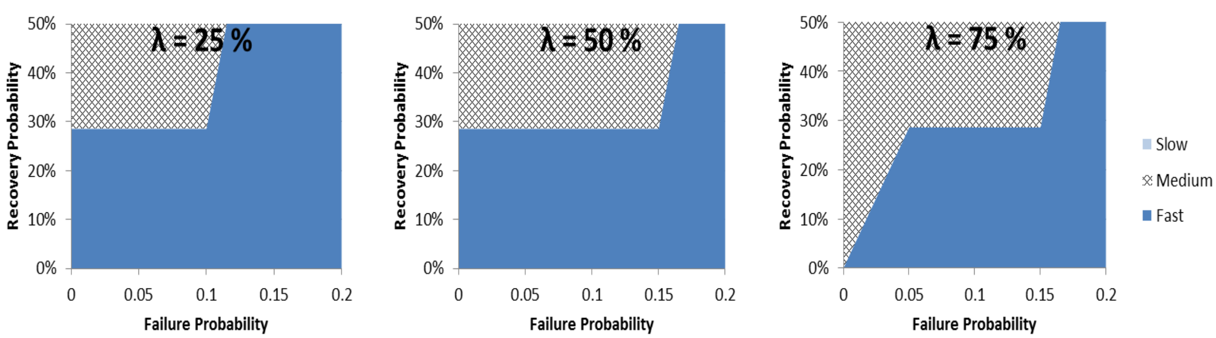


Figure 3.13 Optimal response speed, Mean risk, Load-dependent models, γ = 97.5%.

The comparison of Figure 3.12 and Figure 3.13 reveals that considering the congestion leads to the selection of faster speed levels for a mean-risk decision maker. The proper selection of the optimal response speed would result in major improvements in the supply chain expected service levels, as presented in Table 3.11.

Table 3.11 Improvements in the expected service level of the supply chain with load-dependent models (%), Mean-risk

Recovery Probability (β)	λ	Failure Probability (α)				
		0.01	0.05	0.1	0.15	0.2
0.2	25%	15.00*	13.30	13.90	14.10	14.00
	50%	16.00*	17.00*	13.90	14.10	14.00
	75%	14.60	17.00*	19.00*	14.10	14.00
0.5	25%	14.70	14.00	14.00	14.70	15.00
	50%	14.70	14.00	13.90	11.30	15.00
	75%	14.70	14.00	13.90	11.30	23.00*

*A slower response speed is selected in load-independent model

The results show that if the relative importance of risk neutrality is low, the improvement in the expected service level as a result of correct estimation of the response speed is low (e.g., 15% for $\beta = 0.2,\ \lambda = 25\%$). This is because the selection of response speed in both load-dependent and load-independent approaches does not change for high probabilities of failure and/or low probabilities of recovery. As the λ increases, this improvement in service level would increase considerably (e.g., 17%, 19% and 23% for λ = 75%). This is due to the fact that as the attitude of the decision maker increases toward risk neutrality, ignoring congestion effects results in the selection of a slower response speed than it should. Accordingly, it is essential to incorporate congestion effects for risk-neutral decision maker.

3.2.3 Discussion of results

The first section of the numerical results shows the improvement in supply chain service level upon considering the impact of congestion in the planning stage. This is achieved by triggering the RMS supplier to provide higher capacity level to cover the shortages due to congestion. Afterwards, the response speed of the RMS is determined with the purpose of improving the supply chain responsiveness once the DMS is disrupted.

For a risk-neutral decision maker, the numerical results show the optimality of the slower response speeds for the lower probability of DMS failure and higher probability of recovery. However, as the failure probability of the DMS increases or the recovery probability decreases, the tendency is toward the faster speeds. Considering the congestion effect leads to the selection of the higher speed levels which increase the expected service level of the supply chain significantly. A risk-aversive decision maker would select faster response speed levels compared to the risk-neutral counterpart. However, we observe that the selection of response speed is indifferent with respect to the congestion effects. This is due to the fact that risk-averse decision maker is sensitive to even the slightest loss. In the mean-risk setting, as the attitude of the decision maker increases toward risk neutrality, significant improvements are observed in the supply chain performance upon considering the congestion.

If the failure and recovery probabilities of the main supplier are accurate, the proposed methodology would give a precise perspective to the supply chain management regarding the selection of the backup source's response speed. The outcome of this selection could affect the configuration of the backup source.

This chapter presents the supply chain design requirements in order to incorporate the contingent sourcing as a risk mitigation strategy to cover supply major disruptions. However, in order to design a robust supply network, the operational risks should also be considered in the design stage. Therefore, in next chapter, we review the supply chain design requirements in order to apply a combination of contingent sourcing and strategic stock to cover major disruptions and operational risks. Furthermore, we consider the scenario in which the available capacity during the response time is random (Matta et al., 2007) and we present the methodology to consider this randomness in decision making stage.

Chapter 4

On the value of response time characteristics in robust design of supply flow

In this Chapter, the objective is to design a robust supply network which could sustain the flow of material supply under operational risks and major disruptions in future (Wieland 2013). We incorporate contingent sourcing and strategic stock as cost-effective risk management strategies to deal with supply interruptions. In this setting, the strategic supply chain design decisions include identifying the optimal level of strategic stock and response speed of volume-flexible back-up supplier. To this end, we develop a decision-making tool which considers the randomness associated with available capacity in addition to congestion impacts during the response time in order to improve the quality of solutions. Furthermore, we evaluate the value of considering the response time characteristics in the strategic design stage of supply network.

We consider a supply chain network with similar configuration to the supply network presented in Chapter 3. There is a single product supply chain that includes a manufacturing plant with dual sourcing. The main supplier is cost-effective as a result of dedicated facilities though prone to disruptions during which it may be partially or completely unavailable. There is a back-up supplier located in a low-risk region that is available when the main supplier is unavailable. The back-up supplier has volume-flexible production facilities where it can scale up its capacity according to a speed related to its configuration. However, this scalability increases the production cost of the back-up supplier.

Demand in the normal periods can be met by the main supplier up to a maximum level of C_d. The raw material $\rho_{s,t}^{d}$ is released into the main supplier at the beginning of the period, which results in the production throughput of $x_{s,t}^{d}$. This throughput is less than the maximum capacity due to queuing effects, resulting in work in process inventory, $\omega_{s,t}^{d}$. In the case of minor disruption occurrences, the strategic stock ν which is provided at the beginning of the planning horizon can cover the losses. When the main supplier fails due to a major disruption, the back-up supplier is used to supply the required flow of material (Tomlin 2006). If the demand is not met within its period, it is considered as lost, represented by $l_{s,t}$.

The back-up supplier increases its capacity to meet the plant demand once a major disruption occurs. The target capacity $\xi_{s,t}$ is gradually achieved within the response time because of the non-steady production during this period (Matta et al. 2007). Therefore a random fraction of the target capacity is available during the response period. In addition to these, shifting the demand to the back-up supplier when it is not fully capable of producing at the required rate during the response time can create an overflow of material. This congestion would decrease the throughput during the response time due to the increase in the lead time. Although there is a broad body of literature about reducing the response time (Koren et al. 1999; Matta et al. 2007) including reduction methodologies (Terwiesch and Bohn, 2001) there is no work which quantifies the importance of considering the response time characteristics at the decision-making stage.

The available capacity of the back-up supplier within the response time is important, since the supply chain incurs shortage costs if the available capacity level during this period is lower than the required capacity. Furthermore, the amount of the available capacity during the response time depends on the back-up supplier’s machine configuration (Wang and Koren, 2012). In addition to this, the strategic stock could be used to cover the losses during minor disruptions as well as the response time (Hopp and Yin, 2006; Schmitt 2011). Therefore, the strategic stock level and the back-up supplier machine configuration should be determined at the design stage of the supply chain with respect to the operational costs of holding, initial investment and shortage in order to have a robust supply flow. Note that the selected configurations would remain fixed during the planning horizon while the capacity might change upon the realization of the different disruption scenarios.

4.1 Solution Methodology

In order to design the robust supply flow, a two-stage multi period robust optimization model (RO) is presented. The set of first-stage decision variables includes the level of strategic stock and the response speed level of the back-up supplier. The set of second-stage decision variables, corresponding to scenario s are levels of back-up supplier production, inventory and lost demand. The list of notations and decision variables are shown in Table 4.1 and

Table 4.2.

Table 4.1 List of notations

Indices	
t	Current time
d	Main supplier
r	Back-up supplier
j	Speed level
i	Level of available capacity during response time
s	Scenario
Input parameters	
$T=\{1,2,...,T'\}$	Planning horizon consisting of T' periods

M	A big number
D_t	Demand at time t
w	Production cost of back-up supplier
o	Shortage cost
A	Strategic stock investment cost
H	Holding cost
C_d	Maximum capacity of main supplier
C_r	Maximum capacity of back-up supplier
$G_{s,t}$	1 if major disruption at time t scenario s, 0 otherwise
$F_{s,t}$	1 if minor disruption at time t scenario s, 0 otherwise
B	Intensity of the minor disruption $\in$ (0,1)
C_j	Investment cost of speed level j
$K_{s,t}^{i}$	1 if available capacity level of i is realized during the response time at time t scenario s, 0 otherwise
$U_{j,i}$	The fraction of the added capacity which is available for speed level j and realized available capacity level i
λ	Goal programming parameters $\lambda \geq 0$
P_s	Probability of scenario s
I	Set of available capacity levels during response time
J	Set of speed levels
S	Set of plausible future scenarios

Table 4.2 List of decision variables

Decision variables	
$x_{s,t}^{d}$	Main supplier production at time t, scenario s
$x_{s,t}^{r}$	Back-up supplier production at time t, scenario s
Δ_s	The variation cost of the scenario s
v	Initial strategic stock
$v_{s,t}$	Strategic stock level at time t, scenario s
$i_{s,t}$	Strategic stock level at the end of a disruption period at time t, scenario s
$\omega_{s,t}^{d}$	Main supplier work in process at time t, scenario s
$\omega_{s,t}^{r}$	Back-up supplier work in process at time t, scenario s
$\rho_{s,t}^{d}$	Amount of raw material released to main supplier at time t, scenario s
$\rho_{s,t}^{r}$	Amount of raw material released to back-up supplier at time t, scenario s
$\tau_{s,t}$	Back-up supplier actual capacity at time t, scenario s
$\xi_{s,t}$	Back-up supplier nominal capacity at time t, scenario s
$u_{s,t}$	Back-up supplier available capacity during the response time at time t, scenario s
$l_{s,t}$	Lost demand at time t, scenario s
y_j	1 if speed level j is selected; 0 otherwise
$b_{s,t}$	1 if demand loss exists at time t, scenario s; 0 otherwise
$q_{s,t}$	1 if there is capacity addition at time t, scenario s; 0 otherwise

$z_{s,t}$	1 if there is capacity deletion at time t, scenario s; 0 otherwise

4.1.1 Robust Stochastic Optimization Model

The objective function Z includes the investment cost of supply chain mitigation strategies, the expected cost of recourse actions with respect to the plausible future scenarios and the expected variation cost of worst case scenarios.

$$Minimize\ Z = \sum_{j \in J} C_j y_j + Av + \sum_{s \in S} P_s \{ \sum_{t \in T} (Hv_{s,t} + Wx^r_{s,t} + Ol_{s,t}) \} + \lambda \sum_{s \in S} P_s \Delta_s \tag{50}$$

The decision variables to be identified at the design stage of supply chain are presented in the first two terms in (50). It includes a binary variable y_j representing different configurations of the back-up supplier $j \in J$ and the level of strategic stock v. The investment costs associated with these decisions are represented by C_j and A respectively. The decisions made in the first stage result in recourse actions with corresponding costs expressed in the third term in (50). These costs include the holding cost of strategic stock level $v_{s,t}$ in period t at the rate of H per unit, production cost of satisfying $x^r_{s,t}$ units of the demand by back-up supplier at a production cost of W per unit and shortage cost of $l_{s,t}$ units at the rate of O per unit. All these cost parameters are expressed in \$/unit/period. In order to achieve a robust supply chain design that performs efficiently under the occurrence of worst case scenarios, the forth term in (50) is incorporated in the objective function. This term identifies the expected difference between the cost of the worst case scenarios and a predetermined threshold. List et al. (2003) incorporate a preselected threshold in their variability measure approach called upper partial moment (UPM).

In this Chapter, we use the expected cost as the threshold to avoid the efforts required to select the threshold in advance (Kazemi Zanjani et al. 2010). The relative importance of the expectation and variation costs is controlled by the parameter λ. For λ = 0, the model would be reduced to a two-stage stochastic optimization model. As λ increases, it reflects the decision maker's risk aversion level (List et al., 2003). The worst case scenarios are identified in (51) as the scenarios in which the total cost of the recourse actions is higher than the expected cost of the recourse actions. This difference is measured by Δ_s where any scenario with a variation above the expected costs is penalized in the objective function.

$$\sum_{t\in T}(Hv_{s,t} + Wx^r_{s,t} + Ol_{s,t}) - \sum_{s\in S} P_s\{\sum_{t\in T}(Hv_{s,t} + Wx^r_{s,t} + Ol_{s,t})\} \leq \Delta_s \qquad \forall\, s \in S \qquad (51)$$

In addition to robustness related constraints, the system constraints should be introduced. Among the possible speed levels in set J, only one should be selected for the back-up supplier (52). The inventory flow balance equation is represented in constraint (53): The demand at any period could be satisfied by the main supplier and/or back-up supplier and through strategic stock in case of a disruption at the main supplier. Unmet demand is assumed to be lost. The disruption scenarios are identified in the model through the random parameters $G_{s,t}$ and $F_{s,t}$ representing major and minor disruptions respectively. The strategic stock level at the end of current period is equal to the previous period if there is no disruption in the current period as indicated in (54). The strategic stock level is selected at the beginning of the planning horizon as part of the first stage decision variables (55). At the end of any period, there could be either strategic stock left or product shortage which is ensured by constraints (56**)** and (57**)**. The product shortage occurs only in the period with disruption (58).

$$\sum_{j\in J} y_j = 1 \qquad (52)$$

$$D_t = x^d_{s,t} + x^r_{s,t} + l_{s,t} + (G_{s,t} + F_{s,t})(v_{s,t-1} - i_{s,t}) \qquad \forall\, s \in S \quad \forall\, t \in T \qquad (53)$$

$$v_{s,t} = (1 - G_{s,t} - F_{s,t})v_{s,t-1} + (G_{s,t} + F_{s,t})i_{s,t} \qquad \forall\, s \in S \quad \forall\, t \in T \qquad (54)$$

$$v_{s,0} = v \qquad \forall\, s \in S \qquad (55)$$

$$l_{s,t} \leq Mb_{s,t} \qquad \forall\, s \in S \quad \forall\, t \in T \qquad (56)$$

$$v_{s,t} \leq M(1 - b_{s,t}) \qquad \forall\, s \in S \quad \forall\, t \in T \qquad (57)$$

$$l_{s,t} \leq M(G_{s,t} + F_{s,t}) \qquad \forall s \in S \quad \forall t \in T \qquad (58)$$

In order to demonstrate the response time characteristics, we represent two notations for the capacity level defined as the nominal capacity and the actual capacity. The nominal capacity determines the amount of capacity that the system is set to reach for the following period, expressed by constraint (59). The actual capacity represents the amount of capacity that is available (60). This will be less than the nominal capacity during the response time since some portion of the added capacity is lost. The available capacity during the response time $u_{s,t}$ is bounded by the random fraction $U_{j,i}$ of the added capacity where $U_{j,i}$ depends on the back-up supplier configuration y_j and the level of the capacity availability realized during response time $K^i_{s,t}$, indicated in (61). Constraint (62) states that either capacity addition or removal is allowed at any period with capacity change. The back-up supplier removes its capacity once the main supplier recovers from a major disruption (63). The back-up supplier production at any period is less than the actual capacity (64).

$$\xi_{s,t} = \xi_{s,t-1} + (q_{s,t} - z_{s,t})C_r \qquad \forall s \in S \quad \forall t \in T \qquad (59)$$

$$\tau_{s,t} = \xi_{s,t-1} + u_{s,t} - z_{s,t}C_r \qquad \forall s \in S \quad \forall t \in T \qquad (60)$$

$$u_{s,t} \leq (\sum_{i \in I} K^i_{s,t} U_{j,i})C_r q_{s,t} + M(1 - y_j) \qquad \forall j \in J \quad \forall s \in S \quad \forall t \in T \qquad (61)$$

$$q_{s,t} + z_{s,t} \leq 1 \qquad \forall s \in S \quad \forall t \in T \qquad (62)$$

$$\xi_{s,t} \leq G_{s,t}C_r \qquad \forall s \in S \quad \forall t \in T \qquad (63)$$

$$x^r_{s,t} \leq \tau_{s,t} \qquad \forall s \in S \quad \forall t \in T \qquad (64)$$

Constraints (65) and (66) represent the workload balance equations. The work in process inventory (WIP) consists of the jobs in the queue or under operation. The maximum workload in any period is bounded by the available capacity during that period, since the utilization of a resource cannot exceed 100% as indicated in (67) and (68). Furthermore, the impact of minor disruptions over the main supplier's capacity is represented in (67). There will be no material release in the main supplier when it is down due to a major disruption (69).

$$\omega_{s,t}^{d} = \omega_{s,t-1}^{d} + \rho_{s,t}^{d} - x_{s,t}^{d} \qquad \forall s \in S \quad \forall t \in T \qquad (65)$$

$$\omega_{s,t}^{r} = \omega_{s,t-1}^{r} + \rho_{s,t}^{r} - x_{s,t}^{r} \qquad \forall s \in S \quad \forall t \in T \qquad (66)$$

$$\omega_{s,t-1}^{d} + \rho_{s,t}^{d} \le (1 - F_{s,t} B) C_d \qquad \forall s \in S \quad \forall t \in T \qquad (67)$$

$$(\omega_{s,t-1}^{r} + \rho_{s,t}^{r}) G_{s,t} \le \tau_{s,t} \qquad \forall s \in S \quad \forall t \in T \qquad (68)$$

$$\rho_{s,t}^{d} \le M(1 - G_{s,t}) \qquad \forall s \in S \quad \forall t \in T \qquad (69)$$

At the beginning of a major disruption, shifting the total demand to the back-up supplier when it is not 100% available will create an overflow. The resulting queue built up will increase the lead time. This congestion would limit the amount of the back-up supply during the response time. In order to determine the appropriate strategic stock and the response speed level of the back-up supplier, the impact of congestion, especially during the response period should, be considered.

