An Analysis of Robustness and Flexibility in Freight Transportation Systems

by

Atikhun Unahalekhaka

B.S. (Electrical and Computer Engineering)
Carnegie Mellon University (2011)

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

Master of Science in Engineering Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Author . . . .
Engineering Systems Division
June 4, 2013

Certified by . . . .
Dr. Chris Caplice
Senior Lecturer, Engineering Systems Division
Executive Director, Center for Transportation and Logistics
Thesis Supervisor

Accepted by . . . .
Olivier L. de Weck
Professor of Engineering Systems and Aeronautics and Astronautics
Chair, Engineering Systems Division Education Committee

An Analysis of Robustness and Flexibility in Freight Transportation Systems

by

Atikhun Unahalekhaka

Submitted to the Engineering Systems Division
on June 4, 2013, in partial fulfillment of the
requirements for the degree of
Master of Science in Engineering Systems

Abstract

Freight transportation is a complex large scale system that operates under a highly dynamic and uncertain environment. Due to the scale and complexity of the system, a highly interdependent set of decisions are made across multiple planning levels. The interaction between tactical level decisions and execution level decisions largely determines the overall effectiveness of the system. This thesis aims to provide an analysis of how the interaction between the degree of robustness of a tactical plan and the flexibility level of an execution policy affects the performance of freight transportation systems in a dynamic and uncertain environment. Such analysis is conducted under two types of transportation systems, including a generalized distribution system and a military logistic system, that operate under demand uncertainty. Execution policies with different levels of flexibility – defined as the degree of freedom to which decisions can be adjusted at the executional level – are obtained by controlling the stickiness level of the different decisions as the decision maker transitions from the tactical plan to the executional plan. Robustness is used as a metric to measure the ability of a system to withstand random changes. Tactical plans with various degrees of robustness are obtained through the use of robust optimization. The performances of tactical plans with various degrees of robustness and execution policies with different levels of flexibility are evaluated through simulation. Results from the analysis on the distribution system show that the optimal level of robustness required from a tactical plan to achieve the lowest expected total cost decreases as the flexibility level of the system increases. Finally, the analysis on the military logistics system shows that the effect of increasing the flexibility level of the execution policy on the performance of the system depends on the uncertainty level of the demand.

Thesis Advisor: Dr. Chris Caplice
Title: Senior Lecturer, Engineering Systems Division & Executive Director, Center for Transportation and Logistics

Acknowledgements

I would like to express my sincere gratitude to my advisor Chris for his understanding and encouragement. This thesis would not have been possible without his guidance and support. I would also like to thank Francisco for his patience, kindness, and willingness to help me.

My gratitude goes to TRANSCOM, who has financially supported this research. Many thanks to the team at the Lincoln Labs, especially Masha and Allison, for introducing me to military logistics and for providing invaluable comments and feedback that have contributed greatly to my thesis.

I would like to thank all my friends at MIT for the wonderful moments we have shared. Special thanks to Lita and Yin-Jin for the friendship and support. Many thanks to my Thai friends in Boston and Cambridge for making me feel at home. I would like to especially thank Kanoon for all the memorable moments and the things that we have been through together. My gratitude also goes to my friends outside the US. Special thanks to Knight and Jew for always being by my side.

Most importantly, I would like to extend my deepest gratitude to my family members for their love and support. Thank you for everything.

This page intentionally left blank

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Freight transportation is a complex large scale system that operates under a highly dynamic and uncertain environment. The key to an effective operation of such system stems from an efficient management of a large set of transportation resources that range from human resources such as drivers and workers at sorting facilities to material resources such as freight, vehicles, and facilities. Due to the scale and complexity of the system, a highly interdependent set of decisions are made across multiple planning levels. The interaction between tactical level decisions and execution level decisions largely determines the effectiveness of the system. The focus of this thesis is on the analysis of how the interactions between tactical plans and execution policies affect the performance of a freight transportation system in a dynamic and uncertain environment.

1.1 Freight Transportation Systems

Freight transportation systems deal with the management of resources and assets to move freight between multiple origins and destinations in a way that achieves a certain performance metric level. Such system appears in a wide variety of context ranging from manufacturing companies that need to move raw materials and intermediate products between facilities or to distribute final products to vendors, to individuals that need to send letters or parcels across cities or countries, to military organizations that need to distribute goods to sustain and prolong various ongoing military operations around the globe. Companies and organizations can internally manage their own transportation operations or may outsource such operations to independent carriers such as railway, shipping, trucking, and parcel service companies.

The most basic goal of a typical freight carrier is to achieve a certain customer service level in the most cost effective manner. One of the primary challenges that these carriers face is the management of resources under a high degree of variability in operating conditions that stems from sources such as vehicle delays due to adverse weather conditions, mechanical breakdowns, and uncertain demand volumes. Accordingly, the effectiveness of a freight transportation operation relies heavily on the interplay between the tactical planning and the operational planning.

Planning at the tactical level deals with the allocation of types of assets or resources to prepare the system appropriately for future operations. Service network design is a critical part of the tactical planning process in consolidated transportation systems (i.e. systems where customer demands are consolidated to achieve economy of scale), where service routes and schedules are established ahead of time and are adjusted accordingly during execution. The types of problem that the service network design is concerned with are the determination of service types and frequencies, vehicle routes and schedules (e.g. design of aircraft schedules in parcel express carriers), traffic distribution or freight routes (e.g. load plan design in LTL operations), operating policies at terminals, and empty balancing strategies to reposition empty vehicles (Crainic, 2000). Tactical level decisions are typically made several months ahead of execution based on a representative planning period. Accordingly, decisions and commitments at this level must be made to accommodate a wide range of possible scenarios that are bound to vary across different time periods. Once established, a tactical plan is repeatedly executed in each time periods across the time horizon. For instance, LTL carriers typically execute the same freight routing decisions daily but makes appropriate adjustments in response to daily volume fluctuations. Operational planning, on the other hand, are made very frequently (daily, weekly, etc.) and determines the specific time-by-time allocation of specific vehicles to specific jobs based on the resources allocated by the tactical plan. There is very little randomness in these models since the missions are happening now and the data used is actual information.

Decisions made in these two planning levels are highly interdependent. The tactical plan provides the environment for the operational plan to operate on: tactical level decisions and commitments restrict the set of decisions that can be made by the operational plan. A good tactical plan should provide an operational plan with a lot of flexibility and options to respond to various unforeseen scenarios. This highlights the importance of considering uncertainty during the tactical planning process. A tactical plan that is generated based on point forecasts or estimate values may lead to an effective system level performance under a certain set of potential future scenarios but would invariably perform poorly as operating conditions deviate from the norm. But even if uncertainty is explicitly incorporated into the tactical planning process, there is still the challenge of addressing the trade-off between the robustness of a plan and the cost (i.e. a trade-off between risk and cost). For instance, under-allocating resources at the tactical level may yield lower upfront cost, but comes with the expense of potential costly adjustments during operation. On the other hand, over-allocating resource may ensure no costly adjustments at the operational level but may require an extremely high upfront cost.

1.2 Thesis Contribution

The overarching goal of this thesis is to investigate the different dimensions of incorporating uncertainty into the planning process of a freight transportation system. Accordingly, this thesis will analyze the following topics:

- Robust optimization method used to address demand uncertainty in transportation planning problems,
- Framework for analyzing the effects of the interactions between tactical plans and execution policies on the performance of a freight transportation system under uncertainty,
- Trade-off between the robustness of a tactical plan and the flexibility available within the execution system,
- Trade-off between the robustness and the cost for adding robustness in transportation planning problems.

1.3 Thesis Outline

The remainder of this thesis is organized as follows: Chapter 2 introduces the concepts of robustness and flexibility as means to address uncertainty and review approaches used for incorporating uncertainty into the planning process. Chapter 3 introduces a framework for analyzing the effects of the interactions between tactical plans with various degrees of robustness and execution policies with various levels of flexibility on the performance of a transportation distribution system. This framework is then applied to the military logistics context in the next two subsequent chapters. Chapter 4 provides background information on the United States Transportation Command and the planning process of the military airlift system. Chapter 5 then applies the framework introduced in Chapter 3 to a simplified network model motivated by the

Air Mobility Command's scheduling processes. Finally, Chapter 6 summarizes the findings from the study and suggests areas for future research.

Chapter 2

Robustness and Flexibility

2.1 Introduction

Flexibility and robustness are the two concepts that are widely used in the study of planning under uncertainty. In this work, flexibility refers to the ability for a system to react or adapt to unexpected changes during the executional phase. The flexibility level of a system is composed of two components – *systemic* flexibility and *operational* flexibility. *System*ic flexibility is the degree to which decisions can be adjusted, changed, or made at the operational level. Such degree of freedom to make adjustments to the plan in response to unforeseen changes is influenced by both exogenous and endogenous factors. For instance, due to restrictions that are beyond the control of the decision makers, it may not always be possible to delay making all decisions to the last minute. Certain resources must be committed and acquired long before execution. For example, air carriers may need to acquire airport slots (i.e. permission to take-off or land at the airport) months or years in advance. Furthermore, there may be a lead time for procuring some types of resources. Accordingly, acquiring these resources during the operational phase may not be feasible. Decision makers may also have some degree of control over this type of flexibility. For instance, airline decision makers could standardize their fleet type so that they have the option to swap aircraft during operation in the face of disruption. In addition, certain resources may be cheaper to procure early and the decision makers may choose to commit to acquiring such resources at the tactical level.

Given the systemic flexibility level that determines the dynamics of the decision making process (i.e. which decisions are committed at the tactical level, and which decisions can be made or adjusted at the operational level), operational flexibility provides the decision makers with options to respond and adapt to changes. Decision makers typically have control over this dimension of flexibility to a great extent by proactively embedding certain attributes into the plan during the resource allocation process. For instance, given that it is feasible for airline to swap aircraft during operation, operational flexibility can be embedded into a tactical plan by designing a schedule with swapping opportunities. Similarly, trucking companies may invest in acquiring extra fleets so that decision makers have more options to respond to demand fluctuations or mechanical breakdowns.

For a given systemic flexibility level, robustness is a metric used to measure the performance of a plan. A plan is considered to be robust if it is able to operate well under various operating conditions without requiring costly adjustments. This section considers two metrics for measuring the robustness of a plan. The first metric is concerned with the probability of feasibility. Infeasibility occurs when the tactical plan doesn't allocate enough resources to accommodate changes that occur at the operational level. Under such metric, a more robust plan is less sensitive to uncertainties and requires less costly adjustments. The second metric that is used to measure the robustness of a plan is the cost for fixing infeasibilities - the lower the cost, the more robust the plan. Under this metric, a robust plan may be more prone to disruption but the cost for fixing such disruption is minimal.