To this end, we present the suppliers as a single server system with Poisson arrivals and general service time distribution (M/G/1 system). The relationship developed using this model allows developing the clearing function (Missbauer 2002). Based on the clearing function in (70), the expected system throughput $E(x_t)$ in any period is a function of the expected workload $E(\omega_{t-1} + \rho_t)$, available capacity τ and parameter k. This parameter is defined in (71) based on the mean $\frac{1}{\mu}$ and the variance $\frac{1}{\sigma^2}$ of the processing time.

$$E(\mathrm{x}_t) = \frac{1}{2}\left[\tau + k + E(\omega_{t-1} + \rho_t) - \sqrt{\tau^2 + 2\tau k + k^2 - 2\tau E(\omega_{t-1} + \rho_t) + 2kE(\omega_{t-1} + \rho_t) + E(\omega_{t-1} + \rho_t)^2}\right] \qquad (70)$$

$$k = \frac{\mu\sigma^2}{2} + \frac{1}{2\mu} \qquad (71)$$

The clearing function in (70) is concave and nonlinear (Missbauer 2002). An outer approximation approach has been implemented to generate a set of lines in order to linearize the robust model. Based on this linearization approach, the throughput of each supplier can be represented depending on their states.

There exist three states for the main supplier. The production of the main supplier in the normal periods could not be more than the expected throughput, which is estimated by its clearing function (72). The set of lines N^d represents the clearing function where A_η is the slope and B_η is the constant value of the line η. In the periods with minor disruption, the set of lines N_F^d in (73) represents the main supplier's throughput due to the reduction in the service rate of the main supplier. In the case of a major disruption occurrence, there is no production at the main supplier incorporated by the term $G_{s,t}$ in (72).

In cases of major disruptions, the backup supplier must ramp up its capacity. In order to show the congestion effect over the back-up supplier production during the response time, the clearing function is represented through a set of planes $N_{j,i}^r$ in (74). This is due to the fact that the change in capacity requires defining the throughput during the response time as a function of both workload and actual capacity which result in a three dimensional clearing function (Ebrahim Nejad et al., 2014). Note that the binary variables y_j and $K_{s,t}^i$ in (74), activate the clearing function associated with selected speed level of j and realized available capacity level of i. After the response time, the back-up supplier can operate at its predetermined capacity level. The impact of congestion over the production of back-up supplier in these periods is illustrated by set of lines N^r in (75). The non-negativity constraints are stated in (76) to (78). The non-anticipativity constraints are in (79).

$$x_{s,t}^d \le (A_\eta(\omega_{s,t-1}^d + \rho_{s,t}^d) + B_\eta + MF_{s,t})(1 - G_{s,t}) \qquad \forall\ \eta \in N^d,\ \forall\ s \in S,\ \forall\ t \in T \quad (72)$$

$$x_{s,t}^d \le A_\eta(\omega_{s,t-1}^d + \rho_{s,t}^d) + B_\eta + M(1 - F_{s,t}) \qquad \forall\ \eta \in N_F^d,\ \forall\ s \in S,\ \forall\ t \in T \quad (73)$$

$$x_{s,t}^r \le A_\eta(\omega_{s,t-1}^r + \rho_{s,t}^r) + B_\eta \tau_{s,t} - G_\eta + M(3 - y_j - K_{s,t}^i - q_{s,t}) \quad \forall j \in J,\ \forall\ \eta \in N_{j,i}^r,\ \forall\ s \in S,\ \forall\ t \in T \quad (74)$$

$$x_{s,t}^r \le A_\eta(\omega_{s,t-1}^r + \rho_{s,t}^r) + B_\eta + Mq_{s,t} \qquad \forall\ \eta \in N^r,\ \forall\ s \in S,\ \forall\ t \in T \quad (75)$$

$$x_{s,t}^{d} \geq 0,\ x_{s,t}^{r} \geq 0,\ l_{s,t} \geq 0,\ v_{s,t} \geq 0,\ i_{s,t} \geq 0,\ \omega_{s,t}^{d} \geq 0,\ \omega_{s,t}^{r} \geq 0,$$
$$\rho_{s,t}^{d} \geq 0,\ \rho_{s,t}^{r} \geq 0,\ \tau_{s,t} \geq 0,\ \xi_{s,t} \geq 0,\ u_{s,t} \geq 0,\ b_{s,t} \in \{0,1\},\ q_{s,t} \in \{0,1\},\ z_{s,t} \in \{0,1\} \qquad \forall s \in S \qquad \forall t \in T \qquad (76)$$

$$\Delta_s \geq 0 \qquad \forall s \in S \qquad (77)$$

$$y_j \in \{0,1\},\ v \geq 0 \qquad (78)$$

$$x_{s,t}^{d} = x_{s',t}^{d},\ x_{s,t}^{r} = x_{s',t}^{r},\ l_{s,t} = l_{s',t},\ v_{s,t} = v_{s',t},\ i_{s,t} = i_{s',t},\ \omega_{s,t}^{d} = \omega_{s',t}^{d},\ \omega_{s,t}^{r} = \omega_{s',t}^{r},$$
$$\rho_{s,t}^{d} = \rho_{s',t}^{d},\ \rho_{s,t}^{r} = \rho_{s',t}^{r},\ \tau_{s,t} = \tau_{s',t},\ \xi_{s,t} = \xi_{s',t},\ u_{s,t} = u_{s',t},\ b_{s,t} = b_{s',t},\ q_{s,t} = q_{s',t},\ z_{s,t} = z_{s',t} \qquad \forall s, s' \in \{S\}_T \qquad (79)$$

4.1.2 Scenario generation

In this section, we explain how the scenarios resulting from disruptions and the random available capacity during the response time can be generated. In order to identify the scenarios within the planning horizon, we incorporate a scenario tree. We define a scenario as the states of the supply flow within the planning horizon. The flow can be provided from the main supplier or back up supplier. The main supplier may be completely or partially available due to a minor disruption. Once the main supplier becomes unavailable due to a major disruption, the back-up supplier would resume the supply flow. Different levels of the available capacity can be realized in the first period of back-up supply due to inherent randomness. We categorize the available capacity during the response time to three different levels: high, normal and low. Furthermore, we assume that the back-up supply is completely available after the response time.

Figure 4.1 represents a snapshot of the scenario tree at the first two periods of the planning horizon. Each node represents a possible status of the supply flow at a given period. The set of supply status includes main supplier fully operational, main supply partially operational, back-up supplier during response time - high capacity availability, back-up supplier during the response time – normal capacity availability, back-up supplier during response time – low capacity availability and back-up supplier fully operational after the response time. Based on the stated assumptions, the following steps are conducted in order to generate scenarios.

- Since the problem considered is a strategic supply chain network design, each period is long enough for the main supplier to recover from a minor disruption or the back-up supplier to reach the full required capacity level.
- The length of a minor disruption is one period.
- The length of a major disruption is not limited and it may last for the whole planning horizon.
- The length of the response time is one period.
- After a major disruption, the first period represents the response time.
- After the response time, the back-up supplier is fully utilized for at least one period.
- After a minor disruption, the main supplier could be fully operational or it may get disrupted as a result of a major disruption.
- Once the major disruption ends, the main supplier becomes fully operational and it provides the required supply for at least one period.

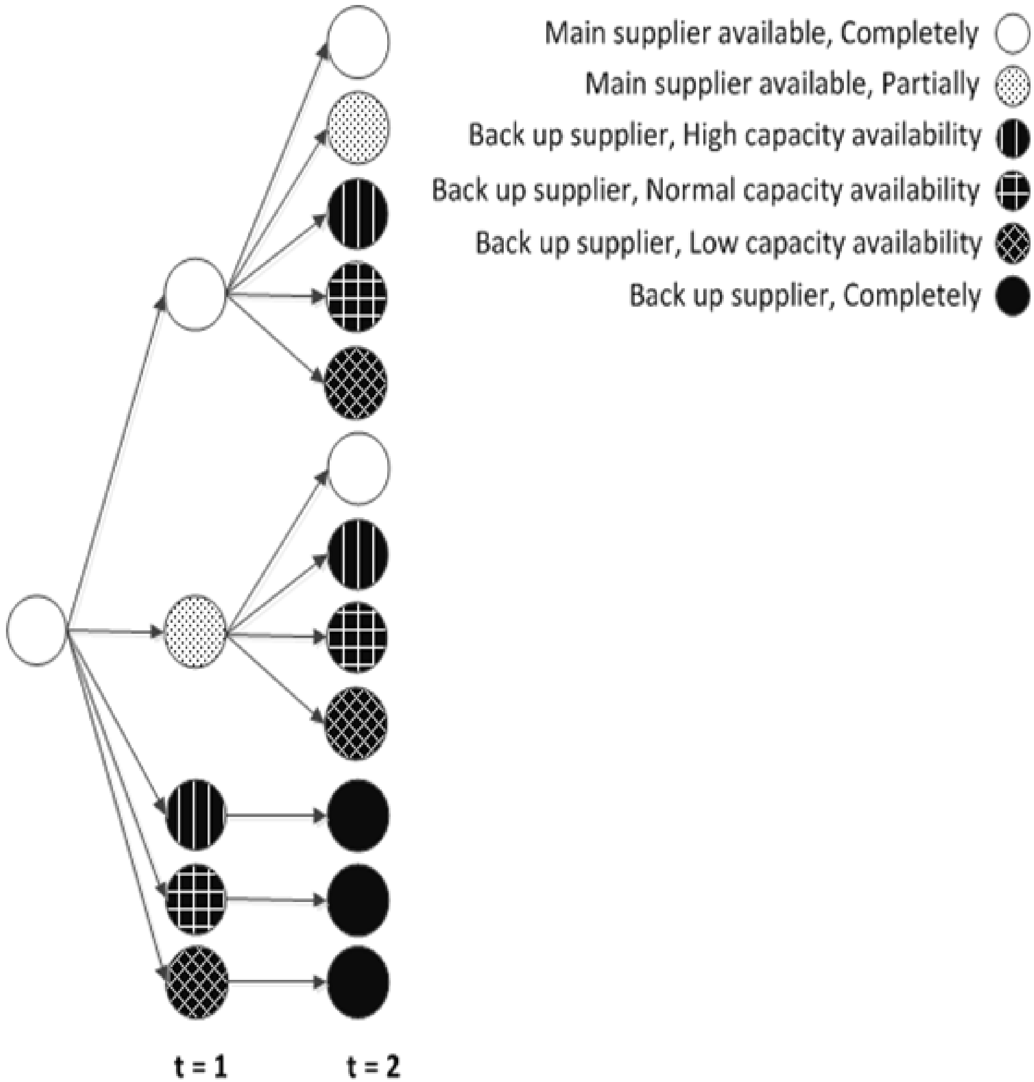


Figure 4.1 A snapshot of the scenario tree

The scenario tree that includes all the possible realizations for major and minor disruptions for multiple periods can significantly increase the required number of constraints in the robust optimization model, resulting in difficulties in solving the problem in a reasonable amount of time. In order to reduce this complexity, we make the following assumptions regarding the realizations of disruptions and recovery strategies without the loss of generality as indicated in Figure 4.1: Once a major disruption occurs, it lasts for at least three consecutive periods and the maximum length of the minor disruption could be one period. Moreover, the probability of each scenario is computed through the product of the transition probabilities of the states. These transition probabilities can be found in Appendix C, Table C.1.

In order to investigate the impact of considering the response time characteristics in the optimal selection of the strategic stock level and the response speed level of the back-up supplier, an illustrative example is presented in section 4.2. Furthermore, a sensitivity analysis is conducted on λ in order to evaluate the impact of different tolerance level of the decision maker towards risk.

4.2 Numerical Experiments

We consider the supply chain associated with a product whose lifecycle lasts for six periods where each period represents two months. Based on this assumption and with respect to the minor and major disruptions and three levels of available capacity during the response time, we identify 162 scenarios over the planning horizon. Furthermore, the demand is assumed to be deterministic and 1800 units per period. Three different speed levels are presented as decision variables: fast, medium and slow which represents manufacturing system configurations of parallel, parallel-serial and serial configurations accordingly. While faster response configuration will perform better, there are investment costs corresponding to the selected response speed and the level of initial strategic stock. The cost parameters which are included in the objective function are defined in Table 4.3.

Table 4.3 Cost parameters

Cost Parameters	***Value***	***Unit***
Investment cost of fast speed	135,000	$
Investment cost of medium speed	90,000	$
Investment cost of slow speed	45,000	$
Strategic stock investment cost	180	$/unit
Holding cost	40	$/unit/period

Production cost of backup supplier	125	$/unit/period
Shortage cost	300	$/unit/period

The main supplier has a capacity of 2200 units while the backup supplier provides 2000 units when the main supplier is breakdown. The fractions of the added capacity which are available during the response time with respect to different scenarios are presented in Table 4.4.

Table 4.4 Fractions of the added capacity available in different scenarios

Level of available capacity during response time	***Response speed level***		
	Fast	Medium	Slow
High	0.933	0.867	0.8
Normal	0.833	0.667	0.5
Low	0.733	0.467	0.2

The fraction of added capacity available during response time increases as higher speed level is selected or a higher level of available capacity is realized during response time. Furthermore, we assume the intensity of the minor disruptions over the main supplier results in 30% loss in the maximum capacity (Wagner and Bode, 2008). Based on the stated assumptions and inputs, the following experiments are conducted. First we evaluate the impact of considering the random capacity availability during the response time in the design of the robust supply flow. We implement this evaluation for different values of the parameter λ. We measure the value of this consideration through the difference in the objective functions. Second, we assess the importance of considering the congestion in the decision process. The results show whether there is any benefit for decision makers to incorporate the response time characteristics in the design of the robust supply flow. The proposed robust optimization model has been implemented in ILOG

CPLEX 12.5. By setting the desired optimality gap to 0.0001, the results have been obtained with an average computation time of 94 seconds.

4.2.1 The impact of considering the available capacity randomness

In order to observe the impact of considering the randomness related to the available capacity during the response time in the design of a robust supply flow, we first execute the robust model (RO) presented in section 4.1.1. Second, we remove the random parameters $K^{i}_{s,t}$ which represent the realized level of the available capacity during response time in constraint (61). In addition to this, we replace the random parameters $U_{j,i}$ by its expected value U_j. We run the modified robust model (MRO). Note that by removing the randomness associated with the available capacity, the number of the scenarios is reduced to the 68 scenarios. We conduct four different experiments representing different probability distributions for the available capacity levels during response time as presented in Table 4.5. The probabilities of the minor disruption, major disruptions and the recovery from major disruptions are set to 0.1, 0.01 and 0.2 respectively in all experiments.

Table 4.5 Experiments based on the probabilities of available capacity levels during response time

Experiment	P_H	P_N	P_L
1	0.6	0.2	0.2
2	0.333	0.333	0.333
3	0.2	0.6	0.2
4	0.2	0.2	0.6

The first experiment represents the situation where it is more probable to observe high level of the available capacity during response time. This can represent a flexible supplier with better responsiveness capabilities which enable a smooth transition. On the other hand, experiment 4 identifies a supplier that is inexperienced; susceptible to increased problems during ramp up. In order to have a robust supply flow, the required strategic stock level and the response speed of the back-up supplier in experiment 1 are presented in Figure 4.2. The slow response speed is identified as the optimal response speed for different values of λ. However; higher strategic stock levels are required as λ increases. This is because of the fact that as λ increases, the model would give more emphasis on minimizing the cost of the worst case scenarios. The results of RO and MRO model are identical in selecting the response speed, although RO model requires lower strategic stock level compared to the MRO model. The reason is due to the fact that MRO solutions are based on a fixed value of the available capacity during the response time. On the other hand, the RO solutions consider all plausible scenarios of available capacity levels during response time. Since the scenarios with high level of added capacity available during response time are more probable in experiment 1, the RO model selects lower strategic stock level.

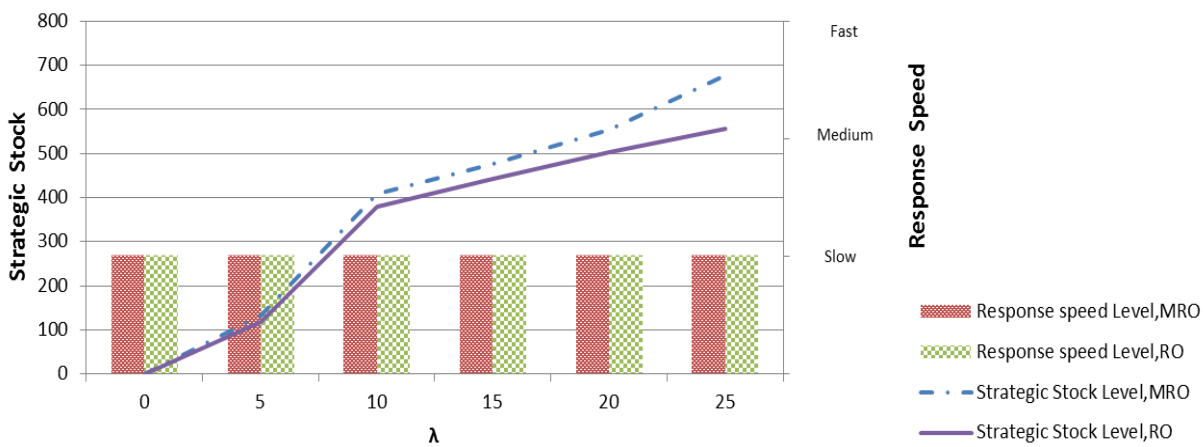


Figure 4.2 Robust design of the supply flow, Experiment 1

The assignment of equal probabilities to the different levels of available capacity during response time provides the results of experiment 2 as shown in Figure 4.3. As illustrated, the higher levels of the response speed and strategic stock are selected compared to the experiment 1. This is due to the reduction in the probability of high level of available capacity during the response time and the increase in the probabilities of normal and low levels. Furthermore, the selection of strategic stock and response speed levels in the RO model become close to MRO model unlike experiment 1. This is a result of having equal probabilities of available capacity during the response time. Therefore, the solution of the RO model is not influenced by any level of the available capacity significantly. The RO model selects lower level of strategic stock in λ =10, 15, 25 and slower speed in λ = 20.

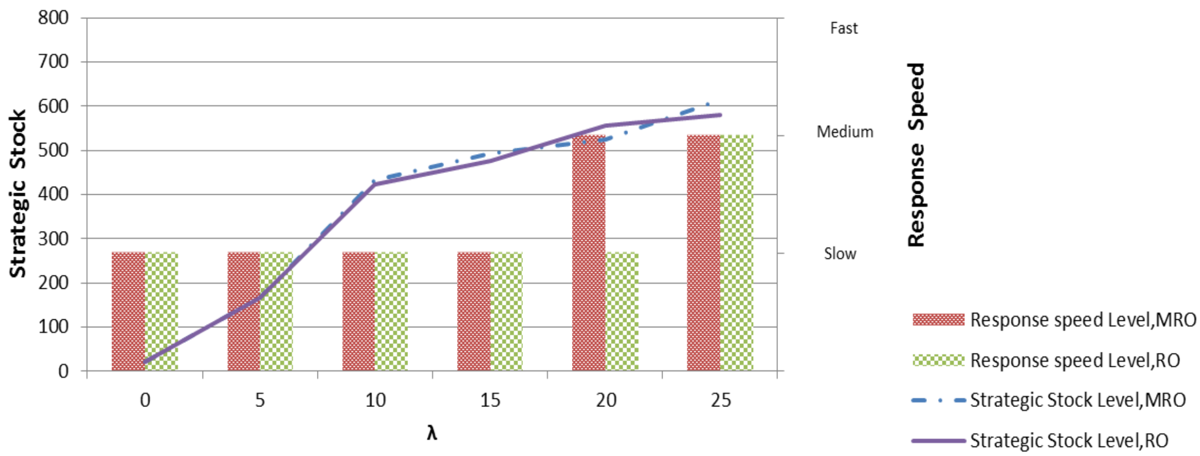


Figure 4.3 Robust design of the supply flow, Experiment 2

The designs of robust supply flow with more chance of having normal level of available capacity during response time are presented in Figure 4.4. More strategic stock level is required compared to the experiment 2 due to the reduction in the probability of having high level of available capacity during response time. On the other hand there is almost no change in the selection of the response speed (except RO model in λ = 20) due to the reduction in the probability of having low level of available capacity during response time. In addition to these, the results of the RO and MRO model are almost identical. This is because of the fact that; MRO solutions are achieved with respect to the expected level of available capacity during response time and RO solutions in this experiment are obtained by assigning more probability to normal level of available capacity during the response time. Since the normal level of capacity availability is equivalent to the expected level of capacity availability, identical solutions are observed.

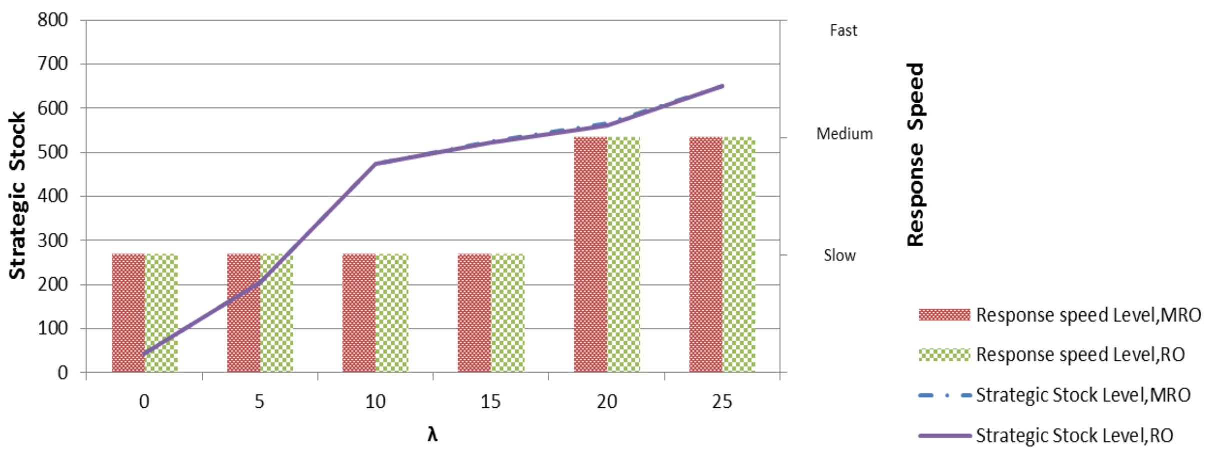


Figure 4.4 Robust design of the supply flow, Experiment 3

As presented in Figure 4.5, more strategic stock and faster speed level are required to achieve a robust supply flow as the probability of having low level of available capacity during the response time increases. In addition to this, the RO model selects higher strategic stock level for λ less than 15 and faster speed level for λ more than 10 compared to the MRO model. This is due to increased chance of having low level of available capacity during the response time.

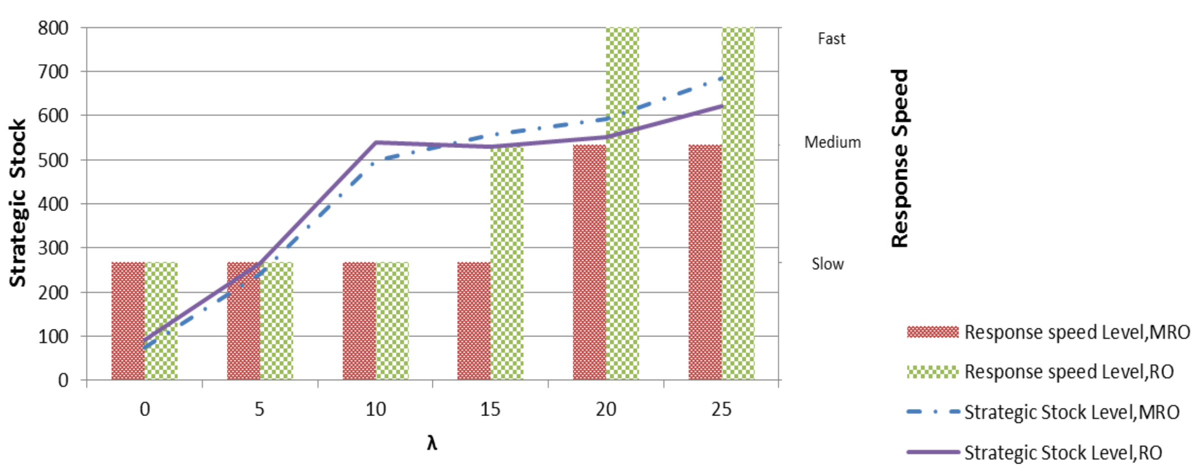


Figure 4.5 Robust design of supply flow, Experiment 4

In order to evaluate the value of incorporating the randomness associated with the available capacity during the response time in the decision making stage, we compare the objective function of the RO model versus the case when the solution of the MRO model is used to compute its performance in RO model. The savings which would be achieved by considering the randomness related to the available capacity during the response time are presented in Table 4.6.

Table 4.6 % Reduction in the objective function by RO solution

	λ					
Experiment	0	5	10	15	20	25
1	0%	0.40%	1.10%	1.60%	2.90%	4.10%
2	0%	0%	0.35%	0.44%	3.90%	0.66%
3	0%	0%	0.03%	0.03%	0.12%	0.05%
4	2.30%	3.30%	4.60%	6.30%	8.50%	11.30%

The results show that the improvement in the objective function is higher in experiments1 and 4 by considering the randomness of the available capacity during the response time. In experiment 1, the reduction is achieved as a result of lower strategic stock level in RO model compared to the MRO, reducing the strategic stock investment and holding costs. On the other hand, the savings in experiment 4 are due to the significant reduction in shortage cost since the RO model select higher strategic stock level and/or response speed compared to the MRO model. The reductions are low in experiments 2 and 3 since the results of the RO and MRO model are close to each other.

4.2.2 The importance of incorporating congestion

In order to assess the effect of considering the congestion impact in the robust design of the supply flow, we remove the constraints (72) to (75) which represent the clearing functions in the RO model. Through these changes, the model would ignore the impact of workload accumulation on production capability. We call the modified model as load independent robust model (LIRO). Figure 4.6 represents the results of RO and LIRO model with respect to the experiment 2.

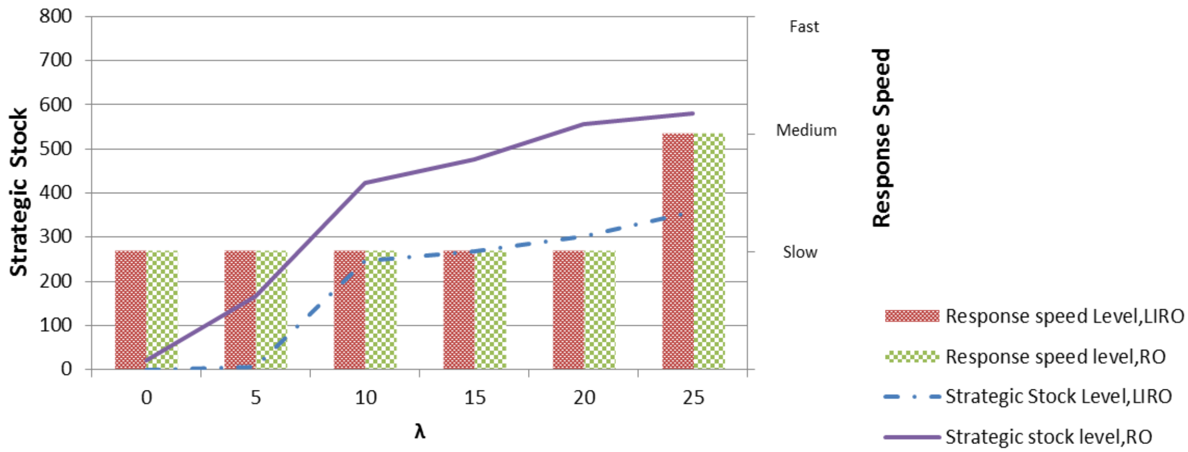


Figure 4.6 Robust design of the supply flow, *RO VS LIRO*, Experiment 2

Ignoring the impact of congestion leads to the underestimation of the required strategic stock level with no changes in the response speed. This would result in degraded performance of the robust solution. We measure the benefit of incorporating congestion as follows: For each scenario, the clearing functions present the actual production quantity of the main and backup suppliers based on the capacity and WIP levels which have been determined in the LIRO model. The actual demand losses are then computed by the difference between demand and actual production quantities obtained using the clearing function. This will give the actual performance of the supply chain if the solution generated from LIRO is implemented. Table 4.7 indicates the reductions achieved in the objective function. This reduction is mainly the result of increased strategic stock levels obtained from RO to cover the impact of congestion.

Table 4.7 Improvement in total costs by considering congestion effects

	λ					
	0	5	10	15	20	25
LIRO	$118,003	$386,595	$573,948	$736,407	$890,307	$1,024,103
RO	$114,226	$367,677	$543,939	$693,298	$834,834	$960,652
% Reduction	3.20%	4.89%	5.23%	5.85%	6.23%	6.20%

The results in Table 4.7 show that incorporating the congestion impact in the decision making process reduces the costs. However, this improvement is decreased for λ greater than 20. This is due to the fact that as the decision maker becomes more risk averse, the differences between strategic stock level of LIRO and RO model decreases as result of focusing on the smaller portion of the worst case scenarios albeit with higher impacts.

4.2.3 Discussion of results

The first section of the numerical experiments evaluates the impact of considering the randomness associated with available capacity during the response time in the design of the supply flow. Two options to represent the available capacity during the response time are investigated. The first option presents the random available capacity by considering different scenarios while the second one considers only the expected available capacity.

For the situation where it is more probable to observe high level of the available capacity during response time, the approach based on considering the random available capacity selects lower level of strategic safety stock. This reduces the supply chain investment cost in strategic safety stock and inventory holding cost albeit with no products shortages since the supply chain is prone to the high level of capacity availability during the response time. On the other hand, there are higher level of strategic safety stock and faster response speed selected as the chance of exposure to the low level of available capacity during the response time increases. This results in lower shortage cost since there are sufficient buffers incorporated in the supply chain structure in order to cover the low level of capacity availability during the response time. Therefore, there is reduction in the total cost of supply chain by considering the randomness associated with the available capacity during the response time in the design of the robust supply flow.

The second part of the numerical experiments investigates the value of considering the congestion impact during the response time when a combination of strategic safety stock and contingent sourcing is applied as supply chain risk mitigation strategies. The results show that the supply chain performance would benefit by considering the congestion impact. This is due to increase in the strategic stock level to cover the shortages which are created as a result of congestion.

The supply chain network design problems considered in Chapter 3 and 4 are assumed to have sufficient data to estimate the probabilities of plausible future scenarios. However, there might problems with limitation of no data availability to estimate the likelihood of future scenarios.

The next chapter proposes a novel approach to design a robust supply flow subject to deep uncertainty associated with operational risks and disruptions.

Chapter 5

A Clustering based Scenario Reduction Approach to Design Robust Supply Network under Operational Risks and Disruptions

In this Chapter, the former supply network design problem (Ebrahim Nejad and Kuzgunkaya, 2015) is considered under the condition in which it might be difficult to estimate the probability of operational risks and disruptions. Therefore, we aim to find an equitable solution by achieving an efficient performance with respect to all plausible future scenarios. For this purpose, we develop a decision making tool which focuses on solution robustness in identifying the supply chain design decisions. In order to address the computational complexity which may result from large scenario set, we propose a novel scenario reduction technique. Finally, we compare the performance of the proposed scenario reduction methodology both with respect to the quality of solutions and the computational time against other approaches available in literature.

We consider a single product supply chain that includes a manufacturing plant with dual sourcing. The main supplier of the manufacturing plant is cost-effective as a result of dedicated facilities though prone to disruptions during which it may be partially or completely unavailable. There is a back-up supplier located in a low-risk region that is available when the main supplier is unavailable. The back-up supplier has volume-flexible production facilities where it can scale up its capacity, however this scalability increases the production cost. Demand in the normal periods can be met by the main supplier. In the case of minor disruption occurrences, the strategic stock which is provided at the beginning of the planning horizon can cover the losses. When the main supplier fails due to a major disruption, the back-up supplier increases its capacity to meet the plant demand. However, the target capacity is gradually achieved within the recovery time because of the non-steady production during this period (Matta et al. 2007). Therefore a random fraction of the target capacity is available during the recovery time. In addition to these, shifting the demand to the back-up supplier when it is not fully capable of producing at the required rate during the recovery time can create an overflow of material. This congestion would decrease the throughput during the recovery time due to the increase in the lead time. The available capacity of the back-up supplier within the recovery time is important and it should be considered in the design stage, since the supply chain incurs shortage costs if the available capacity level during this period is lower than the required capacity. The amount of the available capacity during the recovery time depends on the back-up supplier's machine configuration (Wang and Koren, 2012). In addition to this, the strategic stock could also be used to cover the losses during the recovery time (Schmitt 2011). Therefore, the appropriate strategic stock level and machine configuration of the back-up supplier are important parameters in the proposed supply chain settings which should be considered in the strategic design stage of the supply chain network.