2.2 Methods for Adding Robustness

Planning models are typically formulated based on some assumed systemic flexibility level, as certain output decisions are designed to be fixed. Given the systemic flexibility level, the aim of these models is to create a plan that achieves a certain level of robustness. When there is completely no systemic flexibility (i.e. all decisions from the tactical level are executed without being changed at the operational level), robustness can be obtained by allocating resources appropriately so that the plan is insensitive to uncertainties. When there is some system flexibility, robustness can be obtained by adding operational flexibility into the plan. Under such setting, the level of operational flexibility in a plan corresponds directly to the robustness level of a plan. Fixing the systemic flexibility level, a plan is more robust with more operational flexibility. Regardless of which metric is used to measure robustness or how much system flexibility exist in the system, robustness always comes with a cost, and there is always a trade-off between robustness and cost. A plan that is extremely robust has a high upfront cost but low adjustment costs, whereas a plan that is not robust may have lower upfront cost but higher adjustment costs. Two classes of optimization methods for adding robustness into a plan are identified from the literature. The first approach explicitly incorporates knowledge of the future into the model to create robust plans that are able to withstand uncertainties and perform well under various unforeseen circumstances. The second approach doesn't utilize any future knowledge but creates plans that are less costly to adjust once infeasibility occurs. These two approaches are discussed in more details below.

2.2.1 Approaches that Utilize Future Knowledge

The key characteristic of this modeling approach is the explicit utilization of future knowledge in the model. The general aim of such model is to create plans that perform well under a wide variety of scenarios without breaking down, i.e. having a high probability of feasibility. Such approach can be further categorized into two classes based on the nature of information that is used in the model: deterministic and stochastic. The deterministic approach requires no probabilistic knowledge of the future, whereas the stochastic approach requires some probabilistic knowledge of the future. The modeling approaches for these two cases are briefly reviewed below.

2.2.1.1 Deterministic Approach

In some cases, it may be difficult to obtain the probability distribution of the different uncertain parameters in the system. For such cases, scenarios representing all the possible outcomes may be generated or bounded uncertain sets may be modeled to represent each uncertain parameter. The approach introduced by Kouvelis and Yu (1997) generated different possible scenarios reflecting the data uncertainty and aims to come up with a solution that works well for all the scenarios. Two methods were introduced for achieving such goal, with one optimizing against the worst instance that might arise due to data uncertainty and the other optimizing over the maximum deviation the optimal robust solution is from the optimal solution for each scenarios. The resulting solution is able to withstand uncertainty since it is designed to perform well under the worst case scenario.

A similar modeling approach that constructs a solution that is optimized against the worst data realizations that are taken from a bounded set has been proposed under the robust optimization framework. This framework can be traced back to the work of Soyster (1973) that considered linear programming models under data uncertainty. The work considered column wise uncertainty where the set of coefficients for each variable in the constraints belong to a certain set (i.e. the coefficients in the A_i column of the $Ax \leq b$ constraint belong to an uncertainty set). Such linear programming can be reformulated as an equivalent linear programming model, but leads to a worst case analysis where the solution is robust against all parameter instances and is too conservative to be used in real applications.

To address such conservativeness in the approach introduced by Soyster (1973), Ben-Tal and Nemirovski (1998, 1999) considered linear programming and convex problems under row wise ellipsoidal uncertainty set (i.e. the rows of the A matrix belong to an ellipsoidal uncertainty set). The approach restricted the uncertain coefficients to simultaneously be at their worst case values, and resulted in a solution that is less conservative than the solution that is obtained from the approach in Soyster (1973). However, the resulting robust counterparts for such problems are generally not linear. For instance, the works showed that linear programming problem under ellipsoidal uncertainty sets can be reformulated as a robust counterpart which is a second order cone problem.

Bertsimas and Sim (2002) addressed such drawback by proposing the concept of "budget of uncertainty" for linear programming problems under row-wise polyhedral uncertainty sets. Each uncertain coefficient is assumed to follow an unknown symmetric distribution, such as a triangular or uniform distribution, with a known expected value, upper bound value, and lower bound value. A Γ parameter was introduced into the uncertainty set to put a bound on the sum of the scaled deviation from the mean for each coefficient. Intuitively, the robust solution is deterministically feasible if at most Γ coefficients are at their worst case values and the remaining coefficients are at their expected values. In addition, by assuming symmetric distribution, the work was able to provide a probabilistic guarantee for feasibility under each Γ value. The attractiveness of this formulation is that the robust counterpart is a linear program. This model has been extended to discrete optimization and network flow problems in Bertsimas and Sim (2003).

Ben-Tal et al. (2004) applied the robust optimization framework to a two stage setting where certain decisions are made prior to the data realization and certain decisions are made after the data realization. To make the problem tractable, they proposed modeling the second stage variables as a linear function of the realized data. Such formulation yields a robust counterpart that is tractable under linear uncertainty sets, and produces a solution that is less conservative than that of a static single stage robust optimization formulation.

2.2.1.2 Stochastic Approach

The second approach requires probabilistic knowledge of the uncertain parameters to construct a solution that is robust against uncertainty. Several frameworks for addressing uncertainty in optimization models under this modeling approach have been proposed in the past. Stochastic programming has been extensively used to deal with uncertainty in optimization models. The most widely used and studied stochastic programming model is the two-stage model. Under such framework, decisions are made in two stages. Decisions in the first stage are made before the uncertainty materializes, and are influenced by the decisions made in the second stage after the uncertain parameters are realized. Accordingly, first stage decisions are assumed to be fixed regardless of how the uncertainty materializes. On the other hand, second stage decisions are made based on the fixed first stage decisions and the realization of the uncertain parameters. The aim of the model is to optimize the sum of a certain objective value based on the first stage design and the expected value of a certain objective function based on the second stage design. A generalization of the two stage model is the multi stage model where decisions are made at different stages after the unknown parameters are observed and the aim is to optimize sum of the expected cost at each stage.

Another approach that incorporates uncertainty into optimization model is the Chance Constrained Programming introduced by Charnes and Cooper (1959). In this approach, the probability of satisfying a constraint is restricted to be greater than a certain threshold. An alternative approach is to maximize the probability of achieving a certain objective value.

The stochastic dynamic programming method models the different possible realizations as states. With probabilistic information, the probability that a state transitions from one to another can be calculated, and dynamic programming can used to determine a policy that optimizes the decision made at each state over a certain time period. This is done by decomposing the problem into simpler sub-problems and recursively optimizing each sup-problem.

Finally, the work by Mulvey et al. (1995) incorporated scenario based approach with goal programming to find a robust solution. The work introduced two ideas of robustness. A solution is *solution robust* if its optimal value remains close for all scenarios. Similarly, a solution is *model robust* if the solution is almost feasible for all the scenarios. These two concepts are incorporated into the model by including both the cost function and the feasibility error function into the objective function. A constraint violation is allowed at a certain cost and is captured through the feasibility error function. To incorporate the decision maker's risk attribute into the objective function, it has been suggested to model the cost function using the mean/variance model or the expected utility model.

2.2.1.3 Related Work in Transportation Planning Work

This subsection presents a summary of some of the works that have applied the discussed techniques to transportation planning.

Lium et al. (2009) investigated the impact of incorporating demand uncertainty into the service network design. The aim of the paper is to get a better understanding of the qualitative differences between a stochastic design and a deterministic design. The study was carried out under a generalized problem of managing a fleet of homogenous capacitated vehicles to move demands from different origins to different destinations under time constraints. Demand uncertainty was addressed through the use of two stage stochastic programming. The first stage formulation determines the vehicle routing decisions, and aims to minimize the cost of empty and loaded vehicle movements and the expected cost based on the recourse decisions made in the second stage. Assuming that the vehicle routes are fixed from the first stage, the second stage formulation determines the number of ad-hoc capacity and cargo routes to deal with the different demand realizations. For each scenario, the aim of the second stage formulation is to deliver the realized demand in a way that minimizes the cost for using ad-hoc capacity. The two stage stochastic model was solved using scenario generation and was tested against a simple problem instance that includes six nodes and seven periods with three different demand uncertainty levels and three different demand correlations. It was shown that on average, the stochastic solution performed 17% better than the deterministic solution. The 2 major structural differences between a stochastic design and a deterministic design that were identified are consolidation and path sharing. Assuming that cargo routes are established before actual demands are known, consolidation provides a protection against uncertainty through the principle of risk pooling –

high demand volume from one customer tends to be offset by low demand volume from another customer. Furthermore, the stochastic design was observed to have more paths leading from the origin to the destination. Such attribute provides more flexibility for decision makers to switch cargo flows if necessary in the face of demand fluctuations. Hoff et al. (2010) extended the formulation and devised a heuristic method to solve an equivalent real life instance problem.

List et al. (2003) proposed a two stage stochastic optimization model to address the fleet sizing problem under demand and fleet productivity uncertainty. The focus of the model is in fleet acquisition and retirement decisions at each planning period subject to the variability in the demand and fleet operating conditions. The model is formulated using the two stage stochastic optimization framework where decisions relating to the fleet sizes, fleet acquisition, and fleet retirements are made in the first stage and fleet allocation decisions are made in the second stage after the demand volumes and fleet productivity levels are known. List et al. (2003) claimed that the traditional two stage stochastic optimization that minimizes the fixed cost of the first stage design and the expected cost of the second stage design may perform well on average but may perform poorly under extreme scenarios. To account for these extreme scenarios, the model added an additional constraint in the second stage formulation to capture the excess cost that goes beyond a certain cost threshold and penalizes such excess cost proportional to a "risk" parameter. The objective function of the first stage formulation then minimizes the sum of the first stage design cost that includes the fleet ownership, acquisition, and retirement costs and the expected value of the sum of the cost of the second stage design cost that includes the fleet operating cost, and penalty cost for not meeting demand, and the scaled penalty cost which is a function of the risk parameter. By adjusting this risk parameter, the model is able to analyze the tradeoff between the cost and risk.

Morton et al. (2002) proposed a multistage stochastic programming model for creating sealift deployment plans that are robust against potential attacks. The aim of the model is to manage a fleet of ships to transport cargos from the port of embarkation (POE) to the port of debarkation (POD), and forwarded to the final destination through ground transportation within the required time window. The model assumes that only a single attack will occur over the planning horizon period, although the attack may affect several ports of debarkations. In the face of an attack, the affected POD(s) shut down for a certain period and inbound and outbound flow of ships are prohibited. Under such circumstances, the model allows ships to be re-routed from one POD to the other, given that the ship has not entered the POD but is just waiting outside. Using probabilistic knowledge on the timing, location, and severity of the attack, the model seeks to minimize the sum of the expected late penalty for transporting all cargos the final destination, expected penalty for undelivered cargos, penalty for unnecessary ship voyage, and penalty for unnecessary ship rerouting.

Lan et al. (2006) presented two modeling approaches for achieving robust airline schedule plans. The first approach involves routing aircraft to minimize the expected propagated delay. The arrival delay of a given flight is broken down into propagated delay and non-propagated delay. Propagated delay is a delay that occurs when the aircraft is delayed because of a delay in its prior leg and is a function of the aircraft's routing. Non-propagated delay is a delay that occurs for any reasons apart from aircraft routing and is independent of the aircraft's routing. The expected length of non-propagated delay for each flight may be obtained from past historical data. Given the non-propagated delay for all the flights in a given sequence, the expected total propagated delay for such sequence of flight may be obtained iteratively. Then from all the feasible sequence of aircraft routes, the objective is to select the sequence of flight routes that that minimizes the expected total propagated delay. The second approach assumes the aircraft route is fixed and develops a method to adjust the departure time within a small time frame of each flight to minimize the number of passenger misconnection. The idea is to allocate the slack time for the existing routing to minimize passenger misconnection without the need for requiring more aircraft. A similar approach was presented in Ahmadbeygi et al. (2009). The aim of this work is to minimize the delay propagation by redistributing the scheduled slack in an existing plan, where the propagation delays are modeled using a propagation tree.