In order to achieve a robust supply chain network, the operational risks and disruptions should be considered in the strategic design stage of the supply chain network (Klibi et al. 2010). An instance of operational risks in the considered supply chain network is machine/equipment breakdown which decrease the production capacity of main supplier during normal periods and backup supplier within the recovery time. As opposed to operational risks, the disruptions significantly reduce the production capacity of main supplier when they occur. In addition to this, the information about disruptions occurrences and their corresponding magnitudes are typically hard to predict or maybe unavailable since they are rare. Therefore, the solution robustness is proposed as our supply chain network design performance measure. The solution robustness could be calculated independent of disruptions probabilities and it measures the difference in performance between the optimal solution and the solution provided by the robust optimization (Govindan et al. 2017). There are different approaches proposed in supply chain network design literature to achieve solution robustness (Kouvelis and Yu 1997, Roy 2010 and Kalai et al. 2012). In this Chapter, a lexicographic aggregator based approach will be applied to achieve solution robustness. This operator reorders the performance vector (e.g. regret) from the worst to the best. Next, it first minimizes the worst regret, then the second worst regret is minimized (provided that the first worst one is as small as possible), then the third worst regret is minimized (provided that the first two worst regrets remain as small as possible) and so on (Sawik 2014). This approach is known to be less conservative than traditional Minmax formulation since it evaluates all plausible future scenarios as opposed to only the worst case one. However, there is computational complexity challenge associated with lexicographic aggregator formulation as the size of the scenario set increases.

The supply chain network design problem considered in this Chapter includes a large scenario set. At one hand considering multiple random parameters including the occurrence of minor and major disruptions in the main supplier and the portion of the added capacity to the back-up supplier which is available during the recovery time and at the other hand, investigating the supply chain performance in a multi period planning horizon increases the size of the scenario set. Therefore, one solution is to decrease the size of the scenario set by selecting a few representative scenarios. The following section describes our solution methodology to identify the optimal strategic stock level and machine configuration of the backup supplier by incorporating a lexicographic aggregator approach to achieve solution robustness while reducing the size of the scenario set by applying a novel scenario reduction technique.

5.1 Solution Methodology

In order to find the optimal back-up supplier machine configuration and the level of strategic stock, a two-step solution methodology is proposed as illustrated in Figure 5.1. The first step is a Mixed Integer Programming (MIP) based scenario clustering model which reduces the set of plausible future scenarios into a smaller set by selecting the most representative scenarios. The second step is a MIP robust SCND model which is developed to achieve solution robustness when scenario probabilities are not available. This model identifies the machine configuration of the back-up supplier and the level of strategic stock by considering the representative scenarios which are achieved in step one. The selections of back-up supplier machine configuration and the level of strategic stock are based on the trade-off which exists between investment cost and operational cost of supply chain with respect to representative scenarios.

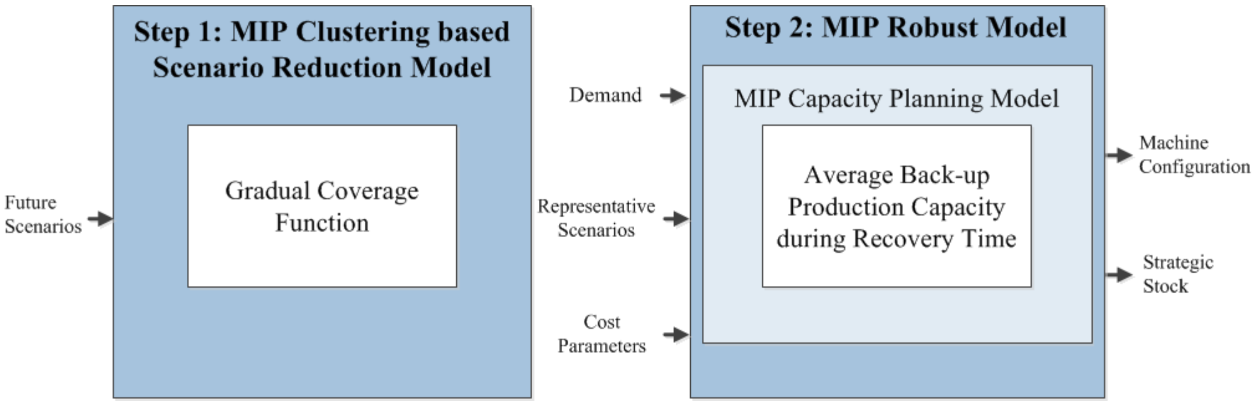


Figure 5.1 Solution Methodology

The MIP robust SCND model in step two is the stochastic version of an MIP deterministic capacity planning model. The contingency capacity plan of supply flow when the main supplier becomes disrupted is generated through MIP capacity planning model. In order to have an estimation of the available production capacity during the recovery time, the impact of work load accumulation over the system throughput is represented in the MIP capacity planning model. The following section presents the first step of the solution methodology which is the MIP clustering based scenario reduction model.

5.1.1 Clustering based MIP Scenario Reduction (CBSR) Model

The MIP robust SCND model may include a large number of random scenarios that makes it very hard to solve. To overcome this difficulty in such problems, some approximation methods have been presented in literature which reduces the dimension of the problem by determining a subset of the scenario set. Li and Floudas (2014) present an MIP scenario reduction model which minimizes the probabilistic distance between the original and reduced input scenario distribution. The probabilistic distance depends on scenario probabilities and distances between scenario values. The distance $d_{s,s'}$ between any two scenarios s, s' is measured based on the following metric:

$$d_{s,s'} = \sum_{q\in Q}\sum_{t\in T}\left|\theta_t^{s,q} - \theta_t^{s',q}\right| \quad (80)$$

Where $t \in T$ represents time, $q \in Q$ is the set of random parameters and $\theta_t^{s,q}$ is a binary variable which becomes 1 if random parameter q occurs at time t in the scenario s and 0 otherwise. Eventually, scenarios are deleted when they are close or have small probabilities. In order to reduce the number of scenarios in this Chapter, an MIP model is proposed which group scenarios into different clusters. This model applies the gradual coverage function of facility location problems in order to calculate the degree of membership of each scenario to the proposed clusters. The covering models in facility location problems follow a similar rationale to the clustering based scenario reduction methodologies (Farahani et al. 2012). For a given set of customer locations, these models identify locations of facilities such that the customer can receive service from each facility whose distance from customer is equal or less than a given critical distance. Among set covering models, there are formulations which allow customers to receive either full and/or partial coverage from single and/or multiple facilities. The gradual coverage formulation fully covers a demand point if its distance from closest facility is less than R^1 . If the distance is between R^1 and R^2 , the demand point will be partially covered and the coverage level provided by facility acts as a decreasing function of the distance from the facility to the customer's location. Finally, if the distance is more than R^2 , the demand point will never be covered (Berman and Krass, 2002). Gendreau et al. (1997) develop a double coverage model based on two coverage radius R^1 and R^2 ($R^1 < R^2$). All demand points must be covered within R^1 and a portion of demands, say α must be covered within R^2. The benefit of partial coverage and/or coverage with multiple facilities is in the lower number of facilities required to cover the set of customers compared to the case where only full and/or single coverage is allowed.

The proposed clustering based MIP scenario reduction model in this Chapter incorporates the partial and multiple coverage techniques. This model groups scenarios into clusters. Next, it identifies cluster center associated with each cluster. A cluster center is a scenario which represents scenarios within that cluster. Finally, cluster centers would represent the reduced scenario set and the remaining scenarios would be eliminated. The objective of this model is to minimize the distance of cluster centers from eliminated scenarios while reducing the size of original scenario set to a desirable level. This model applies the following three steps to group scenarios which are based on the gradual coverage formulation. The distance between scenarios is measured by the metric proposed by Li and Floudas (2014).

I. The scenario s will be fully covered by scenario s' if the distance between scenario s and s' is less than a primary admissible tolerance.

II. The scenario s will be partially covered by scenario s' if the distance between scenario s and s' is more than primary admissible tolerance and less than a secondary admissible tolerance. The amount of partial coverage is a decreasing function of the distance from the scenario s to the scenario s'.

III. The scenario s will not covered by scenario s' if the distance between them is more than the secondary admissible tolerance.

The following tables represent the parameters and variables incorporated into the model.

Table 5.1 List of Decision Variables in Clustering based MIP Scenario Reduction Model

Decision Variables	Definition
$X_{s'}$	1 if scenario s' is a cluster center, else 0
$Y_{ss'}$	1 if scenario s is covered by scenario s', else 0

Table 5.2 List of Input Parameters in Clustering based MIP Scenario Reduction Model

Input Parameters	Definition
S	The set of all scenarios
s, s'	Index of scenarios $s, s' \in S$

$w_{ss'}$	Coverage level of scenario s by scenario s' $\begin{Bmatrix} 1 & if\ d_{ss'} \le l_s \\ f_{ss'} & if\ l_s < d_{ss'} \le u_s \\ 0 & if\ d_{ss'} > u_s \end{Bmatrix}$
$d_{ss'}$	distance between scenario s and s'
l_s	the primary admissible tolerance for scenario s
u_s	the secondary admissible tolerance for scenario s
$f_{ss'} = \frac{d_{ss'} - l_s}{u_s - l_s}$	the coverage decay function which represent the amount of partial coverage that scenario s' provides for scenario s
UB	upper bound on level of coverage per scenario across all cluster centers
LB	lower bound on level of coverage per scenario across all cluster centers

The objective function of the model minimizes the total distance between cluster centers and other members of clusters which are covered by cluster centers. The decision variable $Y_{ss'}$ determines whether scenario s is covered by scenario s' and the input parameter $w_{ss'}$ represents the portion of scenario s which could be covered by scenario s'.

$$Minimize \sum_{s' \in S} \sum_{s \in S} w_{ss'} d_{ss'} Y_{ss'} \tag{81}$$

The constraint (82) limits the size of reduced scenario set to the desirable level of N. The decision variable $X_{s'}$ determines whether scenario s' is a cluster center. The total number of cluster centers should be equal to N. The constraint (83) guarantees that scenario s will be assigned to scenario s' if scenario s' is identified as a cluster center.

Subject to:

$$\sum_{s' \in S} X_{s'} = N \tag{82}$$

$$Y_{ss'} \leq X_{s'} \qquad \forall s,\ s' \in S \qquad (83)$$

$$\sum_{s' \in S} w_{ss'} Y_{ss'} \leq UB \qquad \forall s \in S \qquad (84)$$

$$LB \leq \sum_{s' \in S} w_{ss'} Y_{ss'} \qquad \forall s \in S \qquad (85)$$

$$Y_{ss'} \leq w_{ss'} M \qquad \forall s, s' \in S \qquad (86)$$

The constraint (84) sets an upper bound on the overall level of coverage provided to a scenario through all cluster centers. The constraint (85) states that the summation of coverage given to any scenario should be more than LB. These two constraints control the trade-off between the accuracy of clustering and the computational time of model. For a given N, the higher values of UB and LB would result in reduced scenario set with closer profile to original scenario set however this will increase the processing time of the model. The constraint (86) guarantees that scenario s will be covered by scenario s' only if the distance between these two scenarios is lower than the secondary admissible tolerance u_s. The cluster centers are representative scenarios.

The MIP robust SCND model identifies the strategic design decisions by considering representative scenarios. It computes the operational cost of a design decision by generating the contingency capacity plans of representative scenarios. The following section describes the MIP contingency capacity planning model.

5.1.2 Capacity Planning Model

The production quantity and capacity levels of main and back-up suppliers, the levels of lost demand and strategic stock within the planning horizon are determined by MIP contingency capacity planning model. In addition to these, the model determines the optimal machine configuration of back-up supplier and the initial level of strategic stock for a given disruption scenario. The list of input parameters and decision variables is presented as follows.

Table 5.3 List of Notations and Decision Variables

Input Parameters	Definition
$T = \{1, 2, ..., T'\}$	Planning horizon consisting of T periods
J	The set of available backup supplier machine configuration
I	The set of plausible capacity loss during recovery time
t	The current time, $t \in T$
j	The machine configuration of backup supplier, $j \in J$
i	Level of capacity loss during recovery time, $i \in I$
C_j	The investment cost of machine configuration j
A	The purchasing/manufacturing cost of initial strategic stock
R	The production cost of main supplier
W	The production cost of back-up supplier
O	The shortage cost
E	The capacity addition cost
H	The inventory holding cost
G_t	1 if a major disruption occurs at time t, else 0
F_t	1 if a minor disruption occurs at time t, else 0
C_d	The maximum capacity level of main supplier
B	The intensity of minor disruption
K_t^i	1 if capacity loss level of i realized during recovery time t
$U_{j,i}$	The fraction of added capacity which is available based on back-up supplier machine configuration of j and realized capacity loss level of i
Decision Variables	
x_t^d	The production quantity of main supplier at time t
x_t^r	The production quantity of back-up supplier at time t

Δ_t^+	The amount of added capacity to back-up supplier at time t
Δ_t^-	The amount of removed capacity from back-up supplier at time t
l_t	The lost demand at time t
ν_t	The level of strategic stock at time t
y_j	1 if backup supplier machine configuration level j is selected, 0 else
ν	The initial level of strategic stock
ω_t^d	The WIP level of main supplier at the end of time t
ρ_t^d	The level of raw material released into main supplier at time t
ω_t^r	The WIP level of back-up supplier at the end of time t
ρ_t^r	The level of raw material released into back-up supplier at time t
ξ_t	The nominal capacity level of back-up supplier at time t
τ_t	The actual capacity of back-up supplier at time t
u_t	The amount of added capacity available during recovery time t

The objective of the model is to minimize the investment cost plus the total operational costs (87). The cost parameters include the investment cost associated with backup supplier machine configuration, the purchasing/manufacturing cost of initial strategic stock, the production cost of main and back-up suppliers, the shortage cost, the capacity addition and strategic stock holding cost.

$$Min \sum_{j\in J} C_j y_j + A\nu + \sum_{t\in T} Rx_t^d + Wx_t^r + E\Delta_t^+ + Ol_t + H\nu_t \quad (87)$$

The flow of material is represented through constraints (88) to (90). The demand at any period is satisfied by production of the main and the back-up supplier and strategic stock. The unsatisfied demand is assumed to be lost (88). The level of strategic stock at any period does not change if there is no disruption (89). The level of strategic stock at the beginning of the planning horizon equals to ν (90). The constraint (91) guarantees that only one configuration would be selected out of possible back-up supplier machine configurations in set J. The impact of disruptions over the main supplier capacity is represented by (92). The level of work in process in the main supplier and the back-up supplier are identified through constraints (93) and (94) respectively. The WIP level at the beginning of each period is the WIP level of the previous period ω_{t-1} plus the amount of raw material released ρ_t. The WIP level at the end of each period ω_t is the difference between the level of WIP at the beginning of that period and the production level x_t.

$$\textit{Subject to}:$$

$$D_t = x_t^d + x_t^r + l_t + (G_t + F_t)(i_t) \quad \forall\, t \in T \tag{88}$$

$$\nu_t = \nu_{t-1} - (G_t + F_t) i_t \quad \forall\, t \in T \tag{89}$$

$$\nu_0 = \nu \tag{90}$$

$$\sum_{j \in J} C_j y_j = 1 \tag{91}$$

$$x_t^d \le (1 - F_t B - G_t) C_d \quad \forall\, t \in T \tag{92}$$

$$\omega_t^d = \omega_{t-1}^d + \rho_t^d - x_t^d \quad \forall\, t \in T \tag{93}$$

$$\omega_t^r = \omega_{t-1}^r + \rho_t^r - x_t^r \quad \forall\, t \in T \tag{94}$$

The capacity balance equations of the back-up supplier are represented in constraints (95) and (96). The constraint (95) indicated that the nominal capacity ξ_t determines the amount of capacity that the system is set to reach. It is equal to the nominal capacity of the previous period plus the amount of added capacity Δ_t^+ or minus the amount of removed capacity Δ_t^-. In the periods where capacity is added, the actual capacity τ_t is the nominal capacity of the previous period plus the amount of added capacity that is available during recovery time u_t. In constraint (96), the actual capacity is equal to the nominal capacity in periods with capacity decrease. The amount of added capacity available during the recovery time is bounded by a fraction $U_{j,i}$ of the added capacity where $U_{j,i}$ depends on the back-up supplier machine configuration y_j and the level of the capacity loss realized during recovery time K_t^i as presented in constraint (97). Note that K_t^i is a binary input parameter and $\sum_{i \in I} K_t^i = 1$. The constraint (98) states that the back-up supplier production at any period is less than the actual capacity. The effect of congestion observed at the back-up supplier during the recovery time is represented with an M/G/1 queueing based nonlinear clearing function which is proposed by (Missbauer, 2002). In order to solve the complexity associated with the nonlinearity of clearing function, we apply an outer approximation. Thus the clearing function is represented through a set of M planes in (99).

$$\xi_t = \xi_{t-1} + \Delta_t^+ - \Delta_t^- \quad \forall\, t \in T \tag{95}$$

$$\tau_t = \xi_{t-1} + u_t - \Delta_t^- \quad \forall\, t \in T \tag{96}$$

$$u_t \le (\sum_{i \in I} K_t^i U_{j,i})\Delta_t^+ + \Psi(1 - y_j) \quad \forall\, t \in T \tag{97}$$

$$x_t^r \le \tau_t \quad \forall\, t \in T \tag{98}$$

$$x_t^r \le A_m(\omega_{t-1}^r + \rho_t^r) + B_m \tau_t - G_m \quad \forall\, t \in T \quad \forall m \in M \tag{99}$$

$$\begin{gathered} y_j \in \{0,1\},\ \nu \ge 0,\ x_t^d \ge 0,\ x_t^r \ge 0,\ l_t \ge 0,\ v_t \ge 0,\ i_t \ge 0,\ \omega_t^d \ge 0,\ \omega_t^r \ge 0 \\ \rho_t^d \ge 0,\ \rho_t^r \ge 0,\ \xi_t \ge 0,\ \tau_t \ge 0,\ \Delta_t^+ \ge 0,\ \Delta_t^- \ge 0,\ u_t \ge 0 \end{gathered} \tag{100}$$

The MIP capacity planning model presented in this section provides the contingency capacity plan of supply network, optimal machine configuration of back-up supplier and initial level of strategic stock for a given disruption scenario. Next section present the last step of the solution methodology presented in Figure 5.1 which is to develop a robust SCND model based on the proposed capacity planning model. This robust SCND model identifies the optimal machine configuration of back-up supplier and initial level of strategic stock by considering all representative disruption scenarios which are selected in step 1.

5.1.3 Robust SCND Model

In this Chapter, we treat each scenario as an objective. Since the scenario probabilities are not available, all objectives can be treated as equally important. Our goal is to achieve solution robustness defined as a solution which remains close to optimal for any occurrence of scenarios. Therefore, we will find a fair solution, in which the relative regrets of all scenarios are as much close to each other as possible. The relative regret of a solution in a given scenario is defined as the difference between the cost of the solution in that scenario and the cost of the optimal solution for that scenario (Snyder and Daskin, 2006).

In multi objective decision making problems where all objectives are equally important to the decision maker, the ordered weighted averaging (OWA) aggregation operator could be applied to achieve a fair solution (Sawik 2014). In such a problem context, the objective is to generate a fair solution, in which all normalized objective function values are as close to each other as possible. The OWA aggregation operator provides the sum of weighted objective function values which have been sorted in the order of the largest value, the two largest values and so on. Liu and Papageorgiou (2013) provide a formulation in order to transform the OWA aggregation operator to the objective function of a linear minimization problem. Later on, Liu et al (2014) prove that assignment of equal weights to the optimization problem in Liu and Papageorgiou (2013) provides a fair as well as a Pareto optimal solution. We apply the formulation proposed in Liu and Papageorgiou (2013) in order to represent the OWA aggregation operator in the objective function. Since all scenarios are equally important, we assign identical weights to all objectives. The formulation of OWA aggregation operator in the robust model according to Liu and Papageorgiou (2013) approach is represented in (101) and (102).

$$Min \sum_{l \text{ in } S} \left(l\lambda_l + \sum_{s \text{ in } S} \delta_{sl} \right) \tag{101}$$

$$Subject\ to:$$

$$\lambda_l + \delta_{sl} \geq f_s \quad \forall l,\ \forall s \in S \tag{102}$$

The above formulation provides minimizing the summation of largest value, the two largest values, the three largest values and so on of outcome values f_s which represents the relative regret of scenario s. The variable λ_l is unrestricted and the non-negative variable δ_{sl} represents the upside deviation of f_s from the value of λ_l. The constraint (103) determines the relative regret of each scenario based on the difference between the cost of the solution in that scenario Z_s(104) and the cost of the optimal solution for that scenario Z_s^*. Note that Z_s^* is an input parameter which is defined as the total cost of MIP contingency capacity planning model for scenario s.

$$f_s = \frac{Z_s - Z_s^*}{Z_s^*} \quad \forall s \in S \tag{103}$$

$$Z_s = \sum_{j \in J} C_j y_j + Av + \sum_{t \in T} Rx_{s,t}^d + Wx_{s,t}^r + E\Delta_{s,t}^+ + Ol_{s,t} + H\nu_{s,t} \quad \forall s \in S \tag{104}$$

$$(88)\ to\ (99) \quad \forall s \in S \tag{105}$$

The proposed MIP robust SCND model applies the ordered weighted averaging (OWA) aggregation operator to achieve solution robustness when probabilities of representative scenarios selected in step 1 are unavailable. In the next section, we present an illustrative case study and experiments to compare the proposed methodology with other approaches from the literature.

5.2 Numerical Experiments

We consider the supply chain associated with a product whose demand is assumed to be deterministic and 6,500 units per period. Three different layout configurations are presented as decision variables: parallel, parallel-serial and serial. The parallel configuration provides higher level of capacity during the recovery time however the better scalability increases the investment cost of parallel configuration (Wang and Koren, 2012) as indicated in Table 5.4.

Table 5.4 List of Cost Parameters

Cost Parameter	Value ($)	Cost Parameter	Value ($)
C_j : The investment cost of machine layout where j is parallel configuration	135,000	H : Holding cost ($/unit)	40
C_j : The investment cost of machine layout where j is parallel-serial configuration	90,000	W : Back-up supplier production ($/unit)	125
C_j : The investment cost of machine layout where j is serial configuration	45,000	R : Main supplier production ($/unit)	25
A : Strategic stock investment ($/unit)	180	O : Shortage ($/unit)	300
		E : Capacity addition ($/unit)	20

The maximum capacity of the main supplier is higher than back-up supplier. These values are set to 10,000 and 7,500 units respectively. The fractions of the added capacity which are available during the recovery time $U_{j,i}$ with respect to the back-up supplier machine configuration and the level of capacity loss are presented in Table 5.5. Furthermore, we assume the intensity of the minor disruptions on the main supplier results in 30% loss in its maximum capacity.

Table 5.5 Fraction of the added available capacity available during recovery time - $U_{j,i}$

Level of capacity loss during recovery	Back-up supplier machine configuration		
	Parallel	Parallel-serial	Serial
High	0.933	0.867	0.8
Normal	0.833	0.667	0.5
Low	0.733	0.467	0.2

Based on the stated assumptions and inputs, the following experiments are conducted. First, the performance of the clustering based MIP scenario reduction model presented in section 5.1.1 is compared against the MIP scenario reduction model proposed by Li and Floudas (2014) both with respect to the quality of selected representative scenarios and the computational time. Next, the solution robustness of back-up supplier machine configuration and initial level of strategic stock determined by robust SCND model proposed in section 5.1.3 is compared against other formulation approaches such as minimizing the average and worst case cost performance, minimizing the average and the standard deviation of the cost and minimizing the worst case cost performance (Diaz et al. 2017).

5.2.1 Investigating the performance of CBSR model

The performance of the proposed clustering based scenario reduction methodology is compared to the MIP formulation proposed by Li et al. (2014) through the following illustrative example. Three cases with different length of product lifecycle are considered ($T_{\text{Small Set}} = 6, T_{\text{Medium Set}} = 8, T_{\text{Large Set}} = 10$ periods). For each case, the original scenario set is developed using the scenario tree approach. The random parameters considered in generating the original scenario set are levels of disruptions and capacity losses during the recovery time. The sizes of the original scenario set associated with small set ($T = 6$), Medium set ($T = 8$) and large set ($T = 10$) are 162, 556 and 1571 scenarios respectively. The objective is to generate different subsets of the original scenario set to test the effectiveness of the proposed scenario reduction algorithm.

The maximum and minimum levels of coverage per scenario represented by UB and LB in the clustering based scenario reduction model are set to 1 and 0.8 accordingly. Therefore, the overall of coverage $\sum_{s' \in S} w_{ss'} Y_{ss'}$ provided to scenario s from cluster center scenarios $s' \in S$ should be at least 0.8 and at most 1. The primary admissible tolerance l_s and the secondary admissible tolerance u_s are calculated based on $f_{\max\ s}$ which is the maximum distance of scenario s from other scenarios. For each scenario, the primary admissible tolerance is set to 25% of $f_{\max\ s}$ and the secondary tolerance is set to 75% of $f_{\max\ s}$. For these settings, the proposed clustering based scenario reduction methodology (CBSR) and the OSCAR scenario reduction algorithm proposed by Li and Floudas (2014) are both applied to perform a comparison. The cumulative probability distributions of the cost of scenarios optimal solution (z_s^*) associated with small, medium and large original scenario sets and the reduced sets generated by CBSR and OSCAR for different levels of reduction are presented in Figure 5.2, 5.3 and 5.4 accordingly.

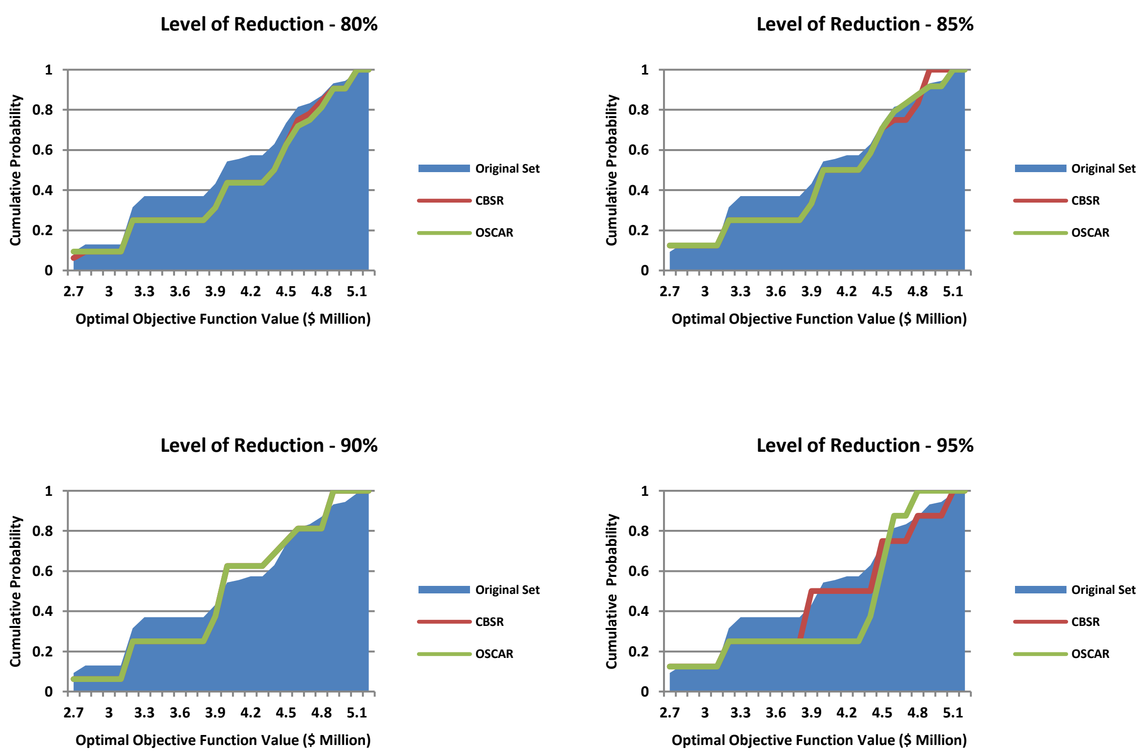


Figure 5.2 Cumulative distribution of cost of optimal solutions, small scenario set (162 scenarios)

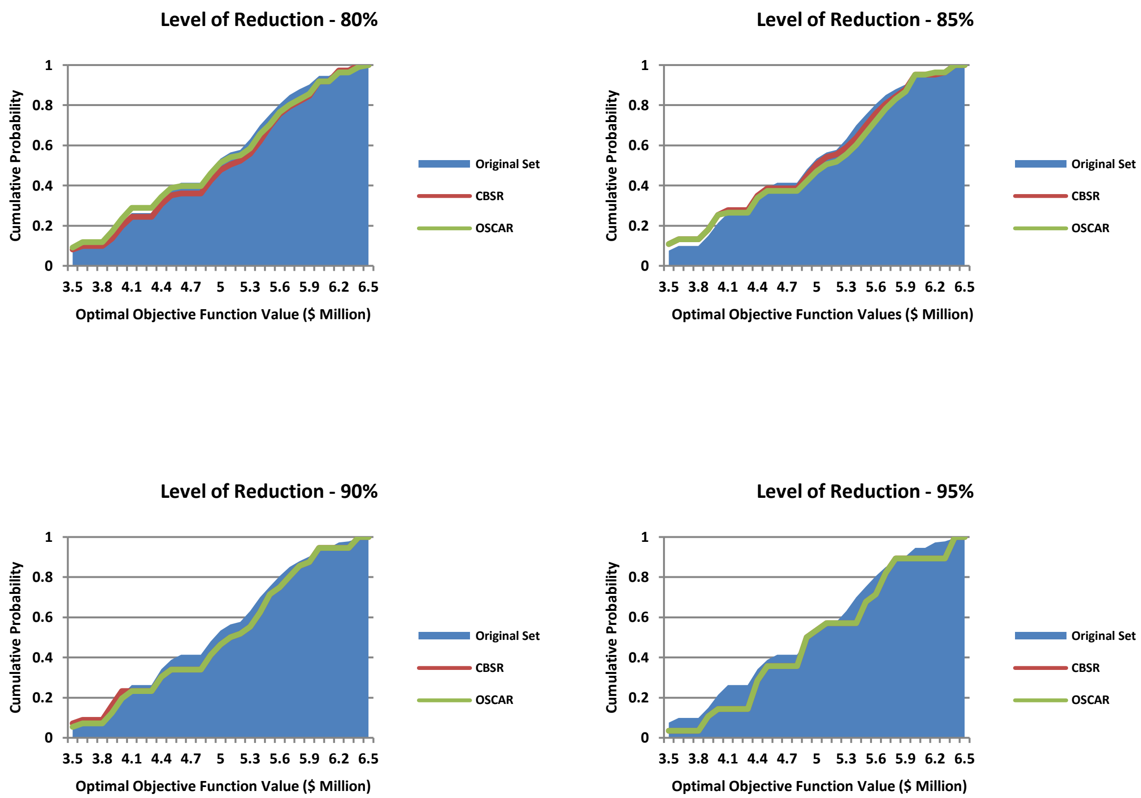
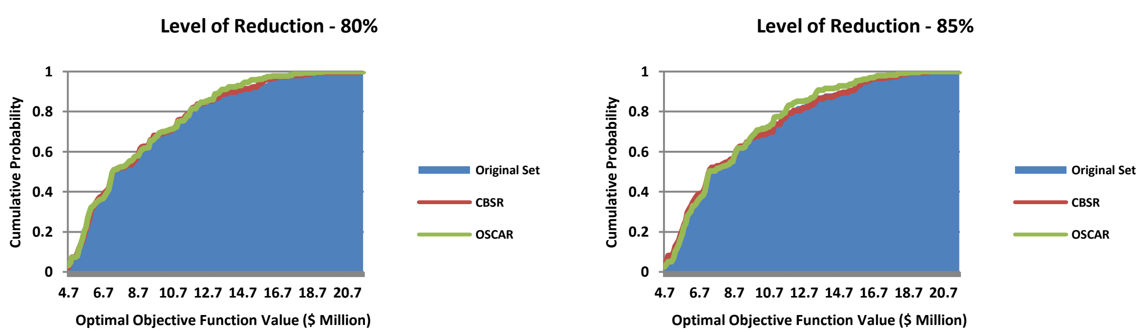


Figure 5.3 Cumulative distribution of cost of optimal solutions, medium scenario set (556 scenarios)



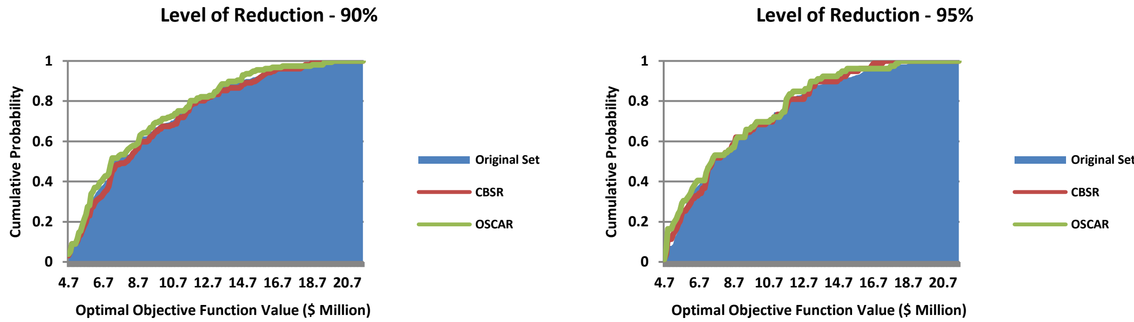


Figure 5.4 Cumulative distribution of cost of optimal solutions, large scenario set (1571 scenarios)

The results in graphs above show that both methods have close performance in identifying reduced scenario sets with cumulative probability distributions of scenarios optimal cost close to original scenario set. In order to further compare the quality of the reduced scenario sets developed by CBSR against OSCAR, statistical measures are computed for the original and reduced distributions. The statistical parameters considered are maximum, minimum, expected value, standard deviation, skewness and kurtosis of the optimal objective function values (z_s^*). The skewness measures the degree of symmetry of a distribution and the kurtosis measures the height and level of sharpness of the central peak of distribution compared to a standard normal distribution. For each parameter, the relative percentage difference between the values obtained from the original scenario set and the reduced scenario sets generated by CBSR and OSCAR are presented in Table 5.6. The lower the relative difference, the closer the reduced scenario set to the original scenario set with respect to the statistical parameter considered. The computation time of each instance used by CPLEX is indicated in the last row.