Mudchanatongsuk et al. (2008) addressed the multi-commodity flow network design problem with single source and single sink under transportation cost and demand uncertainty using the adjustable robust optimization approach. Capacity allocation decisions are made in the first stage prior to the realizations of the uncertain data, and the flow for each commodity is made in the second stage and depends on the realization of the data. Similar to the approach that this work will be presenting in chapters 3 and 5, the cargo flows were modeled as fractional flows of the realized demand values. The work considered demand and transportation costs under polyhedral and ellipsoidal uncertainty sets and outlined a column generation method for solving the robust counterpart for the formulation under polyhedral transportation cost uncertainty. The solution was validated against a set of simulated data taken uniformly from the uncertainty set. The simulation results suggested that the robust solution performed only marginally better than the deterministic solution with respect to the mean cost and the standard deviation for small network structures, but performed relatively better under larger networks especially in the case with high uncertainty level.

This subsection provided a survey of the different approaches for incorporating uncertainty into the transportation planning process based on the techniques introduced in the previous two subsections. Accordingly, these approaches can be categorized based on the nature of the information that is used. The approaches presented by Lium et al. (2009), Morton et al. (2002), List et al. (2003), and Lan et al. (2006) all required some probabilistic knowledge regarding the uncertain parameters. Lium et al. (2009) applied the two stage stochastic programming framework to address demand uncertainty in a generalized service network design problem. The results suggested that the stochastic design has more consolidation between the commodity types and more paths leading from the origin to destination for each commodity types than that of the deterministic design. Morton et al. (2002) proposed a multistage stochastic programming model for creating sealift deployment plans that are robust against potential attacks. List et al. (2003) applied the two stage stochastic optimization framework to address the fleet sizing problem under demand and fleet productivity uncertainty. To account for the extreme scenarios that the solution obtained from traditional two stage stochastic optimization may perform poorly on, the model added an additional constraint in the second stage formulation to capture the excess cost that goes beyond a certain cost threshold and penalized such excess cost proportional to a "risk" parameter. Lan et al. (2006) presented two modeling approaches for achieving robust airline schedule plans. The first approach involved routing aircraft to minimize the expected propagated delay. Fixing the aircraft routes, the second approach developed a method to adjust the departure time within a small time frame of each flight to minimize the number of passenger misconnections. The approach proposed by Mudchanatongsuk et al. (2008), on the other hand, applied the robust optimization framework to address the multi-commodity flow network design problem under transportation cost and demand uncertainty, and didn't utilize any probabilistic information. In such approach, uncertainties were modeled as polyhedral and ellipsoidal uncertainty sets, and the solution was constructed to be deterministically feasible for all data realizations that are taken from these sets.

2.2.2 Approaches that do not Utilize Future Knowledge

This section presents the second modeling framework for adding operational flexibility and robustness into a plan. Although the framework doesn't explicitly incorporate future knowledge into the model, it provides decision makers with options to minimize recovery cost or the impact of disruptions. Such framework is widely seen in the robust airline scheduling literature and can be further categorized into two approaches based on the objective of the model. The first approach aims to provide the decision makers with more recovery options in the face of disruptions. This is achieved by explicitly adding more opportunities for the decision makers to perform certain recovery tools or by lowering the cost of performing some of the recovery tools. The second approach isolates the effect of disruption so that the impact is minimized. This is generally achieved by partitioning the schedule into independent sub-schedules so that disruptions in one sub-schedule don't propagate to the other. This is especially useful for alleviating the impact of airport closure or reduction in airport capacity where a lot of aircraft may be disrupted. The methods for embedding operational flexibility into the system based on these two approaches are briefly discussed in the next three subsections.

Section 2.2.2.1: Maximize Recovery Options

The work by Ageeva (2000) introduced the concept of aircraft swapping opportunities as a mean to incorporate flexibility into airline schedules during the aircraft routing stage. There is an aircraft swapping opportunity when two aircraft meet twice between their scheduled maintenance visits. This gives the decision makers an opportunity to swap an aircraft during disruption and then later swap back to normal operations.

The model created a flexible plan with an optimal solution by generating multiple optimal solutions and selecting the solution with the highest number of aircraft swapping opportunities. Multiple optimal solutions are obtained by re-solving the aircraft routing model and adding new constraints each time to exclude previously chosen solutions. The aim of the aircraft routing problem is to assign routes (sequence of connecting flight legs) to specific aircraft so that the maintenance requirements are met at the lowest total maintenance cost. The maximum required flight time between maintenance visits is three to four days but airlines generally require a maintenance check every 40-45 hours.

When the number of flights and aircraft in consideration is large, it is possible to obtain several optimal solutions. For instance, consider the schedule in Table 2.1. Assume that this schedule is a segment taken from a hypothetical optimal aircraft routing solution which requires Aircraft 2 to perform a maintenance check at Station 4 at the end of Monday. By reassigning flights to different aircraft and ensuring that the maintenance requirement for Aircraft 2 is met, it is possible to obtain a more flexible schedule at no additional maintenance cost as shown in Table 2.2. As illustrated in Figure 2.1 and Figure 2.2, the schedule in Table 2.2 is more flexible than the schedule in Table 2.1 because there is an opportunity for Aircraft 1 and Aircraft 3 to swap at Station 1 on the Monday and then swap back at Station 3 the following day. Having an optimal schedule with the greatest number of aircraft swapping opportunities provides the decision makers with more recovery options, thereby making the recovery process easier and less costly.

Shebalov and Klabjan (2006) introduced the idea of crew move-up as a mean of introducing flexibility to the airline schedule during the crew assignment process. This concept is analogous to the idea of aircraft swapping presented by Ageeva (2000). There is an opportunity for a crew-move up when two crews from different flights can feasibly be swapped. In order for a crew swapping to be feasible, the two involved crews must overlap at an airport within a certain time threshold, start from the same crew base, and finish their respective itineraries at the same day and at the same crew base.

To incorporate this concept into the crew assignment model, an additional objective of maximizing the number of move-up crews is introduced. Note that there is a clear tradeoff between maximizing the number of move-up crew and minimizing the crew assignment cost since an optimized crew assignment is more likely to have fewer number of move-up crews. To better manage this tradeoff, the conventional crew assignment model was first solved to obtain the most optimal crew assignment cost. Then the number of move-up crews is maximized with the additional constraint of limiting the total crew assignment cost to be below a certain factor of the optimal crew assignment cost.

Based on an experimental analysis using simulation and a crew recovery decision support system on a representative airline data under different random disruption scenarios, the model was able to increase the number of crew move-up by 4 folds from increasing the optimal cost by

5%. It has also been observed that it is more beneficial to have a few crew move-ups for many flights than having many crew move-ups for a few flights. This is captured in the model by adding a constraint that limits the number of crew move-up for each flight. By adjusting this parameter along with the cost factor parameter, the flexible crew model was able to achieve a lower operational cost, fewer deadheads, fewer uncovered legs compared to the conventional crew model.

	Flight	Origin	Destination	Departure Time	Arrival Time	Day
Aircraft 1						
	1	Station 2	Station 1	15:00	16:00	
	2	Station 1	Station 3	17:00	19:00	Monday
	3	Station 3	Station 2	20:00	22:00	
	4	Station 2	Station 4	8:00	9:00	
	5	Station 4	Station 3	10:00	11:00	Tuesday
	6	Station 3	Station 2	12:00	14:00	
Aircraft 2						
	7	Station 4	Station 1	15:00	16:00	
	8	Station 1	Station 2	18:00	19:00	Monday
	9	Station 2	Station 4	21:00	22:00	
			Maintenance Check			
	10	Station 4	Station 2	8:00	9:00	
	11	Station 2	Station 1	10:00	11:00	Tuesday
	12	Station 1	Station 2	13:00	14:00	
Aircraft 3						
	13	Station 2	Station 3	15:00	17:00	
	14	Station 3	Station 2	18:00	20:00	Monday
	15	Station 2	Station 1	21:00	22:00	
	16	Station 1	Station 3	8:00	10:00	Tuesday
	17	Station 3	Station 4	12:00	13:00	

Table 2.1: Flight schedule for 3 aircraft taken from a hypothetical optimal routing solution

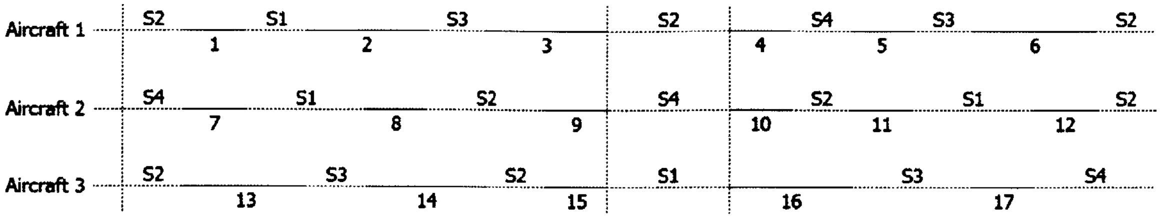

Figure 2.1: Illustration for the schedule in Table 1. Note that there is no swapping opportunity in this case. It is not possible for Aircraft 1 and Aircraft 2 to swap at Station 1 on Monday and swap back at station 2 on Tuesday because of the maintenance requirement of Aircraft 2 at Station 4 at the end of Monday.

	Flight	Origin	Destination	Departure Time	Arrival Time	Day
Aircraft 1						
	1	Station 2	Station 1	15:00	16:00	Monday
	2	Station 1	Station 3	17:00	19:00	
	3	Station 3	Station 2	20:00	22:00	
	4	Station 2	Station 4	8:00	9:00	Tuesday
	5	Station 4	Station 3	10:00	11:00	
	6	Station 3	Station 2	12:00	14:00	
Aircraft 2						
	13	Station 2	Station 3	15:00	17:00	Monday
	14	Station 3	Station 2	18:00	20:00	
	9	Station 2	Station 4	21:00	22:00	
			Maintenance Check			
	10	Station 4	Station 2	8:00	9:00	Tuesday
	11	Station 2	Station 1	10:00	11:00	
	12	Station 1	Station 2	13:00	14:00	
Aircraft 3						
	7	Station 4	Station 1	15:00	16:00	Monday
	8	Station 1	Station 2	18:00	19:00	
	15	Station 2	Station 1	21:00	22:00	
	16	Station 1	Station 3	8:00	10:00	Tuesday
	17	Station 3	Station 4	12:00	13:00	

Table 2.2: New schedule with added flexibility

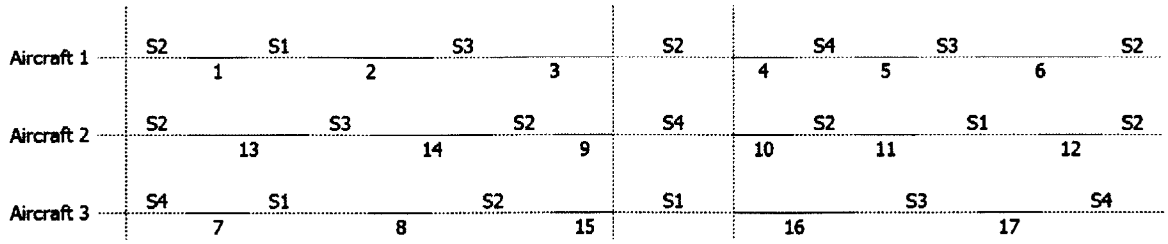

Figure 2.2: Illustration for the schedule in Table 2. There is an opportunity for Aircraft 1 and Aircraft 3 to swap at Station 1 on Monday and later swap back at Station 3 on Tuesday.