Table 5.6 Statistics on scenario reduction results

	Level of Reduction	**80%**		**85%**		**90%**		**95%**	
Measure	**Scenario Set \ Method**	**CBSR**	**OSCAR**	**CBSR**	**OSCAR**	**CBSR**	**OSCAR**	**CBSR**	**OSCAR**
# of Scenarios	**Small**	32		24		16		8	
	Medium	111		83		56		28	
	Large	314		236		157		79	
Maximum	**Small**	1.54%	1.54%	5.00%	1.54%	5.21%	5.21%	1.54%	7.57%
	Medium	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%

	Large	0.37%	10.23%	7.74%	10.23%	0.37%	7.74%	20.17%	16.00%
Minimum	**Small**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Medium	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Large	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Expected Value	**Small**	5.08%	5.26%	3.01%	3.01%	1.64%	1.64%	3.43%	5.41%
	Medium	2.11%	2.74%	0.57%	1.41%	2.28%	2.58%	3.34%	3.34%
	Large	1.03%	2.26%	0.74%	1.23%	3.43%	3.86%	0.49%	2.68%
Standard Deviation	**Small**	4.17%	2.14%	3.88%	3.24%	14.53%	14.53%	0.39%	5.07%
	Medium	4.62%	8.07%	8.47%	9.65%	3.44%	0.84%	1.32%	1.32%
	Large	3.54%	8.87%	1.78%	8.36%	5.15%	4.19%	5.79%	5.19%
Skewness	**Small**	246.46%	253.83%	256.54%	225.10%	101.04%	101.04%	209.53%	685.14%
	Medium	107.25%	105.65%	65.47%	124.84%	128.27%	104.42%	70.93%	70.93%
	Large	0.40%	12.83%	7.50%	3.57%	9.95%	16.55%	26.38%	15.15%
Kurtosis	**Small**	38.54%	39.57%	41.70%	45.45%	53.33%	53.33%	61.38%	150.61%
	Medium	0.69%	11.81%	11.10%	12.11%	8.80%	8.94%	22.41%	22.41%
	Large	13.28%	107.59%	138.41%	14.91%	83.66%	94.15%	187.61%	178.72%
Time (s)	**Small**	12	14	15	23	24	28	30	35
	Medium	161	2,000	153	1,822	307	7,440	330	2,040
	Large	6,060	16,380	7,500	22,920	9,300	36,240	10,800	41,640

The results show that the relative difference between the minimum value of original scenario set and the minimum values of reduced set generated by CBSR and OSCAR is zero across different sizes of original scenario set and levels of reduction. Therefore, both scenario reduction methods cover the minimum performance. For the maximum value, the relative difference in both methods increases as the level of reduction increases. This is a result of limiting the number of representative scenarios being used as we increase the level of reduction leading to the maximum value of reduced scenario set to be significantly different than original scenario set. On the other hand, the relative difference of maximum value of reduced scenario set generated by CBSR is lower.

For the expected value, the relative differences of reduced scenario sets developed by CBSR are consistently smaller than their counterparts achieved by OSCAR. Thus, the expected values of reduced scenario sets developed by CBSR are closer to original values especially for larger sizes of original scenario set. However, OSCAR provides the standard deviation of scenario values in the reduced set closer to original set for small size of original set. Both methods have similar performance for medium and large size of original scenario set. The skewness of reduced scenario set distributions is closer to original distribution for large size of scenario set and both methods have similar performance across different levels of reduction. This is because of higher number of scenarios which provide a similar degree of symmetry compared to the original one. For kurtosis, the relative differences of reduced scenario set distributions achieved by CBSR and OSCAR are close across different sizes of original scenario set and levels of reduction.

The computational time which is required to achieve reduced scenario sets increases as the level of reduction and the size of original scenario set increases. The higher level of reduction requires more scenarios to be eliminated which increase the computational efforts. On the other hand, the size of the problem increases with the size of original scenario set which results in long computational time. Furthermore, CBSR provides the reduced scenario set in significantly shorter amount of time compared to OSCAR. This is due to partial and multiple coverage capabilities embedded in CBSR formulation which allows a scenario to be fully or partially covered by multiple representative scenarios. Therefore, this technique requires less computational efforts compared to other approaches where a given scenario is limited to be fully covered by only one representative scenario. According to results discussed earlier, the quality of reduced scenario sets achieved by CBSR is close to OSCAR. Therefore, the short computational time performance of CBSR has least impact on the quality of the reduced scenario set developed.

5.2.2 Robust Formulation Assessment

In this section, the performance of the proposed ordered weighted averaging (OWA) aggregation operator in the objective function of the robust model is compared with respect to other formulation options. The set of alternative robust objective function formulation options considered in this study includes minimizing the average cost and worst case cost performance (AVG-WC), minimizing the average cost and the standard deviation of the cost of the solution across different scenarios (AVG-STD) and minimizing the worst case cost performance of the solution (Diaz et al. 2017). The minimization of the standard deviation aims to reduce the variation in cost, whereas the minimization of the maximum cost tries to minimize the worst case scenario cost derived from design decisions. First we execute the MIP robust SCND model presented in section 5.1.3. Next, the following models are executed for the formulation options elaborated above. This comparison analysis is conducted over large size original scenario set which includes 1571 scenarios.

5.2.2.1 Minimizing the Average and Worst Case Cost (AVG-WC)

The input parameter Ω in this model represents the number of scenarios in the scenario set.

$$Min \ \frac{\sum_{s \ in \ S} Z_s}{\Omega} + W \tag{106}$$

Subject To:

$$W \geq Z_s \quad \forall s \in S \tag{107}$$

$$(88) \ to \ (99), \ (104) \tag{108}$$

5.2.2.2 Minimizing the Average and Standard Deviation of the Cost (AVG-STD)

The decision variable $\Delta_s^+ \geq 0$ represents the deviation of each scenario cost from the average cost.

$$Min \ \frac{\sum_{s \ in \ S} Z_s}{\Omega} + \sum_{s \ in \ S} \Delta_s^+ \tag{109}$$

Subject To:

$$\Delta_s^+ \geq Z_s - \frac{\sum_{s \ in \ S} Z_s}{\Omega} \quad \forall s \in S \tag{110}$$

$$(88) \ to \ (99), \ (104) \tag{111}$$

5.2.2.3 Minimizing the Worst Case Cost (MinMax)

$$Min \ W \tag{112}$$

Subject to:

$$W \geq Z_s \qquad \forall s \in S \tag{113}$$

$$(88)\ \mathit{to}\ (99),\ (104) \tag{114}$$

In order to investigate the performance of the ordered weighted averaging (OWA) aggregation operator based formulation against proposed robust formulation options, we benchmark the solutions returned by each model with respect to three robustness measures called precision to regret, precision to average (Cunha and Covas, 2008) and relative deviation (Bernard 2010). The precision to regret represents the difference between the cost of the solution obtained to the cost of the optimal solution of each scenario and it is given by:

$$\text{Precision to Regret (\%)} = \frac{\sum_{s\ in\ S}\left(1 - \frac{\left|Z_s - Z_s^*\right|}{Z_s^*}\right)}{\Omega} \tag{115}$$

where Z_s represents the cost of scenario s achieved by solving robust model and it is calculated by constraint (104) and Z_s^* is the optimal cost of scenario s calculated by solving the capacity planning model presented in section 4.2 for scenario s. The precision to average represents the deviation of the cost of the solution obtained to the average of cost of optimal solutions across all scenarios and it is computed as follows.

$$\text{Precision to Average (\%)} = \frac{\sum_{s\ in\ S}\left(1 - \frac{\left|Z_s - Z_{\text{average}}\right|}{Z_{\text{average}}}\right)}{\Omega} \tag{116}$$

$$Z_{\text{average}} = \frac{\sum_{s\ in\ S} Z_s^*}{\Omega} \tag{117}$$

The relative deviation is the value of maximum relative regret across the scenarios where the relative regret is the difference between the costs of the solution obtained to the cost of the optimal solution of each scenario. The relative deviation is calculated by:

$$\text{Relative Deviation (\%)} = Max_{s \ in \ S} \frac{\left| Z_s - Z_s^* \right|}{Z_s^*} \qquad (118)$$

Next, the three robust models elaborated above and the robust formulation with ordered weighted averaging (OWA) aggregation operator in the objective function of the model are executed in order to determine the optimal machine configuration of back-up supplier and the initial level of safety stock. The set of input parameters considered are operational and investment costs presented in Table 5.4 and the fraction of the added capacity which is available during recovery time based on the machine configuration of back-up supplier (Table 5.5). The results show that the optimal back-up supplier machine configuration across all robust formulations considered is the parallel configuration. The selection of the initial strategic stock levels with respect to different levels of scenario reduction is presented in Figure 5.5.

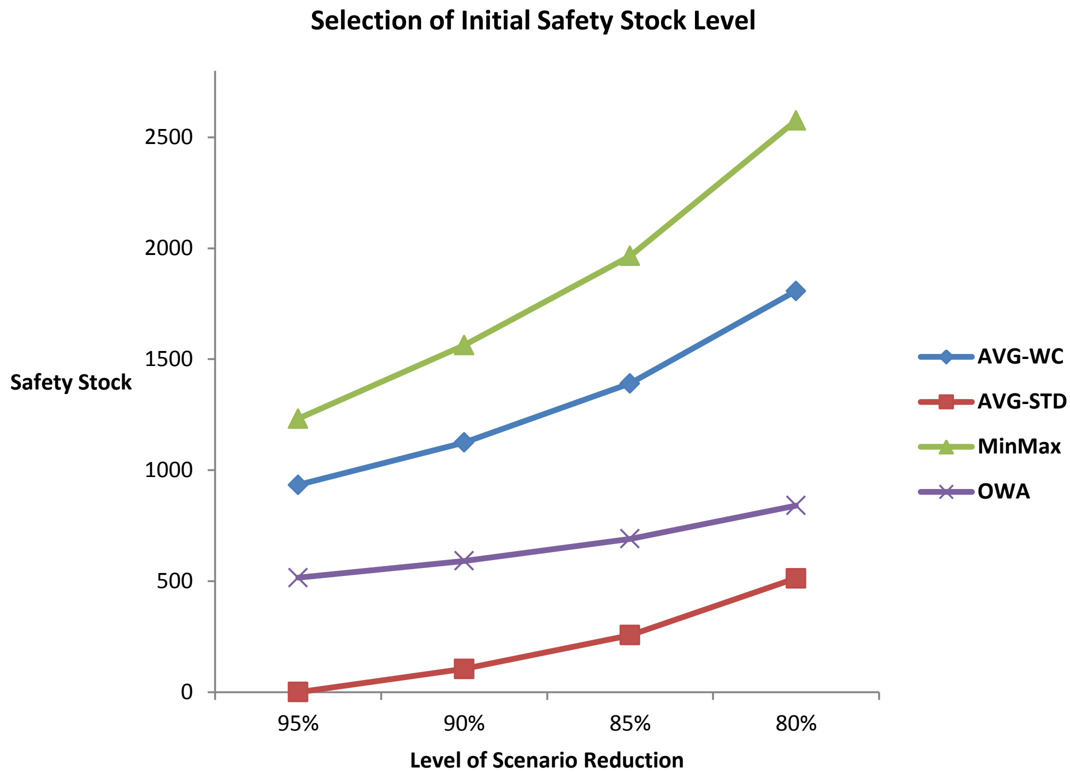


Figure 5.5 Safety Stock Selections of Robust Formulations

The safety stock level increases in all four robust formulations as the level of scenario reduction decreases. This happens since the number of scenarios with major disruption impacts increases in the reduced scenario set as the size of reduced set increases. Therefore, higher level of safety stock is required to cover these high impact disruption scenarios. The MinMax robust formulation represents the highest level of safety stock selection across different levels of scenario reduction. The reason is in MinMax formulation's inherent characteristic which focuses only on worst case scenarios. Therefore, it selects the highest level of safety stock to minimize the impact of extremely disruptive scenarios. The AVG-WC robust model represents the second highest levels of safety stock selection across different levels of scenario reduction. The selected safety stock levels are lower than MinMax since AVG-WC model is formulated to minimize average cost in addition to worst case cost. However, the level of safety stock selected by AVG-WC formulation is higher compared to OWA based robust formulation. Furthermore, the difference in selected level of safety stock between these two approaches increases as the level of scenario reduction decreases. The OWA operator based robust formulation minimizes scenarios costs in the decreasing order of scenario impact (e.g. first the worst case cost, second the summation of two worst case costs and so on). Therefore, as the number of scenarios increases due to decrease in the level of scenario reduction, the performance of OWA based formulation becomes closer to minimization of average cost. This behavior results in safety stock levels selected by OWA based model to become closer to AVG-STD formulation safety stock selections and further away from AVG-WC and MinMax selections as the level of scenario reduction decreases. The lowest level of safety stock is selected by AVG-STD robust model. This is due to AVG-STD objective function formulation which aims to minimize the average cost and the positive deviation of each scenario cost from average. The minimization of this formulation requires less safety stock since the objective is to minimize the deviation of each scenario cost from average rather than worst case scenarios' costs.

The quality of solutions provided by each formulation is measured with respect to three robustness benchmarks; precision to regret, precision to average and relative deviation. The results are represented in Table 5.7.

Table 5.7 Performance of Robust Formulations with Respect to Robustness Measures

Robust Formulation	Level of Scenario Reduction	Precision to Regret	Precision to AVG	Relative Deviation
AVG-WC	95%	76.15%	66.94%	18.24%
	90%	75.44%	65.88%	17.69%
	85%	73.60%	65.01%	15.66%
	80%	71.10%	63.13%	12.07%
AVG-STD	95%	75.83%	70.13%	28.71%
	90%	77.02%	72.56%	24.79%
	85%	79.01%	75.81%	21.49%
	80%	82.13%	80.20%	18.98%
MinMax	95%	72.13%	65.12%	16.05%
	90%	70.03%	63.17%	14.41%
	85%	68.29%	60.57%	11.09%
	80%	63.49%	57.38%	8.68%

OWA	95%	79.74%	68.43%	20.11%
	90%	81.33%	70.88%	19.66%
	85%	82.51%	74.10%	18.84%
	80%	84.90%	79.16%	17.50%

As the level of reduction decreases, the results show that the precision to regret and precision to average performance of MinMax and AVG-WC formulations degrades however the relative deviation improves. The reason is in the objective of these two approaches which aims to minimize the worst case scenario cost. As the size of reduced scenario set increases, the number of scenarios with major disruption impact increases. Therefore, the solution achieved by MinMax and AVG-WC would alleviate worst case scenario impact and has a low performance with respect to average performance (precision to average) and optimal performance of each scenario (precision to regret). Furthermore, the magnitude of improvement in relative deviation and decline in precision to regret and average of MinMax is higher than AVG-WC. The reason lies in MinMax formulation which focuses only on worst case scenarios.

The precision to regret and the precision to average of both OWA based formulation and AVG-STD model increases as the level of reduction decreases. The pace of improvement in precision to regret is higher in OWA based formulation compared to AVG-STD model. This behavior is observed since OWA based formulation identifies solutions such that the distance to optimal solution of each scenario is targeted to be minimized. Therefore, it has the highest precision to regret at different levels of scenario reduction and among all robust formulations considered. On the other hand, AVG-STD model represents the highest precision to average across all levels of scenario reduction. This is due to the fact that in this approach the objective aims to minimize the positive deviation of each scenario cost from average cost. Furthermore, OWA based formulation has the second highest precision to average at different levels of reduction. The relative deviation of OWA and AVG-STD model improves as the size of reduced scenario set increases. However, the magnitude of improvement in both approaches is lower compared to MinMax and AVG-WC. Furthermore, the OWA based formulation levels of improvement are higher compared to AVG-STD model since this approach minimize the impact of worst case scenarios in the decreasing order of impact.

5.2.3 Discussion of Results

The first section of the numerical experiments investigate the performance of the proposed clustering based scenario reduction (CBSR) methodology against the MIP based approach presented by Li et al. (2014) called OSCAR. The results show that CBSR provides the reduced scenario set in significantly shorter amount of time compared to OSCAR. This is due to partial and multiple coverage capabilities embedded in CBSR formulation. Furthermore, the quality of reduced scenario sets achieved by CBSR is close to OSCAR. For the minimum value, both scenario reduction methods cover the minimum value of the original set. However, the maximum and expected values of reduced scenario sets developed by CBSR are closer to original values especially for larger sizes of original scenario set. On the other hand, OSCAR provides the standard deviation of scenario values in the reduced set closer to original set for small size of original set. Furthermore, both methods have similar performance across different levels of reduction for skewness and kurtosis. Therefore, the short computational time performance of CBSR has least impact on the quality of the reduced scenario set developed.

The second part of the numerical experiments compares the performance of the proposed ordered weighted averaging (OWA) aggregation operator in the objective function of the robust model against other robust formulations such as minimization of the average cost and worst case cost performance (AVG-WC), minimization of the average cost and the standard deviation of the cost of the solution across different scenarios (AVG-STD) and minimization of the worst case cost performance of the solution (MinMax).

The results show that the OWA based formulation could provide solutions with a fair level of solution robustness compared to other formulations considered in this Chapter. This is based on highest level of precision to regret which is presented by OWA based formulation across different sizes of reduced scenario set in addition to second highest level of precision to average provided by OWA based model. Finally, the OWA based formulation has a fair performance with respect to relative deviation.