Smith and Johnson (2006) introduced the idea of imposing station purity to the fleet assignment model to make the solution more flexible to crew planning, maintenance planning, and operations. In the formulation, the number of fleet types serving a given non-hub airport is limited to be below the airport's "purity level". By limiting the number of fleet type serving a non-hub station, the number of airports visited by each feet type decreases and the number of lonely fleets (fleet that serves a non-hub airport only once or twice during the planning horizon) decreases. Commonality in fleet types at the airport increases swapping opportunities for both the aircraft and crews and also decreases maintenance cost due to the reduction in the need for spare parts. This concept was captured in the fleet assignment model by limiting the number of fleet types serving each station. Two purity levels were experimented with the maximum purity level being set to the minimum required number of fleet types at an airport and the moderate purity level being set to 1 or 2 depending on the size of the airport.

The impact of imposing station purity into the schedule was tested against a set of disruption scenarios on a published data of a mid-sized international carrier and a US domestic carrier. Imposing purity decreases the expected revenue from as low as 1% to as high as 20% for both the maximum and moderate purity levels. However, significant reduction in maintenance cost and crew cost contributes to an increase in the total expected profit. Although no operational analysis was performed, it is expected that introducing station purity would help save recovery cost, thereby further increasing the total profit.

Section 2.2.2.2: Isolate Impact of Disruption

Kang (2004) introduced the idea of a degradable schedule to isolate the effect of disruptions. Flight legs are partitioned into several independent sub-networks with different priority levels.

Sub-networks are prioritized based on the revenue of the flight legs they contain with the maximum revenue sub-network having the highest priority. In the face of disruptions, the order that delays and cancellations are assigned to is the reverse of the order of the priority (i.e. flight legs in the sub-network with the lowest priority level are first delayed or cancelled). Degradability can be incorporated into the schedule at the flight scheduling, fleet assignment, or aircraft routing stage of the airline scheduling process. In any case, aircraft and crews are assigned independently for each sub-network. In order to ensure that there are enough flight legs in a sub-network so that the routing problem is feasible, there is a maximum limit of number of flight legs in a sub-network. Since this problem was designed to hedge against disruptions caused by reduction in airport capacity due to severe weather, the maximum number of flight legs in the highest priority sub-network was assigned to be less than the maximum airport capacity under bad weather. The aim of the model is to assign the flight legs into different sub-networks such that the total weighted revenue is maximized.

The performance of a degradable schedule was tested using the MIT Extensible Air Network Simulation (MEANS) and a recovery system to simulate the operation under a representative normal weather day and bad weather day to determine the delays and recovery costs involved. The tested flight schedule was based on a representative airline data and was decomposed into 2 independent sub-networks. The simulation results show that the degradable schedule performed comparably to the traditional schedule under good weather but experiences more flight cancellations under bad weather. However, since most of the cancellations from the degradable schedule are from the lower revenue generating flights in the second sub-network, the total recovery cost for the degradable schedule is still lower than that of the conventional schedule.

Section 2.2.2.3: Maximize Recovery Options/ Isolate Impact of Disruption

Rosenberger et al. (2004) incorporated flexibility into the fleet assignment model by considering cancellation cycles and hub connectivity. In the face of disruptions or irregular operations, flight cancellations may be unavoidable. Under such circumstances, airline decision makers are usually forced to cancel a flight cycle when cancelling a single flight to maintain flight balance. This cycle is known as a "cancellation cycle" and is defined as a sequence of legs in a rotation in which the first leg departs from the same airport as which the last leg lands. An aircraft rotation with a shorter cancellation cycle decreases the impact of disruption since it allows the airline decision makers to cancel smaller number of flights while maintaining flow balance. Since flight cancellations become less costly, the number of attractive recovery options increases.

To illustrate the idea, consider the two assignments of Aircraft 1 shown in Table 2.3 and Table 2.4. Suppose that on the day of operation, Aircraft 1 breaks down before operating Flight 1 at Hub 1 and won't be available until 12:00. Assignment 1 would require the decision maker to cancel Flights 1, 2, 3, and 4 (cancellation cycle). On the other hand, Assignment 2 would require the decision maker to cancel just Flights 1 and 2 (cancellation cycle). In this case, Assignment 2 has shorter cancellation cycle and is more flexible than Assignment 1.

In addition, since airline schedules are particularly sensitive to disruption at hubs, flexibility may be added by isolating hubs from the schedule. Hub isolation may be measured by "hub connectivity" which is defined as the number of legs in a rotation where the aircraft begins at a hub, visits spokes, and ends at another hub. Airline schedules with lower hub connectivity are less sensitive to disruptions. Assignment 1 has a hub connectivity of 8 (each route has a hub connectivity of 2) whereas Assignment 2 has a hub connectivity of 0. So in this case, Assignment 2 is more flexible than Assignment 1. To illustrate this idea, suppose that Hub 1 gets shut down for the day, then the only flight that could operate normally under Assignment 1 would be Flight 5. On the other hand, Flights 5, 2, 3, and 8 would still be able to operate normally under Assignment 2. Overall, Assignment 2 is more flexible than Assignment 1 because it has shorter cancellation cycle and lower hub connectivity.

The work related these two concepts together by showing that airline schedules with limited hub connectivity have shorter cancellation cycles. This translates to either limiting the total hub connectivity or maximizing the number of short cancellation cycles in the fleet assignment model. Two optimization models for incorporating these concepts into the schedule were presented. The first model minimized the total hub connectivity in the first stage then it minimized the fleet assignment cost while constraining the hub connectivity to be below a certain factor of the minimized hub connectivity in the second stage. The second model first minimized the fleet assignment cost then it minimized hub connectivity while constraining the fleet assignment cost to be below a certain factor of the optimal fleet assignment cost.

The performance of the flexible fleet assignment model was tested using the simulation of airline operations (SIM) and a recovery system on 3 different sets of representative airline data against different disruption scenarios. The aircraft routings for each of the fleet assignments were assigned using a simple routing algorithm and the crew assignment was excluded from the analysis. The results from the simulation show that the results from the flexible fleet assignment models yielded fewer cancellations and better on time performance compared to the conventional fleet assignment model.

Route	Flight	Origin	Destination	Departure Time	Arrival Time
Aircraft 1					
11	1	Hub 1	Spoke	8:00	10:00
	2	Spoke	Hub 2	11:00	12:00
12	3	Hub 2	Spoke	13:00	14:00
	4	Spoke	Hub 1	17:00	19:00
Aircraft 2					
13	5	Hub 2	Spoke	8:00	9:00
	6	Spoke	Hub 1	11:00	13:00
14	7	Hub 1	Spoke	14:00	16:00
	8	Spoke	Hub 2	17:00	18:00

Table 2.3: Schedule Assignment 1. Note that each route has hub connectivity of 2 so the total hub connectivity for this assignment is 8.

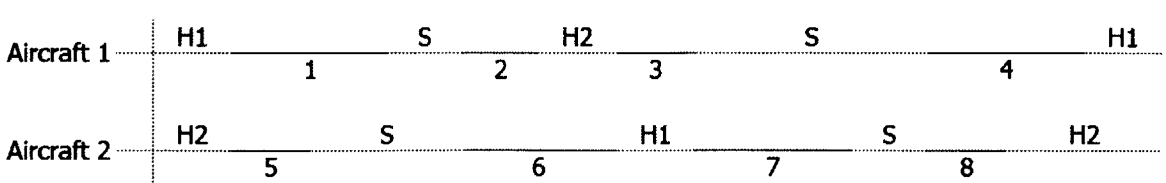

Figure 2.3: Illustration of schedule in Table 2.3

Route	Flight	Origin	Destination	Departure Time	Arrival Time
Aircraft 1					
11	1	Hub 1	Spoke	8:00	10:00
	6	Spoke	Hub 1	11:00	13:00
12	7	Hub 1	Spoke	14:00	16:00
	4	Spoke	Hub 1	17:00	19:00
Aircraft 2					
13	5	Hub 2	Spoke	8:00	9:00
	2	Spoke	Hub 2	11:00	12:00
14	3	Hub 2	Spoke	13:00	14:00
	8	Spoke	Hub 2	17:00	18:00

Table 2.4: Schedule Assignment 2. Note that the total hub connectivity for this assignment is 0

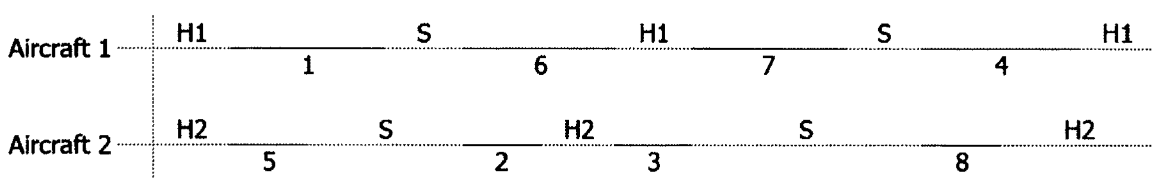

Figure 2.4: Illustration of schedule in Table 2.4

Section 2.2.2.4: Summary

The previous three subsections provided a summary of some of the methods for adding operational flexibility into airline schedules. Different approaches were classified based on two objectives: maximize recovery options and isolate impact of disruption. The major flexibility attributes that are embedded into airline schedules to maximize recovery options are based on the idea of swapping resources –aircraft or crews – in the face of disruption. The approach introduced by Ageeva (2000) generated multiple aircraft routing optimal solutions and selected the solution with the highest number of aircraft swapping opportunity. Shebalov and Klabjan (2006) maximized the number of crew swapping opportunities while constraining the total crew assignment cost to be below a certain factor of the optimal crew assignment cost. Smith and Johnson (2006) introduced the concept of "station purity" that puts a limit on the number of fleet types that can be accommodated at each airport. This increases the commonality of fleet types at each airport, and thereby increases swapping opportunities for both aircraft and crews. The concept of swapping resources to hedge against uncertainty is appropriate for dealing with small scale flight delays or when flight delays do not exceed the specified time threshold that the two resources overlap. However, this method is not appropriate for protecting the system from large scale flight delays and cancellations due to sources such as major mechanical breakdowns or severe weather conditions that lead to airport closures.

Kang (2004) introduced the idea of a degradable schedule to isolate the effect of disruptions. Flight legs are partitioned into independent sub-networks with different priority levels. The priority level is used to determine the order that flight legs are delayed and cancelled in the face of disruptions. This concept was designed to hedge against disruptions caused by reduction in airport capacity due to severe weather, and the maximum number of flight legs in the highest priority sub-network was assigned to be less than the maximum airport capacity under bad weather. Rosenberger et al. (2004) introduced the concept of hub connectivity and cancellation cycle as means to simultaneously make certain recovery option cheaper and to isolate the impact of disruption. Schedules with lower hub connectivity are less sensitive to disruptions due to hub closure, and schedules with shorter cancellation cycle are less costly to cancel flights. The work showed that airline schedules with limited hub connectivity have shorter cancellation cycles, and introduced a method to reduce hub connectivity in fleet assignment model. As opposed to the concept of swapping resources in the face of small scale flight delays, these two concepts of hub connectivity and cancellation cycle are more suited for protecting the system against larger scale disruptions due to sources such as airport closures and mechanical breakdowns.

Chapter 3
Analysis of Robustness and Flexibility in a Distribution System

The effectiveness of a transportation operation is highly dependent on the interaction between a tactical plan and an execution policy. An execution policy determines the setting for which a tactical plan gets translated into an executable plan under various operating conditions. The execution policy is composed of two components: *systemic* flexibility and recourse action. The systemic flexibility determines the degree to which decisions can be adjusted, changed, or made at the operational level. The recourse policy determines the set of allowable actions that are used to fix infeasibilities during operation. The sets of recourse actions typically vary across different transportation fields and may take the form of outsourcing excess demand to third party companies in trucking transportation services or delaying and cancelling flights in the airline industry. Regardless of the recourse action, there is always a certain penalty cost associated with fixing infeasibilities.

The aim of this chapter is to study the interaction and trade-off between the level of systemic flexibility and the level of robustness of a tactical plan in a distribution system. From this point on, the terms *systemic flexibility* and *flexibility* will be used interchangeably. The metric that will be used to measure robustness in this chapter is the probability of feasibility. Under this metric, a plan is more robust if it is less sensitive to uncertainties and requires less costly adjustments. Consider a transportation system that is operating at some initial flexibility level. At the one extreme, a very rigid system with no flexibility may require a tactical plan with a certain level of robustness to achieve a high level of performance. However, as the flexibility level of the system increases, a tactical plan with a lower degree of robustness (relative to the initial flexibility level) might achieve the same or even better level of performance since more decisions are allowed to be changed at the operational level. Assuming that the degree of robustness of a plan is measured with respect to a system with no flexibility (i.e. all tactical decisions are fixed and must be executed at the operational phase), and the metric for measuring the performance of the system is the final operating cost, the questions that this chapter aims to address are:

- Under the system with no flexibility, how robust should the tactical plan be?
- Given that the flexibility level of the system increases, how does the appropriate level of robustness in the tactical plan change?
- How do the performances of the system compare under various levels of flexibility and different degrees of robustness of a tactical plan?

3.1 Overview

This chapter presents a framework for analyzing the effect of the interaction between the degree of robustness of a tactical plan and the flexibility level of an execution policy on the performance of the system. The study will be conducted under a generalized hypothetical distribution problem with demand uncertainty, and the analysis will be achieved through simulating tactical plans with different degrees of robustness under different execution policies with varying degrees of flexibility. To simplify the analysis, the study assumes that decisions are only made in two stages. A tactical plan is generated prior to the demand realizations. After the demand values are simultaneously realized, an execution model is then used to translate the tactical plan into an executable plan based on the execution policy. To accommodate the analysis that requires different flexibility settings, it is convenient to consider a strategic/tactical level problem that is concerned with the determination of the major decisions - fleet size, vehicle routes, and cargo routes – that are required in a distribution problem. In this chapter, vehicle routes refer to the paths that the vehicles take to fulfill customer requirements and cargo routes refer to the paths that lead each cargo from its corresponding origin to destination. Execution policies with various degrees of flexibility can be achieved by adjusting the degrees to which these three decisions at tactical level are allowed to change at the execution level.

Crainic (2000) distinguished "frequency" " service network design models from "dynamic" service network design models in the tactical planning process. "Frequency" service network design models addresses strategic/tactical level decisions such as the determination of type and frequency of services over the planning horizon period. This type of model is appropriate under steady state setting, where demands are generally assumed to follow the same pattern uniformly over the planning horizon. Under such assumption, the model is typically formulated using static spatial network representation where the time aspect can be captured in the definition of the service. On the other hand, "Dynamic" service network design models are more concerned with design of operating schedules that captures the level of details that are required for an actual operation. This type of model captures detailed movements of cargos and vehicles through space and time, and is typically formulated using time space network representation. The model that will be considered in this chapter fits into the description of the frequency service network design model.

To perform an analysis that addresses the trade-off between robustness and flexibility, there needs to be a metric for measuring the flexibility level of a system and the robustness level of a tactical plan. In this study, the flexibility level is discretized into three levels depending on the degree to which the three decisions at tactical level are allowed to change at the execution level. In the least flexible case, all the tactical level decisions are fixed and can't be changed at the executional level. In the most flexible case, the vehicle routes and the cargo routes are allowed to be re-optimized. The degree of robustness will be measured using the concept of budget of uncertainty that was introduced in Bertsimas and Sim (2002). Bertsimas's robust optimization approach provides a systematic method for adding extra buffer and capacity with different degree of robustness that can controlled by a Γ parameter. The robustness level of a tactical plan can then be measured based on the Γ values - the higher the Γ value, the higher the probability of the plan being feasible and thus the higher the robust level. In this study, such degree of robustness is measured with respect to the lowest flexibility level of the system, where all tactical decisions are assumed to be fixed.

The problem description and a deterministic tactical planning model are first introduced in Section 3.2. A robust formulation that applies the concept of budget of uncertainty (Bertsimas & Sim, 2002) to address demand uncertainty in the planning model is presented in Section 3.3. Section 3.4 introduces the different executing policies with various flexibility levels and their corresponding execution optimization model. Finally, Section 3.5 uses the different components that were presented to introduce a framework for analyzing the trade-off between robustness and flexibility.

3.2 Deterministic Formulation

This section presents a deterministic distribution planning model for a representative planning period. Under the steady state assumption that the demands follow the same pattern uniformly over the planning horizon, the problem can be modeled as a multi-commodity network optimization problem. Given a transportation network structure defined by (N, E), where N is the set of nodes and E is the set of arcs, the decision maker is faced with the problem of determining the optimal fleet size, vehicle routes, and the cargo routes to accommodate a set of requirements, each of which is characterized by its demand volume$(\bar{d}_r)$, origin node$(source_r)$, and destination node$(sink_r)$, subject to spatial and capacity constraints. The model assumes that cargo processing times at transshipment points are instantaneous and that the travel times between nodes are known with certainty. Vehicles of different fleet types differ in the capacity but are assumed to have the same travel times between nodes. Finally, since the formulation is modeled using static network that neglects the time dimension, consolidation opportunities may be overestimated. The deterministic formulation is given below:

Set:

N : Set of node in the network

E : Set of arc in the network

A : Set of fleet type

R : Set of requirements

Input Data:

$\bar{d}_r$: Expected demand volume for requirement r

$source_r$: Origin node for requirement r

$sink_r$: Destination node for requirement r

$\tau_{i,j}$: Time for each vehicle to travel from node i to node j

Cap_a : Capacity for vehicle type a

γ_a : Fixed cost for operating one unit of vehicle type a

$\varphi_{a,i,j}$: Cost for moving vehicle type a from node i to node j

$\beta_{r,i,j}$: Cost for moving one unit of requirement r from node i to node j

T : Planning horizon period

Variable:

V_a : Fleet size for vehicle type a

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

The deterministic formulation(NPM) is presented below:

$$minimize_{V,x,f} \sum_{a \in A} \gamma_a \times V_a + \sum_{a \in A} \sum_{(i,j) \in E} \varphi_{i,j} \times x_{a,i,j} + \sum_{r \in R} \sum_{(i,j) \in E} \beta_{r,i,j} \times \bar{d}_r \times f_{r,i,j}$$

st:

$$\sum_{(i,j) \in E} f_{r,i,j} - \sum_{(j,i) \in E} f_{r,j,i} = \begin{cases} 1 \ , j = sink_r \\ -1, j = source_r \\ 0, otherwise \end{cases} \quad \forall r \in R, j \in N \tag{3.2.1}$$

$$\sum_{(i,j) \in E} x_{a,i,j} = \sum_{(j,i) \in E} x_{aj,i} \quad \forall j \in N, a \in A \tag{3.2.2}$$

$$\sum_{r \in R} \bar{d}_r \times f_{r,i,j} \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \quad \forall (i,j) \in E \tag{3.2.3}$$

$$\sum_{(i,j) \in E} \tau_{i,j} \times x_{a,i,j} \leq V_a \times T \quad \forall a \in A \tag{3.2.4}$$

$$f_{r,i,j} \geq 0 \quad \forall (i,j) \in E, a \in A \tag{3.2.5}$$

$$x_{a,i,j} \geq 0, x_{a,i,j} \in \mathbb{Z} \quad \forall (i,j) \in E, a \in A \tag{3.2.6}$$

$$V_a \geq 0, V_a \in \mathbb{Z} \quad \forall a \in A \tag{3.2.7}$$

The objective function minimizes the sum of the fixed fleet size cost, the vehicle flow cost, and the cargo flow cost. Constraint (3.2.1) ensures a conservation of cargo flow for each requirement at the source node, where there is an artificial supply of one, the sink node, where there is an artificial demand of one, and the regular node, where the inflow must equal to the outflow. Constraint (3.2.2) enforces the conservation of flow for each vehicle type. Constraint

(3.2.3) is the capacity constraint that limits the number of cargos flowing through the arc to be less than the capacity that is determined by the number of vehicles flowing through it. Constraint (3.2.4) makes sure that the total vehicle usage time is less than the total vehicle available time for each fleet type. Constraints (3.2.5), (3.2.6), and (3.2.7) are non-negativity constraints. In this formulation, the vehicle routing and the fleet size decision variables are modeled as integer variables, whereas the fractional cargo routing variables take on values between 0 and 1.