Chapter 6

Conclusion

In an era of globalization where supply chains are dispersed across the world in order to benefit from lower manufacturing and supply costs as well as better access to the global markets, it is important to consider the risk of disruptions in the design of supply chain. The focus of this thesis is on disruptions which impact the supply network. We incorporate two well-known risk mitigation strategies, the strategic safety stock and contingent sourcing in the supply network in order to mitigate the impact of operational risks and disruptions. Our design problem includes the challenge to identify the optimal levels of strategic safety stock and the response speed of the back-up supplier in order to create a robust supply flow. Furthermore, we assume that the back-up supplier could invest in its layout configuration in order to improve the response speed. Our first contribution is in Chapter 3 where we consider the impact of congestion over production capacity of back-up supplier during the response time. To this end, the clearing functions are incorporated into the supply chain contingency capacity planning model. The results in Chapter 3 show that considering congestion is especially critical for risk-neutral decision makers in mitigating against disruptions. Furthermore, there is significant improvement in supply chain service level as the congestion effects are taken into account

We solve our design problem under two plausible scenarios which represent the quality of information that might be available to estimate the probabilities of disruptions. The first scenario considers the situation where there is enough information available to estimate the probability of disruptions. This situation could for instance represent a supply network which source from suppliers located in Far East. For this case, we propose a two-stage robust optimization model in Chapter 4 where the first stage decision variables are levels of back-up supplier's response speed and strategic safety stock. Our contribution is to represent the randomness associated with disruptions and available capacity during the response time into the robust optimization model. This model minimizes the summation of expected cost and the variation from expected cost in order to provide solution robustness. The results in Chapter 4 show the optimality of the faster response speed as the failure probability increases or recovery probability decreases. Furthermore, higher level of strategic stock and faster response speed level are required as the probability of the lower level of capacity availability during the response time increases. Finally, we demonstrate that it is worthwhile to consider the complexity associated with modelling of the randomness associated with the response time characteristics due to the improvements in the supply chain performance.

The second scenario represents the situation where there is no level of information available to estimate the probability of disruptions. This situation could represent a newly established supply network. In this situation, we achieve design decisions with solution robustness such that the optimal solution would be close to the optimal solution of any scenario as much as possible. To this end, we employ the Ordered Weighted Averaging (OWA) aggregation operator in the objective function of our two-stage robust optimization model in Chapter 5. The implementation of the robust optimization models with lexicographic based formulation is known to be difficult for large problems. Since we consider a strategic supply chain network design problem which includes a large scenario set, there is computational complexity in solving the two-stage robust optimization model. To overcome this challenge, we propose a novel clustering based MIP scenario reduction model called CBSR in Chapter 5. This proposed scenario reduction model groups scenarios into clusters based on a probabilistic distance metric and it identifies the cluster center associated with each cluster. Finally, the cluster centers will represent the reduced scenario set and the remaining scenarios will be removed. Our contribution is in incorporating the gradual coverage function which is used frequently in location covering models into the scenario reduction formulation. In this context, the gradual coverage function determines the degree of coverage provided by a cluster center to a given scenario. This functionality provides partial coverage and/or coverage with multiple cluster centers capabilities. The performance of the proposed scenario reduction methodology is compared to other approaches available in literature with respect to the quality of solutions and computational time. First, the clustering based MIP scenario reduction model (CBSR) is benchmarked against a MIP based scenario reduction formulation (OSCAR) developed by Li and Floudas (2014). The results indicate that CBSR model has significantly shorter computational time to generate the reduced scenario set compared to OSCAR. Furthermore, the quality of reduced sets achieved by CBSR is at least as good as the reduced sets developed by OSCAR. Next, the performance of the proposed OWA based robust formulation is compared against other formulations with objective of minimizing the average and worst case performance (AVG-WC), minimizing average performance and standard deviation (AVG-STD) and minimizing the worst case performance (MinMax). Three benchmarks are considered to conduct the comparison analysis including: precision to regret, precision to average and relative deviation. The results show that the OWA based robust formulation provides solution with highest precision to regret (solution robustness) and it has the second best performance with respect to precision to average and a fair performance with respect to relative deviation.

The tools developed in this thesis could support the supply chain practitioners in order to design robust supply networks under all plausible future scenarios and with respect to the different levels of data availability.

Future Research Directions

The future research could be conducted in the following directions.

- The proposed supply chain design decisions could be determined for the case when there is information available about the possibilities of operational risks occurrences such as machine breakdown however the probabilities of major disruption occurrences such as earthquake are unavailable. This situation is called partial data availability.

- A scenario reduction approach can be developed for the case where there is correlation between scenarios within the original scenario set. This could represent a situation where there is positive correlation between the higher frequency of machine breakdown and getting closer to the end of product life cycle.
- In this thesis, we assume a single product supply chain to assess volume flexibility. Within the same disruption management strategies, considering the interplay between the demands of multiple products to assess process flexibility could be a future research direction.

References

Baghalian, A., Rezapour, S., and Zanjirani Farahani, R. (2013), "Robust supply chain network design with service level against disruptions and demand uncertainties: A real-life case", *European Journal of Operational Research,* Vol. 227 No.1, pp. 199-215.

Berger PD, Gerstenfeld A, Zeng AZ (2004), "How many suppliers are best? A decision-analysis approach". *Omega*, Vol. 32, pp. 9-15.

Bernard, R. (2010), "Robustness in operational research and decision aiding: A multi-faceted issue." *European Journal of Operational Research*, Vol. 200 No. 3, pp. 629-638.

Bilsel, R.U., Ravindran, A. (2011), "A multiobjective chance constrained programming model for supplier selection under uncertainty", *Transportation Research Part B: Methodological*, Vol. 45, pp. 1284-300.

Bundschuh, M., Klabjan, D., Thurston, D.L. (2003), "Modeling robust and reliable supply chains". Working paper, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Chiu, SL. (1994), "Fuzzy model identification based on cluster estimation". *Journal of Intelligent and Fuzzy System*, Vol. 2, pp. 267-278.

Chopra, S., Sodhi, M.S., (2004), "Managing risk to avoid supply-chain breakdown". *MIT Sloan Management Review*, Vol. 46, pp. 52–61.

Crainic, T.G., Hewitt, M., and Rei, W. (2014), "Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design." *Computers & Operations Research*, Vol. 43: pp. 90-99.

Deif, A.M., ElMaraghy, W.H. (2006), "A Systematic Design Approach for Reconfigurable Manufacturing Systems". *In: Anonymous Advances in Design : Springer*, p. 219-228.

Diaz, J.E., Handl, J., and Xu, D.L. (2017), "Evolutionary robust optimization in production planning–interactions between number of objectives, sample size and choice of robustness measure." *Computers & Operations Research,* Vol. 79: pp. 266-278.

Dupačová, J., Gröwe-Kuska, N., and Römisch, W. (2003), "Scenario reduction in stochastic programming", *Mathematical programming,* Vol. 95 No.3, pp. 493-511.

Ebrahim Nejad, A., Niroomand, I. and Kuzgunkaya, O. (2014), "Responsive contingency planning in supply risk management by considering congestion effects", Omega, Vol.48, pp. 19-35.

Ebrahim Nejad, A. and Kuzgunkaya, O. (2015), "On the value of response time characteristics in robust design of supply flow", *Journal of Manufacturing Technology Management*, vol.26 No.2, pp. 213-230.

Fang, J., Zhao, L., Fransoo, J.C., Van Woensel, T. (2012), "Sourcing strategies in supply risk management: An approximate dynamic programming approach", *Computers & Operations Research*, Vol. 40, No. 5, pp. 1371-1382.

Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., and Goh, M. (2012), "Covering problems in facility location: A review", *Computers & Industrial Engineering*, Vol. 62 No. 1, pp. 368-407.

Farahani, R.Z., Rezapour, S., Drezner, T., Fallah, S. (2013), "Competitive supply chain network design: An overview of classifications, models, solution techniques and applications". *Omega*, Vol. 45, pp.92-118.

Gaspar-Cunha, A., and Covas, J.A. (2008), "Robustness in multi-objective optimization using evolutionary algorithms", *Computational Optimization and Applications*, Vol. 39 No. 1, pp. 75-96.

Ghadge, A., Dani, S., and Kalawsky, R. (2011), "Systems Thinking for Modeling Risk Propagation in Supply Networks", *5th International Conference on Industrial Engineering and Engineering Management, IEEE, Singapore*, pp. 1685-1689.

Gong, J., Mitchell, J.E., Krishnamurthy, A., Wallace, W.A. (2014), "An interdependent layered network model for a resilient supply chain", *Omega*, Vol.46, pp. 104-116.

Govindan, K., and Fattahi, M. (2017), "Investigating risk and robustness measures for supply chain network design under demand uncertainty: A case study of glass supply chain", *International Journal of Production Economics*, Vol. 183, pp. 680-699.

Govindan, K., Fattahi, M., and Keyvanshokooh, E. (2017), "Supply chain network design under uncertainty: A comprehensive review and future research directions." *European Journal of Operational Research*, Vol. 263, No. 1, pp. 108-141.

Hale, T., Moberg, C.R. (2005), "Improving supply chain disaster preparedness: A decision process for secure site location". *International Journal of Physical Distribution & Logistics Management*, Vol. 35, pp. 195-207.

Hallgren, M., and Olhager, J. (2009), "Flexibility configurations: Empirical analysis of volume and product mix flexibility." *Omega*, Vol. 37, No.4, pp. 746-756.

Hamta, N., Akbarpour Shirazi, M., Behdad, S. and Ehsanifar, M. (2017), "A novel bi-level stochastic programming model for supply chain network design with assembly line balancing under demand uncertainty", *Journal of Industrial and Systems Engineering*, Vol. 10 No. 2, pp. 87-112.

Heckmann, I., Comes, T., and Nickel, S. (2015), "A critical review on supply chain risk–Definition, measure and modeling", *Omega*, Vol. 52, pp. 119-132.

Heitsch, H., and Werner, R. (2003), "Scenario reduction algorithms in stochastic programming." *Computational optimization and applications* Vol. 24, No. 2, pp. 187-206.

Hendricks, K.B., Singhal, V.R. (2005), "An Empirical Analysis of the Effect of Supply Chain Disruptions on Long-Run Stock Price Performance and Equity Risk of the Firm". *Production and Operations Management*, Vol.14, pp. 35-52.

Hohenstein, N.O., Feisel, E., Hartmann, E., and Giunipero, L. (2015), "Research on the phenomenon of supply chain resilience: a systematic review and paths for further investigation", *International Journal of Physical Distribution & Logistics Management,* Vol. 45, pp. 90-117.

Hopp, W.J. and Yin, Z. (2006), "Protecting supply chain networks against catastrophic failures", Working Paper, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, IL.

Hopp, W.J., Iravani, S.M. and Liu, Z. (2012), "Mitigating the Impact of Disruptions in Supply Chains" in Gurnani, H., Mehrotra, A. and Ray, S. (Eds.), *Supply Chain Disruptions; Theory and Practice of Managing Risk,* Springer, New York, pp. 21-49.

Hosseini, S., Morshedlou, N., Ivanov, D., Sarder, M. D., Barker, Kash., Abdullah Al Khaled, A. (2019), "Resilient Supplier Selection and Optimal Order Allocation Under Disruption Risks", *International Journal of Production Economics,* In Press.

Iakovou., E., Dimitrios., V., and Anastasios., X. (2010), "A stochastic inventory management model for a dual sourcing supply chain with disruptions." *International Journal of Systems Science,* Vol. 41, No.3, pp. 315-324.

Iris, H., Comes, T., and Nickel, S. (2015), "A critical review on supply chain risk–Definition, measure and modeling", *Omega*, Vol. 52, pp. 119-132.

Jeihoonian, M., Kazemi Zanjani, M., and Gendreau, M. (2017), "Closed-loop supply chain network design under uncertain quality status: Case of durable products", *International Journal of Production Economics*, Vol. 183, pp. 470-486.

Jordan, W.C., Graves, S.C. (1995), "Principles on the benefits of manufacturing process flexibility". *Management Science*, Vol.41, pp. 577-94.

Kalai, R., Lamboray, C., and Vanderpooten, D. (2012), "Lexicographic α-robustness: An alternative to min–max criteria", *European Journal of Operational Research*, Vol. 220 No. 3, pp. 722-728.

Kamalahmadi, M., Parast, M.M. (2017), "An assessment of supply chain disruption mitigation strategies", *International Journal of Production Economics*, vol. 184, pp. 210-230.

Karmarkar, U.S. (1989), "Capacity loading and release planning with work-in-progress (WIP) and leadtimes". *Journal of Manufacturing and Operations Management*, Vol.2, pp. 105-123.

Kazemi Zanjani, M., Ait-Kadi, D. and Nourelfath, M. (2010), "Robust production planning in a manufacturing environment with random yield: A case in sawmill production planning", *European Journal of Operational Research*, Vol. 201 No. 3, pp. 882-891.

Keyvanshokooh, E., Ryan, S.M., and Kabir, E. (2016), "Hybrid robust and stochastic optimization for closed-loop supply chain network design using accelerated Benders decomposition." *European Journal of Operational Research*, Vol. 249 No.1, pp. 76-92.

Khatami, M., Mahootchi, M., and Zanjirani Farahani, R. (2015), "Benders’ decomposition for concurrent redesign of forward and closed-loop supply chain network with demand and return uncertainties", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 79, pp. 1-21.

Kim, J., Byung Do, C., Yuncheol, K., Bongju, J. "Robust optimization model for closed-loop supply chain planning under reverse logistics flow and demand uncertainty." *Journal of cleaner production*, Vol.196, pp. 1314-1328.

Kim, S., Uzsoy, R., (2008), "Exact and heuristic procedures for capacity expansion problems with congestion". *IIE Transactions*, Vol. 40, pp. 1185-1197.

Klibi, W., and Martel, A. (2012), "Modeling approaches for the design of resilient supply networks under disruptions", *International Journal of Production Economics*, Vol. 135, pp. 882-98.

Klibi, W., Martel, A. and Guitouni, A. (2010), "The design of robust value-creating supply chain networks: a critical review", *European Journal of Operational Research,* Vol. 203 No. 2, pp. 283-293.

Koren, Y., Heisel, U., Jovane, F., Moriwaki, T., Pritschow, G., Ulsoy, G. and Van Brussel, H. (1999), "Reconfigurable manufacturing systems", *CIRP Annals-Manufacturing Technology,* Vol. 48 No. 2, pp. 527-540.

Kouvelis, P. and Li, J. (2012), "Contingency Strategies in Managing Supply Systems with Uncertain Lead-Times", *Production and Operations Management,* Vol. 21 No. 1, pp. 161-176.

Kouvelis, P., and Yu, G. (2013), "Robust Discrete Optimization and its Applications", *Springer Science and Business Media,* Vol. 14.

Li, S., He, Y., Chen, L. (2017), "Dynamic strategies for supply disruptions in production-inventory systems", *International Journal of Production Economics,* Vol. 194, pp. 88-101.

Li, Z., Floudas, C.A. (2016), "Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: II. Sequential reduction", *Computers & Chemical Engineering,* Vol. 84, pp. 599-610.

Li, Z., Floudas, C.A. (2014), "Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization", *Computers & Chemical Engineering,* Vol. 70, pp. 50-66.

List, G.F., Wood, B., Nozick, L.K., Turnquist, M.A., Jones, D.A., Kjeldgaard, E.A. and Lawton, C.R. (2003), "Robust optimization for fleet planning under uncertainty", *Transportation Research Part E: Logistics and Transportation Review,* Vol. 39 No. 3, pp. 209-227.

Liu, S., and Papageorgiou, L.G. (2013), "Multiobjective optimisation of production, Distribution and Capacity Planning of Global Supply Chains in the Process Industry", *Omega: The International Journal of Management Science*, Vol. 41, pp. 369–382.

Liu, S., Sawik, T., and Papageorgiou, L.G. (2014), "Corrigendum to Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry", *Omega: The International Journal of Management Science*, Vol. 44 No. 1, pp. 149-153.

Lodree, J.r., Emmett, J., Selda, T. (2009), "Supply chain planning for hurricane response with wind speed information updates." *Computers & Operations Research*, Vol. 36, No.1, pp. 2-15.

Matta, A., Tomasella, M. and Valente, A. (2007), "Impact of ramp-up on the optimal capacity-related reconfiguration policy", *International Journal of Flexible Manufacturing Systems*, Vol. 19 No. 3, pp. 173-194.

Matta, A., Tomasella, M. and Valente, A. (2007), "Impact of ramp-up on the optimal capacity-related reconfiguration policy", *International Journal of Flexible Manufacturing Systems*, Vol. 19 No.3, pp. 173-194.

Mehrabi, M.G, Ulsoy, A.G., Koren,Y. (2000b), "Reconfigurable manufacturing systems and their enabling technologies", *International Journal of Manufacturing Technology and Management*, Vol. 1, pp. 114–131.

Missbauer, H. (2002), "Aggregate order release planning for time-varying demand", *International Journal of Production Research,* Vol. 40 No. 3, pp. 699-718.

Mulvey, John M., Robert J. Vanderbei, and Stavros A. Zenios. (1995), "Robust optimization of large-scale systems", *Operations research*, Vol. 43 No.2, pp. 264-281.

Niroomand I, Kuzgunkaya O, Bulgak AA. (2012), "Effect of System Configuration and Ramp up Time on Manufacturing System Acquisition under Uncertain Demand", *In: Proceedings-*

IOEM: International Conference on Industrial Engineering and Operation Management, pp. 1578-87.

Pahl, J., Voß, S., Woodruff, D.L. (2007), "Production planning with load dependent lead times: an update of research". *Annals of Operations Research*, Vol. 153, pp. 297-345.

Qi, L. (2013), "A continuous-review inventory model with random disruptions at the primary supplier", *European Journal of Operational Research,* Vol. 225 No. 1, pp. 59-74.