3.3 Robust Formulation

In this section, a robust formulation that addresses the uncertainty in the demand volume in each requirement is considered. Each demand value is assumed to follow an unknown distribution that has a mean value $\bar{d}$ and an upper bound value $\bar{d} + \hat{d}$. The uncertainty set is constructed based on the concept of budget of uncertainty (Bertsimas & Sim, 2002), where a Γ parameter is used to limit the sum of the scaled deviation from the mean for each of the demand values. Intuitively, the robust model constructs a solution that is deterministically feasible if at most Γ of the demand values are at their worst case values while the other demand values are at their nominal values. The uncertainty set used in the formulation is given by(3.3.1).

$$\mathrm{U}^{D}(\Gamma) = \begin{Bmatrix} \tilde{d}_i = \bar{d}_i + \hat{d}_i z_i, \;\; i \in R \\ \sum_{i \in R} z_i \leq \Gamma \\ 0 \leq z_i \leq 1, i \in R \end{Bmatrix} \qquad (3.3.1)$$

When the demand volumes are uncertain, it does not make much sense to model the cargo flow decision variables as explicit unit flows since the actual demand volume is not known. This prompts for a process that makes decisions in two stages, where decisions in the first stage are made prior to the realizations of the demand values and decisions in the second stage are made after the realizations of the demand values. Applying such framework to this setting, the decisions relating to the fleet sizes and the vehicle routes are made in stage one and are assumed to be fixed for all demand realizations, whereas cargo flows decisions are made in the second stage and are dependent on the demand realizations. The Adjustable Robust formulation proposed by Ben-Tal et al. (2004) that applied the two stage framework to robust optimization provides a good starting point. Using such formulation, the cargo flow for each requirement can be formulated as a linear function of the realization of the demand values as shown in (3.3.2), where α_r and $\beta_{r,k,i,j}$ are variables.

$$C_{r,i,j} = \alpha_r + \sum_{k \in R} \tilde{d}_k \times \beta_{r,k,i,j} \qquad (3.3.2)$$

However, for the particular setting considered in this work, the formulation in (3.3.2) is not applicable since it is not possible to enforce the non-negativity constraints in (3.3.3). Note that α_r and $\beta_{r,k,i,j}$ are allowed to be negative.

$$\alpha_r + \sum_{k \in R} \tilde{d}_k \times \beta_{r,k,i,j} \geq 0 \qquad \forall \tilde{d} \in \mathrm{U}^D(\Gamma) \quad (3.3.3)$$

The formulation in (3.3.2) is only applicable when demand realizations are assumed to be taken from the uncertainty set as specified by the Γ term. However, the Γ parameter was designed to makes the robust solution less conservative by reducing the size of the original uncertainty set. Thus, if the original uncertainty set is considered to be the sample space that contains all the possible realizations, then there will always be a demand realization from the original uncertainty set that falls outside the uncertainty set for each Γ parameter. When such situations arise, there is no guarantee that the cargo flow will be positive.

Thus, a slight variation to (3.3.2) is made to represent the cargo flow as:

$$C_{r,i,j} = \tilde{d}_r \times f_{r,i,j}$$

The cargo flow through each arc for each requirement is dependent only on the requirement's demand realization. In this case, the $f_{r,i,j}$ variable represents the fraction of the demand that flows through the arc, and takes a value between 0 and 1. When the demand realizations fall outside the uncertainty set, there is a chance that the demand exceeds the allocated capacity, and a recourse action can be used to fix such infeasibility. As opposed to the stochastic programming method that allows the recourse actions to be explicitly incorporated into the formulation, the idea of having a penalty for infeasibility is not compatible with the robust optimization framework. To understand this concept better, it is important to differentiate between the original uncertainty set that is the actual sample space and the uncertainty sets for each Γ value that are subsets of the original uncertainty set. When a robust model is solved for each uncertainty set corresponding to a certain Γ value, it only considers the realizations that belong to that particular uncertainty set. This means that it is impossible to take into account of the realizations that are outside of such uncertainty set but are in the sample space. To fix this problem, a simulation will be used to evaluate the performance of the tactical plan for each Γ value under the realizations that are taken from the sample space.

Under demand uncertainty, the cargo flow is modeled as a fractional flow of the realized demand value. This is compatible with the deterministic formulation(NPM), and only a few modifications are required in the robust formulation.

The full formulation of the robust model($ROPM$) is given below:

$$minimize \sum_{a\in A} \gamma_a \times V_a + \sum_{a\in A}\sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + \underset{\tilde{d} \in U^D(\Gamma)}{maximize} \sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \tilde{d}_r \times f_{r,i,j}$$

st:

$$\sum_{(i,j)\in E} f_{r,i,j} - \sum_{(j,i)\in E} f_{r,j,i} = \begin{cases} 1 \ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \quad \forall r \in R, j \in N$$

$$\sum_{(i,j)\in E} x_{a,i,j} = \sum_{(j,i)\in E} x_{,aj,i} \quad \forall j \in N, a \in A$$

$$\sum_{r\in R} \tilde{d}_r \times f_{r,i,j} \le \sum_{a\in A} Cap_a \times x_{a,i,j} \quad \forall (i,j) \in E, \tilde{d} \in U^D(\Gamma) \qquad (3.3.4)$$

$$\sum_{(i,j)\in E} \tau_{i,j} \times x_{a,i,j} \le V_a \times T \quad \forall a \in A$$

$$f_{r,i,j} \ge 0 \qquad \forall (i,j) \in E, a \in A, r \in R$$

$$x_{a,i,j} \ge 0, x_{a,i,j} \in \mathbb{Z} \qquad \forall (i,j) \in E, a \in A$$

$$V_a \ge 0, V_a \in \mathbb{Z} \qquad \forall a \in A$$

Note that the only differences between the robust formulation and the nominal formulation are the objective function and the capacity constraints. The robust formulation can be re-expressed as a linear programming model by considering the following theorems.

<u>Theorem 3.1:</u>

Constraint (3.3.4) can be re-expressed as:

$$\sum_{r\in R}(\bar{d}_r \times f_{r,i,j}) + \Gamma \times z_{i,j} + \sum_{r\in R} p_{r,i,j} \le \sum_{a\in A} Cap_a \times x_{a,i,j} \quad \forall (i,j) \in E$$

$$z_{i,j} + p_{r,i,j} \geq \hat{d}_r \times f_{r,i,j} \qquad \forall\,(i,j) \in E, \qquad r \in \mathrm{R}$$

$$p_{r,i,j} \geq 0 \qquad \forall\,(i,j) \in E,\ \ r \in \mathrm{R}$$

$$z_{i,j} \geq 0 \qquad \forall\,(i,j) \in E, \qquad r \in \mathrm{R}$$

Proof:

Constraint (3.3.4) can be expressed as:

$$\underset{\tilde{d}\, \in\, \mathrm{U}^D(\Gamma)}{maximize} \sum_{r \in \mathrm{R}} \tilde{d}_r \times f_{r,i,j} \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \qquad \forall (i,j) \in E$$

This equals to:

$$\underset{z\, \in\, \mathrm{U}^D(\Gamma)}{maximize} \sum_{r \in \mathrm{R}} \bar{d}_r \times f_{r,i,j} + \sum_{r \in \mathrm{R}} \left(\hat{d}_r z_r \times f_{r,i,j}\right) \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \qquad \forall (i,j) \in E$$

For a given set of $(i,j) \in E$ and $f_{r,i,j}$, consider the auxiliary optimization problem:

$$maximize \sum_{r \in \mathrm{R}} \left(\hat{d}_r z_r \times f_{r,i,j}\right)$$

st:

$$\sum_{r \in \mathrm{R}} z_r \leq \Gamma$$

$$0 \leq z_r \leq 1 \quad \forall r \in \mathrm{R}$$

The dual of the formulation above can be expressed as:

$$minimize\ \ \Gamma \times z + \sum_{r \in \mathrm{R}} p_r$$

st:

$$z + p_r \geq \hat{d}_r \times f_{r,i,j} \qquad \forall r \in \mathrm{R}$$

$$p_r \geq 0 \qquad \forall r \in \mathrm{R}$$

$$z \geq 0$$

Thus, constraint (3.3.4) can be replaced with:

$$\sum_{r \in \mathrm{R}} \left(\bar{d}_r \times f_{r,i,j}\right) + \Gamma \times z_{i,j} + \sum_{r \in \mathrm{R}} p_{r,i,j} \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \qquad \forall\,(i,j) \in E$$

$$z_{i,j} + p_{r,i,j} \geq \hat{d}_r \times f_{r,i,j} \qquad \forall\,(i,j) \in E, \qquad \forall r \in \mathrm{R}$$

$$p_{r,i,j} \geq 0 \qquad \forall\,(i,j) \in E,\ \ \forall r \in \mathrm{R}$$

$$z_{i,j} \geq 0 \qquad \forall\,(i,j) \in E, \qquad \forall r \in \mathrm{R}$$

Theorem 3.2:

The objective function of $(ROPM)$ can be re-expressed as:

$$minimize \sum_{a\in A} \gamma_a \times V_a + \sum_{a\in A}\sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \bar{d}_r \times f_{r,i,j} + \Gamma \times s + \sum_{r\in R} q_r$$

st:

$$s + q_r \geq \hat{d}_r \times \sum_{(i,j)\in E} \beta_{r,i,j} \times f_{r,i,j} \qquad \forall\, r \in R$$

$$q_r \geq 0 \qquad \forall\, r \in R$$

Proof:

Consider the maximization problem in the objective function:

$$\underset{\tilde{d}\,\in\, U^D(\Gamma)}{maximize} \sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \tilde{d}_r \times f_{r,i,j}$$

This equals to:

$$\sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \bar{d}_r \times f_{r,i,j} + \underset{z\,\in\, U^D(\Gamma)}{maximize} \sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \hat{d}_r \times z_r \times f_{r,i,j}$$

Consider the auxiliary maximization problem:

$$maximize \sum_{r\in R}\left(\hat{d}_r \times z_r \times \sum_{(i,j)\in E} \beta_{r,i,j} \times f_{r,i,j} \right)$$

st:

$$\sum_{r\in R} z_r \leq \Gamma$$

$$0 \leq z_r \leq 1 \quad \forall r \in R$$

The dual can be expressed as:

$$minimize\ \Gamma \times s + \sum_{r\in R} q_r$$

st:

$$s + q_r \geq \hat{d}_r \times \sum_{(i,j)\in E} \beta_{r,i,j} \times f_{r,i,j} \qquad \forall r \in R$$

$$q_r \geq 0 \qquad \forall r \in R$$

$$s \geq 0$$

Thus, the objective function of $(ROPM)$ can be replaced with:

$$minimize \left(\sum_{a \in A} \gamma_a \times V_a + \sum_{a \in A} \sum_{(i,j) \in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r \in R} \sum_{(i,j) \in E} \beta_{r,i,j} \times \bar{d}_r \times f_{r,i,j} + \Gamma \times s + \sum_{r \in R} q_r \right)$$

st:

$$s + q_r \geq \hat{d}_r \times \sum_{(i,j) \in E} \beta_{r,i,j} \times f_{r,i,j} \qquad \forall\, r \in R$$

$$q_r \geq 0 \qquad \forall r \in R$$

As a result, $(ROPM)$ can be reformulated as an equivalent LP counterpart:

$$minimize \left(\sum_{a \in A} \gamma_a \times V_a + \sum_{a \in A} \sum_{(i,j) \in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r \in R} \sum_{(i,j) \in E} \beta_{r,i,j} \times \bar{d}_r \times f_{r,i,j} + \Gamma \times s + \sum_{r \in R} q_r \right)$$

st:

$$s + q_r \geq \hat{d}_r \times \sum_{(i,j) \in E} \beta_{r,i,j} \times f_{r,i,j} \qquad \forall\, r \in R$$

$$\sum_{(i,j) \in E} f_{r,i,j} - \sum_{(j,i) \in E} f_{r,j,i} = \begin{cases} 1 \ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \qquad \forall \mathrm{r} \in \mathrm{R}, \mathrm{j} \in \mathrm{N}$$

$$\sum_{(i,j) \in E} x_{a,i,j} = \sum_{(j,i) \in E} x_{,aj,i} \qquad \forall \mathrm{j} \in \mathrm{N}, a \in \mathrm{A}$$

$$\sum_{r \in \mathrm{R}} (\bar{d}_r \times f_{r,i,j}) + \Gamma \times z_{i,j} + \sum_{r \in \mathrm{R}} p_{r,i,j} \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \qquad \forall\, (i,j) \in E$$

$$z_{i,j} + p_{r,i,j} \geq \hat{d}_r \times f_{r,i,j} \qquad \forall\, (i,j) \in E, \qquad \forall\, r \in R$$

$$\sum_{(i,j) \in E} \tau_{i,j} \times x_{a,i,j} \leq V_a \times T \quad \forall a \in \mathrm{A}$$

$$p_{r,i,j} \geq 0 \qquad \forall\, (i,j) \in E,\ r \in R$$

$$z_{i,j} \geq 0 \quad \forall (i,j) \in E, \quad r \in R$$

$$q_r \geq 0 \quad \forall r \in R$$

$$s \geq 0$$

$$f_{r,i,j} \geq 0 \quad \forall (i,j) \in E, a \in A$$

$$x_{a,i,j} \geq 0, x_{a,i,j} \in \mathbb{Z} \quad \forall (i,j) \in E, a \in A$$

$$V_a \geq 0, V_a \in \mathbb{Z} \quad \forall a \in A$$

3.4 Execution Policies

This section introduces the various execution policies that will be used to translate a tactical plan to an executable plan after the demand values are realized. Several execution policies with two different recourse actions and three flexibility levels are considered. The first recourse policy outsources extra demands to third party companies, and a high penalty fee is incurred for each excess demand unit at each arc. The second recourse policy allows for new vehicle routes to be added given that the total vehicle usage time doesn't exceed the total vehicle available time. If adding extra vehicle routes is still not enough to accommodate the excess demands, new vehicles are rented at a high cost and additional routing decisions must be determined. Furthermore, three degrees of flexibility are considered for each of the two recourse policies. The flexibility level determines the extent to which decisions from the tactical plan can be changed at the executional level. In the least flexible system, all decisions from the tactical plan are fixed at the executional level. In the most flexible system, only the fleet size decision is fixed from the tactical plan and the vehicle and cargo routes are allowed to be re-optimized in response to the realized demand values at the executional level. The details for each flexibility level are summarized in Table 3.5. With two recourse policies and three flexibility level, there are six execution policies. These execution policies will be introduced in the following two subsections. Note that the demand values are known with certainty when decisions are made at this point in time. Let $\dddot{d}_r$ be the realized demand value for requirement r, and let $x_{a,i,j}$, $f_{r,i,j}$, and V_a be the outputs from the tactical planning model.

	Flexibility Level 1	Flexibility Level 2	Flexibility Level 3
Fleet size	Fixed	Fixed	Fixed
Vehicle Route	Fixed	Fixed	Open
Cargo Route	Fixed	Open	Open

Table 3.5: Three levels of flexibility

3.4.1 Recourse Policy 1

The recourse action under this policy is to outsource excess demand to third party companies. A high penalty cost is incurred for each unit of unmet demand. Let $pen_{i,j}$ be the penalty cost for a unit of unmet demand on arc (i,j). The three execution policies under this recourse policy are presented below.

Policy 1 Flexibility Level 1(P1F1):

Decisions on the fleet size, vehicle routes, and cargo routes are fixed. A large penalty cost is incurred if the demand realization on each cargo route exceeds the allocated capacity on the vehicle routes. The total operating cost can be calculated using the following expression:

$$z_{p1f2} = \sum_{(i,j)\in E} pen_{i,j} \times max\left[0, \left(\sum_{r\in R} \ddot{d}_r \times f_{r,i,j} - \sum_{a\in A}(Cap_a \times x_{a,i,j})\right)\right]$$

(TC_{p1f1}):

$$Total\ Cost = \sum_{a\in A} \gamma_a \times V_a + \sum_{a\in A}\sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r\in R}\sum_{(i,j)\in E} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j} + z_{p1f2}$$

Policy 1 Flexibility Level 2(P1F2):

Under this policy, the fleet sizes and vehicle routes from the planning model are fixed. However, there is an added flexibility that allows the cargo routes to be re-optimized in response to the realization of the demands. A penalty cost is incurred if the planned fleet size and the vehicle routes are not able to accommodate the demand. The execution optimization model under P1F2 policy is given by:

Input from tactical planning model:

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

V_a : Fleet size for vehicle type a

Decision Variable:

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

$s_{i,j}$: Unmet demand for arc (i,j)

(EM_{p1f2}):

$$z_{p1f2} = minimize_{f,s} \sum_{r\in R} \sum_{(i,j)\in E} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j} + \sum_{(i,j)\in E} pen_{i,j} \times s_{i,j}$$

$$\sum_{(i,j)\in E} f_{r,i,j} - \sum_{(j,i)\in E} f_{r,j,i} = \begin{cases} 1 \ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \quad \forall r \in R, j \in N$$

$$\sum_{r\in R} \ddot{d}_r \times f_{r,i,j} \leq \sum_{a\in A} (Cap_a \times x_{a,i,j}) + \ s_{i,j} \quad \forall (i,j) \in E$$

$$f_{r,i,j} \geq 0 \qquad \forall (i,j) \in E, a \in A$$

$$s_{i,j} \geq 0 \quad \forall (i,j) \in E$$

(TC_{p1f2}):

$$Total\ Cost = \sum_{a\in A} \gamma_a \times V_a + \sum_{a\in A} \sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + z_{p1f2}$$

Policy 1 Flexibility Level 3(P1F3):

Only the fleet size is assumed to be fixed, and the vehicle routes and cargo routes are allowed to be re-optimized in response to the demand realizations. A penalty cost is incurred if the planned fleet size is not able to accommodate the demand. The execution optimization model under P1F3 is presented below:

Input from tactical planning model:

V_a : Fleet size for vehicle type a

Decision Variable:

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

$s_{i,j}$ Unmet demand for arc (i,j)

(EM_{p1f3}):

$$z_{p1f3} = minimize_{f,x,s} \left(\sum_{a \in A} \sum_{(i,j) \in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r \in R} \sum_{(i,j) \in E} \beta_{r,i,j} \times \dddot{d}_r \times f_{r,i,j} + \sum_{(i,j) \in E} pen_{i,j} \times s_{i,j} \right)$$

$$\sum_{(i,j) \in E} f_{r,i,j} - \sum_{(j,i) \in E} f_{r,j,i} = \begin{cases} 1\ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \quad \forall \mathrm{r} \in \mathrm{R}, \mathrm{j} \in \mathrm{N}$$

$$\sum_{(i,j) \in E} x_{a,i,j} = \sum_{(j,i) \in E} x_{,aj,i} \qquad \forall \mathrm{j} \in \mathrm{N}, a \in \mathrm{A}$$

$$\sum_{r \in R} \dddot{d}_r \times f_{r,i,j} \leq \sum_{a \in A} (Cap_a \times x_{a,i,j}) + s_{i,j} \quad \forall (i,j) \in E$$

$$\sum_{(i,j) \in E} \tau_{i,j} \times x_{a,i,j} \leq V_a \times T \quad \forall a \in \mathrm{A}$$

$$f_{r,i,j} \geq 0 \qquad \forall (i,j) \in E, a \in \mathrm{A}$$

$$x_{a,i,j} \geq 0, x_{a,i,j} \in \mathbb{Z} \qquad \forall (i,j) \in E, a \in \mathrm{A}$$

$$s_{i,j} \geq 0 \quad \forall (i,j) \in E$$

(TC_{p1f3}):

$$Total\ Cost = \sum_{a \in A} \gamma_a \times V_a + z_{p1f3}$$

3.4.2 Recourse Policy 2

Under this recourse policy, the recourse action is to add extra vehicle routes using the existing capacity until the total vehicle usage time fills the total vehicle available time. If there are still infeasibilities, additional vehicles must be rented at a high cost to fill in the extra demand, and new routings for these extra vehicles must also be determined. Let γ_pen_a be the penalty cost for renting an additional vehicle of type a.

Policy 2 Flexibility Level 1(P2F1):

Under this policy, the planned fleet size, vehicle route, and cargo routes are assumed to be fixed. Given the assumption for this policy, the execution model under this flexibility level is given by:

Input from tactical planning model:

V_a : Fleet size for vehicle type a

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

Decision Variable:

V_r_a : Additional type a fleet needed

$x\_r_{a,i,j}$: Number of extra trips that vehicle type a travels from node i to node j

(EM_{p2f1}):

$$z_{p2f1} = minimize_{V\_r,x\_r} \sum_{a \in A} \gamma\_pen_a \times V\_r_a + \sum_{a \in A} \sum_{(i,j) \in E} \alpha_{i,j} \times x\_r_{a,i,j}$$

$$\sum_{(i,j)\in E} x\_r_{a,i,j} = \sum_{(j,i)\in E} x\_r_{a j,i} \qquad \forall \text{j} \in \text{N}, a \in \text{A}$$

$$\sum_{r\in \text{R}} \ddot{d}_r \times f_{r,i,j} \le \sum_{a\in \text{A}} Cap_a \times \left(x_{a,i,j} + x\_r_{a,i,j}\right) \qquad \forall (i,j) \in E$$

$$\sum_{(i,j)\in E} \tau_{i,j} \times \left(x_{a,i,j} + x\_r_{a,i,j}\right) \le (V_a + V\_r_a) \times T \quad \forall a \in \text{A}$$

$$x\_r_{a,i,j} \ge 0, x\_r_{a,i,j} \in \mathbb{Z} \qquad \forall (i,j) \in E, a \in \text{A}$$

(TC_{p1f1}):

$$Total\ Cost = \sum_{a\in \text{A}} \gamma_a \times V_a + \sum_{a\in \text{A}} \sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + \sum_{r\in \text{R}} \sum_{(i,j)\in E} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j} + z_{p2f1}$$

Policy 2 Flexibility Level 2(P2F2):

In this policy, flexibility is added to allow the cargo flows to be re-optimized. Given the assumption for this policy, the execution model under this flexibility level is given by:

Input from tactical planning model:

V_a : Fleet size for vehicle type a

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

Decision Variable:

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

V_r_a : Additional type a fleet needed

$x\_r_{a,i,j}$: Number of extra trips that vehicle type a travels from node i to node j