Qiu, R., and Yizhi, W. (2016), "Supply chain network design under demand uncertainty and supply disruptions." *Scientific Programming*, Vol. 5.

Revilla, E., and Maria Jesus, S. "The impact of risk management on the frequency of supply chain disruptions: A configurational approach." *International Journal of Operations & Production Management,* Vol. 37, No. 5, pp. 557-576.

Rezapour, S., Srinivasan, R., Tew, J., Allen, J. K., Mistree, F. (2018). "Correlation between strategic and operational risk mitigation strategies in supply networks", *International Journal of Production Economics,* Vol. 201, pp. 225-248.

Rink, D.R., Swan, J.E. (1979) "Product life cycle research: A literature review". *Journal of business Research*, Vol. 7, pp. 219-242.

Rong, Y., L. V. Snyder, Z.-J. M. Shen. (2009c). "Bullwhip and reverse bullwhip e®ects under rationing game". Working paper, Lehigh University, Bethlehem, PA.

Bernard, R. (2010), "Robustness in operational research and decision aiding: A multi-faceted issue." *European Journal of Operational Research*, Vol. 200, No. 3, pp. 629-638.

Sadghiani, N., Torabi, S.A, Sahebjamnia, S. "Retail supply chain network design under operational and disruption risks." *Transportation Research Part E: Logistics and Transportation Review*, Vol. 75, pp. 95-114.

Sahebjamnia, N., Torabi, S.A., Mansouri, S.A. (2018), "Building organizational resilience in the face of multiple disruptions", *International Journal of Production Economics*, Vol. 197, pp. 63-83.

Sawik, T. (2012), "Selection of resilient supply portfolio under disruption risks". , Vol. 41, pp. 259-269.

Sawik, T. (2014), "On the robust decision-making in a supply chain under disruption risks", *International Journal of Production Research,* Vol. 52 No.22, pp. 6760-6781.

Schmitt, A.J. (2011), "Strategies for customer service level protection under multi-echelon supply chain disruption risk", *Transportation Research Part B: Methodological*, Vol. 45 No. 8, pp. 1266-1283.

Schmitt, A.J., Singh, M. (2012), "A quantitative analysis of disruption risk in a multi-echelon supply chain", *International Journal of Production Economics,* Vol. 139 No. 1, pp. 22-32.

Schmitt, A.J., Tomlin, B. (2012), "Sourcing strategies to manage supply disruptions" in Gurnani, H., Mehrotra, A. and Ray, S. (Eds.), *Supply Chain Disruptions; Theory and Practice of Managing Risk,* Springer, New York, pp. 51-72.

Schmitt, A.J., Lawrence, V.S. (2012), "Infinite-horizon models for inventory control under yield uncertainty and disruptions." Computers & Operations Research Vol.39, No.4, pp. 850-862.

Sethi,.A.K, Sethi, S.P. (1990), “Flexibility in Manufacturing, A Survey”, International Journal of Flexible Manufacturing System, Vol. 2, pp. 289-328.

Sheffi, Y., (2005) "The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage". MIT Press, Cambridge, MA, first edition.

Snyder, L.V., Daskin, M.S. (2006), "Stochastic p-robust location problems", *IIE Transactions*, Vol. 38 No. 11, pp. 971-985.

Snyder, L.V., Zümbül, A., Peng, P., Ying, R., Schmitt, A.J., Sinsoysal, B. "OR/MS models for supply chain disruptions: A review." IIE Transactions, Vol. 48, No. 2, pp. 89-109.

Šutienė, K., Makackas, D., Pranevičius, H. (2010), "Multistage K-means clustering for scenario tree construction." *Informatica*, Vol. 21 No. 1, pp. 123-138.

Tang, C.S. (2007), "Robust strategies for mitigating supply chain disruptions", *International Journal of Logistics: Research and Applications,* Vol. 9 No. 1, pp. 33-45.

Terwiesch, C., E Bohn, R. (2001), "Learning and process improvement during production ramp-up", *International Journal of Production Economics,* Vol. 70 No. 1, pp. 1-19.

Tomlin, B. (2006), "On the value of mitigation and contingency strategies for managing supply chain disruption risks", *Management Science,* Vol. 52 No. 5, pp. 639-657.

Tomlin, B., Wang, Y. (2010), "Operational strategies for managing supply chain disruption risk" in Kouvelis, P., Boyabatli, O., Lingxiu, D. and Li, R. (Eds.), *Handbook of Integrated Risk Management in Global Supply Chains,* John Wiley and Sons, New York, pp. 79-101.

Torabi, S.A., Baghersad, M., and Mansouri, S.A. (2015), "Resilient supplier selection and order allocation under operational and disruption risks", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 79, pp. 22-48.

Vidyarthi, N., Elhedhli, S., Jewkes, E. (2009), "Response time reduction in make-to-order and assemble-to-order supply chain design". *IIE Transactions*, Vol. 41, pp. 448-466.

Wagner, S.M., Christoph, B. "An empirical examination of supply chain performance along several dimensions of risk." Journal of business logistics 29.1 (2008): 307-325.

Wai, L.C. and Wongsurawat, W. (2012), "Crisis management: Western Digital's 46-day recovery from the 2011 flood disaster in Thailand", *Strategy & Leadership,* Vol. 41 No. 1, pp. 34-38.

Wang, W. and Koren, Y. (2012), "Scalability planning for reconfigurable manufacturing systems", *Journal of Manufacturing Systems,* Vol. 31 No. 2, pp. 83-91.

Wieland, A. (2013), "Selecting the right supply chain based on risks", *Journal of Manufacturing Technology Management,* Vol. 24 No. 5, pp. 652-668.

Wong, H-Y. and Rosenhead J., (2000). "A rigorous definition of robustness analysis." *Journal of the operational research society*, Vol.51, pp. 176-182.

Zeballos, L.J., Carlos A.M., Barbosa-Povoa, A.P., Augusto, Q. (2014) "Multi-period design and planning of closed-loop supply chains with uncertain supply and demand." *Computers & Chemical Engineering* Vol. 66, pp. 151-164.

Appendix A

This section describes the outer approximation approach which has been used to replace the clearing function in the MIP model by a set of lines. As shown in Figure A.1, the approximated throughput for a given work load $(\omega_{t-1} + \rho_t)$ is

$$f(\omega_{t-1} + \rho_t) = \min\left[A_\eta(\omega_{t-1} + \rho_t) + B_\eta \right] \qquad \eta \in \{1, 2, ..., \mathrm{N}\} \qquad \text{(A.1)}$$

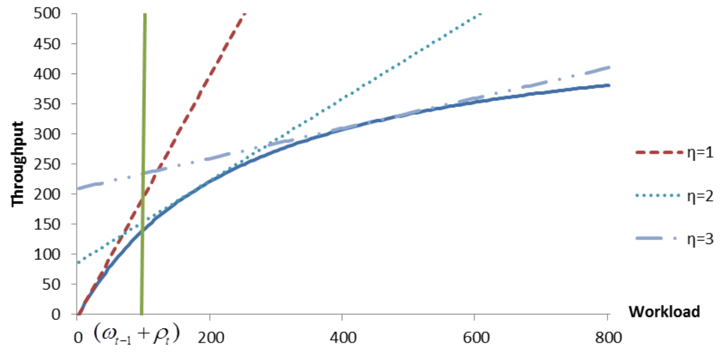


Figure A.1 The clearing function approximated by lines, N=3.

Each line is tangent to the curve at a certain point. Since this is an outer approximation, the tangent points should be selected to minimize the estimation error. For that purpose, the subtractive clustering introduced by Chiu (1994) is employed.

For each capacity scenario, the clearing function (24) is represented as a set of points. The subtractive clustering separates these points into clusters based on a predetermined radius. Each cluster has a cluster center which represents the tangent point of a line to the clearing function.

In the case where the supplier has a fixed capacity, the clearing function depends only on the workload. The cluster centers $(\omega_{t-1}^{'} + \rho_t^{'}, E(x_t^{'}))$ are then used to determine parameters of the approximation lines as follows:

$$A = \frac{\partial E(x_t)}{\partial(\omega_{t-1} + \rho_t)}(\omega'_{t-1} + \rho'_t) \tag{A.2}$$

$$B = E(x'_t) - A(\omega'_{t-1} + \rho'_t) \tag{A.3}$$

The clearing function depends on workload and capacity during the response time. In this case, the clearing function is estimated through a set of planes as it is indicated in Figure A.2.

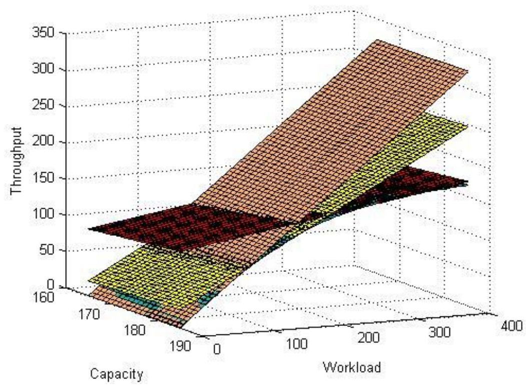


Figure A.2 The clearing function approximated by planes, V =3.

Therefore, the approximated throughput for the given work load $(\omega_{r,t-1} + \rho_{r,t})$, capacity level τ_t and capacity scenario (i, j) is computed as follows:

$$f(\omega_{r,t-1} + \rho_{r,t}, \tau_t) = \min\left[A_\upsilon(\omega_{r,t-1} + \rho_{r,t}) + B_\upsilon \tau_t - G_\upsilon\right] \qquad \upsilon \in \{1, 2, ..., V_{i,j}\} \tag{A.4}$$

The subtractive clustering is employed to determine the points where the planes are tangent to the clearing function. These points are the cluster centers $(\omega'_{r,t-1} + \rho'_{r,t}, \tau'_t, E(x'_t))$ which are used to specify the parameters of planes.

$$A = \frac{\partial E(x_t)}{\partial(\omega_{t-1} + \rho_t)}(\omega'_{r,t-1} + \rho'_{r,t}) \tag{A.5}$$

$$B = \frac{\partial E(x_t)}{\partial(\tau_t)}(\tau'_t) \tag{A.6}$$

$$G = A(\omega'_{r,t-1} + \rho'_{r,t}) + B\tau'_t - E(x'_t) \tag{A.7}$$

In this method, the estimation error is controlled by changing the cluster radius. As the cluster radius decreases, the estimation error improves as a result of the increase in the number of clusters. On the other hand, this leads to an increase in the number of approximation lines. In this thesis, the *subclust* function in MATLAB is employed to find the cluster centers. The parameters of this function are cluster radius, quash factor, accept ratio and reject ratio which are set to 0.5, 1.25, 0.5 and 0.15 accordingly. As a result of these settings, the average of the maximum error for all the clearing functions used in the numerical study is equal to 3.86.

Appendix B

In this section, the disruption scenarios are generated based on the assumption that there could be a maximum of one disruption within the planning horizon (T). The probability of each disruption scenario is computed through the Markov discrete time geometric distribution. Based on the Markov chain states, the DMS supplier has two states: Failure and operational. The transition probabilities from one state to another are as follows:

$$P\{Operational \rightarrow Operational\} = 1-\alpha \quad ; \quad P\{Operational \rightarrow Failure\} = \alpha$$
$$P\{Failure \rightarrow Failure\} = 1-\beta \quad ; \quad P\{Failure \rightarrow Operational\} = \beta \qquad \text{(B.1)}$$

The parameter α represents the probability of a failure state following an operational state and the parameter β represents the probability of an operational state following a failure state.

As illustrated in Figure B.1, the scenario with no disruption occurrence is created as a result of transition from one operational state to another consecutively.

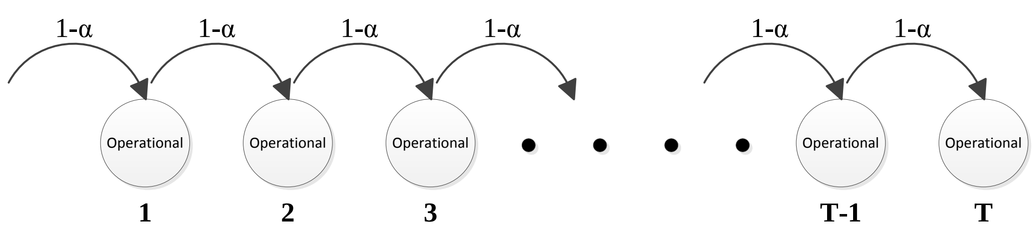


Figure B.1 The no-disruption scenario.

Therefore, the probability of the scenario with no disruption is:

$$P_{No\ disruption}(m,n) = \underbrace{(1-\alpha)(1-\alpha).....(1-\alpha)}_{T\ times} = (1-\alpha)^T \qquad \text{(B.2)}$$

The scenario corresponding to a disruption which occurs at time *m*, lasting *n* periods, is presented in Figure B.2.

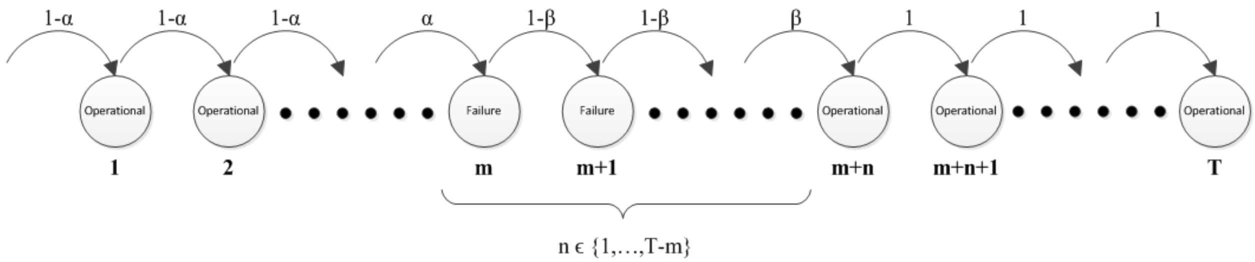


Figure B.2 A disruption scenario with time of occurrence = m, length = n .

This scenario is generated through the following transitions:

I. $m-1$ time(s) transition among operational states.

II. a transition from the last operational state to failure state.

III. $n-1$ time(s) transition among the failure states.

IV. a transition from the last failure state to operational state.

Hence, the probability of occurrence for the scenario with a disruption at time m lasting $n \in \{1,...,T-m\}$ is:

$$P_{Disruption}(m\ ,\ n) = (1-\alpha)^{m-1}\alpha(1-\beta)^{n-1}\beta \qquad \forall n \le T-m \qquad \text{(B.3)}$$

Since we have assumed that the disruption frequency within the planning horizon equals one, the DMS stays operational once it recovers from the failure state. Therefore, the transition probabilities after the end of disruption are equal to one.

For the scenarios where the DMS is in the failure state at period $T^{'}$ (Figure B.3), it could transit to the next period in either failure or operational status when an infinite number of periods are considered. Since we limit the evaluation of the catastrophic disruptions within the planning horizon *T*, the probability of such scenarios would be computed until the transition into the period *T*.

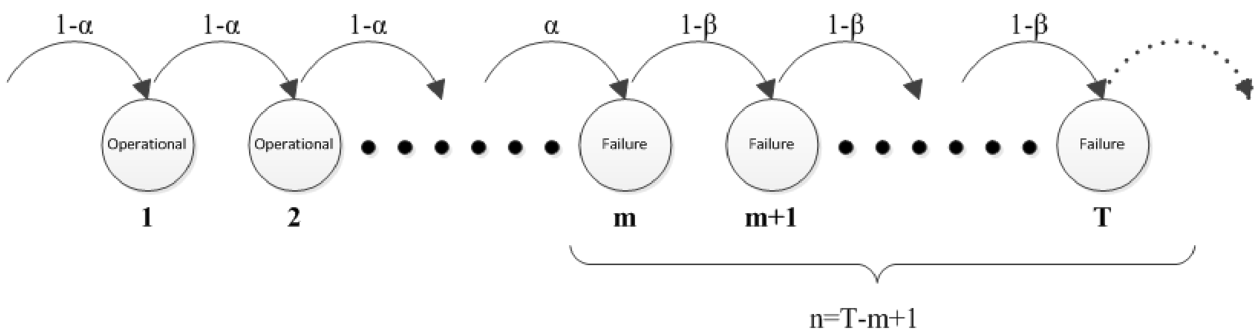


Figure B. 3 A disruption scenario with time of occurrence = *m*, *n+m=T+1*.

$$P_{Disruption}(m\ ,\ n) = (1-\alpha)^{m-1}\alpha(1-\beta)^{n-1} \qquad \forall n+m=T+1 \qquad \text{(B.4)}$$

Appendix C

Table C.1 Transition probabilities

State $f-1$	***State*** f	***Probability***
Main supplier available, Completely	Main supplier available, Partially	π
-	Main supplier unavailable	α
-	Low capacity availability during response time	P_L
-	Normal capacity availability during response time	P_N
-	High capacity availability during response time	P_H
Main supplier available, Completely or Partially	Back up supplier, low capacity availability during response time	αP_L
Main supplier available, Completely or Partially	Back up supplier, normal capacity availability during response time	αP_N
Main supplier available, Completely or Partially	Back up supplier, high capacity availability during response time	αP_H
Main supplier available, Completely	Main supplier available, Completely	$1-\alpha-\pi$
Main supplier unavailable	Main supplier available, Completely	β
Main supplier unavailable	Main supplier unavailable	$1-\beta$