(EM_{p2f2}):

$$z_{p2f2} = minimize_{V\_r,x\_r,f} \left(\sum_{r} \sum_{(i,j)\in E} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j} + \sum_{a\in \text{A}} \gamma\_pen_a \times V\_r_a + \sum_{a\in \text{A}} \sum_{(i,j)\in E} \alpha_{i,j} \times x\_r_{a,i,j} \right)$$

$$\sum_{(i,j)\in E} f_{r,i,j} - \sum_{(j,i)\in E} f_{r,j,i} = \begin{cases} 1\ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \qquad \forall r \in R, j \in N$$

$$\sum_{(i,j)\in E} x\_r_{a,i,j} = \sum_{(j,i)\in E} x\_r_{aj,i} \qquad \forall j \in N, a \in A$$

$$\sum_{r\in R} \ddot{d}_r \times f_{r,i,j} \le \sum_{a\in A} Cap_a \times (x_{a,i,j} + x\_r_{a,i,j}) \qquad \forall (i,j) \in E$$

$$\sum_{(i,j)\in E} \tau_{i,j} \times (x_{a,i,j} + x\_r_{a,i,j}) \le (V_a + V\_r_a) \times T \quad \forall a \in A$$

$$f_{r,i,j} \ge 0 \qquad \forall (i,j) \in E, a \in A$$

$$x\_r_{a,i,j} \ge 0, x\_r_{a,i,j} \in \mathbb{Z} \qquad \forall (i,j) \in E, a \in A$$

(TC_{p2f2}):

$$Total\ Cost = \sum_{a\in A} \gamma_a \times V_a + \sum_{a\in A} \sum_{(i,j)\in E} \alpha_{i,j} \times x_{a,i,j} + z_{p1f2}$$

Policy 2 Flexibility Level 3(P2F3):

Only the fleet size from the planning model is sticky and the vehicle and cargo routes are allowed to be re-optimized. Given the assumption for this policy, the execution model under this flexibility level is given by:

<u>Input from tactical planning model:</u>

V_a : Fleet size for vehicle type a

<u>Decision Variable:</u>

$f_{r,i,j}$: Fraction of demand of requirement r that flows from node i to node j

$x_{a,i,j}$: Number of times vehicle type a travels from node i to node j

V_r_a : Additional type a fleet needed

(EM_{p2f3}):

$$z_{p2f3} = minimize_{V\_r,x,f} \sum_{a \in A} \gamma\_pen_a \times V\_r_a + \sum_{a \in A} \sum_{(i,j) \in E} \alpha_{i,j} \times x_{a,i,j}$$

$$+ \sum_{r \in R} \sum_{(i,j) \in E} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j}$$

$$\sum_{(i,j) \in E} f_{r,i,j} - \sum_{(j,i) \in E} f_{r,j,i} = \begin{cases} 1 \ , j = pod_r \\ -1, j = poe_r \\ 0, otherwise \end{cases} \quad \forall r \in R, j \in N$$

$$\sum_{(i,j) \in E} x_{a,i,j} = \sum_{(j,i) \in E} x_{aj,i} \quad \forall j \in N, a \in A$$

$$\sum_{r \in R} \ddot{d}_r \times f_{r,i,j} \leq \sum_{a \in A} Cap_a \times x_{a,i,j} \quad \forall (i,j) \in E$$

$$\sum_{(i,j) \in E} \tau_{a,i,j} \times x_{a,i,j} \leq (V_a + V\_r_a) \times T \quad \forall a \in A$$

$$f_{r,i,j} \geq 0 \quad \forall (i,j) \in E, a \in A$$

$$x_{a,i,j} \geq 0, x_{a,i,j} \in Z \quad \forall (i,j) \in E, a \in A$$

(TC_{p2f3}):

$$Total\ Cost = \sum_{a \in A} \gamma_a \times V_a + z_{p2f3}$$

3.5 Experimental Analysis

In this section, a method for analyzing the interaction and tradeoff between the robustness level of a tactical plan and the flexibility level of a system is introduced based on the tactical planning models and execution policies introduced in the previous sections. Such analysis is achieved by simulating and comparing the performances of tactical plans with different levels of robustness against the execution policies of various flexibility levels.

3.5.1 Simulation Procedure

The simulation procedure that will be used in the analysis is introduced in $(SimProc)$ and Figure 3.5.

$(SimProc)$:

The simulation procedure for a given policy k proceeds in the following steps:

1. Solve the tactical optimization model
2. Simulate demand values for each requirement
3. Given the results from the tactical planning model in step 1 and demand realizations in step 2, solve the execution model (EM_k)
4. Compute total cost by applying the equation (TC_k)
5. Repeat steps 2-4 for i iterations

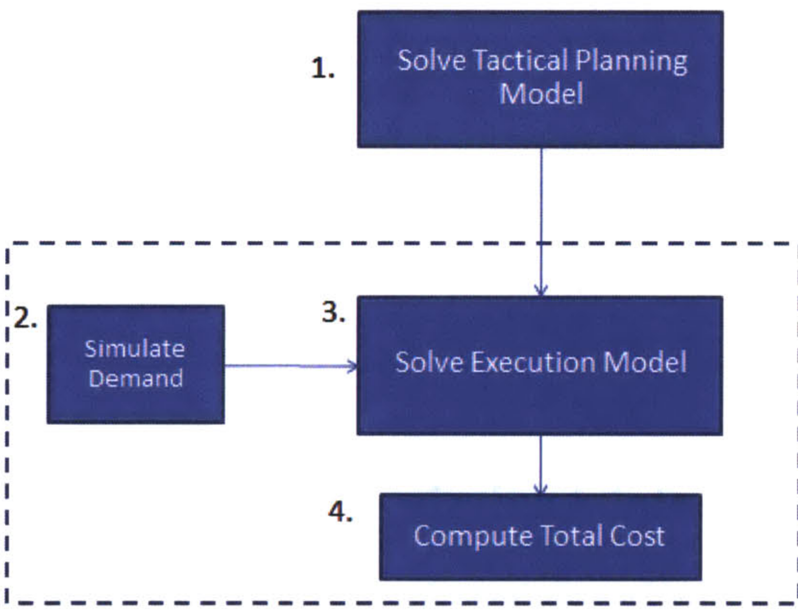

Figure 3.5: The *SimProc* Simulation Procedure

3.5.2 Experimental Design

To study the interaction and trade-off between the robustness level of a tactical plan and the flexibility level of a system, a simple hypothetical problem is considered. The network configuration of the problem is shown in Figure 3.6, where the numbers on the arcs represent the travel times in hour. The problem assumes that there is only 1 fleet type with capacity of 10 units and that there is no variability in the travel times between the nodes. In addition, there is no cargo routing cost which means that the costs for a loaded and unloaded vehicle are the same. The parameter values and the demand information are summarized in Table 3.6 and Table 3.7. For this particular problem, 5 levels of robustness in a tactical plan are considered. In robust level 0, the tactical plan is solved using the deterministic model(NPM). In robust levels 1, 2, 3, and 4, the $(ROPM)$ model is solved using Γ values of 0.25, 0.5, 0.75, and 1 respectively. The results from the optimization for the different robust levels are summarized in Table 3.8. The fleet utilization is the ratio between the total fleet usage time and the total fleet available time.

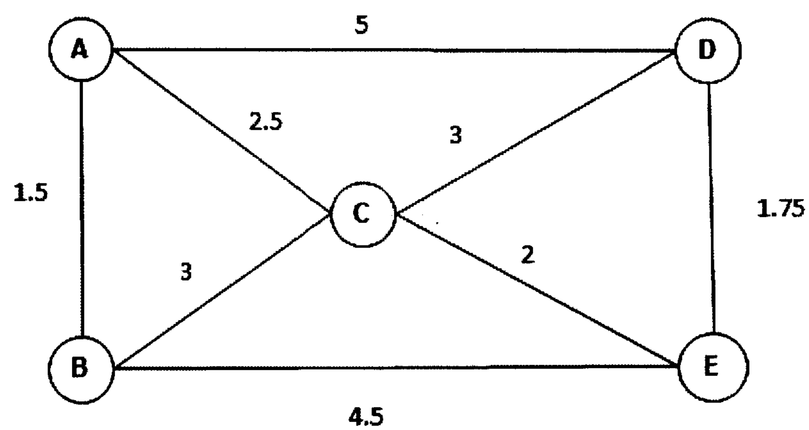

Figure 3.6: Network configuration used in the experiment

Aircraft Usage Cost	Variable Cost	Time Horizon	Fleet Capacity
$50,000	$1000 / hours	50 hours	10 units

Table 3.6: Values of parameters used in the experiment

Origin-Destination	Expected Value	Maximum Deviation
A-C	100	20
A-D	120	24
A-E	80	16
B-C	90	18
B-D	100	20
B-E	140	28

Table 3.7: Demand Requirements

Robust Level	Fleet Size	Utilization	Cost
0	8	99%	$1,217,750
1	9	94.1%	$1,290,875
2	9	96.6%	$1,313,500
3	9	97.5%	$1,324,500
4	9	99.99%	$1,344,625

Table 3.8: Optimization Results

3.5.3 Simulation Setup & Results

For the purpose of analysis, a symmetric triangular distribution around the mean value with a lower bound and an upper bound of one maximum deviation below and above the mean value is used to simulate the demand values. A penalty cost of $3,000 / excess demand unit/ hour is used for recourse policy 1 and a penalty cost of $500,000 / extra vehicle is used for recourse policy 2. For each execution policy, the simulation procedure $(SimProc)$ was executed for 1000 iterations for each five levels of robustness.

3.5.3.1 Results for Recourse Policy 1

Using $(SimProc)$, the simulation procedure for P1F1 proceeds directly from step 2 to step 4 since this execution policy does not require solving any execution model. Equation (TC_{p1f1}) is used in step 4 to calculate the total cost. Under P1F2, the optimization model (EM_{p1f2}) is used for step 3 and the equation (TC_{p1f2}) is used to calculate the total cost for step 4. Finally, the optimization model (EM_{p1f3}) is used for step 3 and the equation (TC_{p1f3}) is used to calculate the total cost for step 4 for P1F3. Note that the tactical plans for robust levels 1, 2, 3, and 4 have the same fleet size so they share the same results under flexibility level 3.

The cumulative distribution functions of the total costs for recourse policy 1 with 3 different levels of flexibility and 5 levels of robustness are presented in Figure 3.7, Figure 3.8, and Figure 3.9, and several statistics from the simulation are presented in Table 3.9, Table 3.10, and Figure 3.10.

For each flexibility level, the cumulative distribution function of the total cost for each of the robust level intersects each other. This suggests that there is no definite answer as to which level of robustness is better than the other, and the choice depends on the risk preference of the decision maker. For instance, suppose that the system is operating at flexibility level 1, a more risk averse decision maker may prefer the tactical plan with robust level 4 since this plan is the most conservative option with the least variance in the final operating cost. The probability of having a total cost greater than $1,344,625 is less than 10%. On the other hand, a risk loving decision maker may prefer the riskier plan with robust level 0 that has a chance of yielding a total cost of as low as $1,217,750 but also a certain chance of yielding a total cost of as high as $1.7 million. Under this option, there is approximately 55% chance that the cost will be below $1,344,625 but 30% chance that the cost will go beyond $1.4 million. On the other hand, there is 0% of going below $1,344,625 for the tactical plan with robust level 4 but less than 2% of going beyond $1.4 million. Thus, there is a trade-off between risk and cost that the decision maker must make.

Given a certain flexibility level, the probability that a penalty is incurred decreases as the robust level increases. Such result is expected since a plan with more robustness is likely to be less sensitive to fluctuations in demand values. However, it may be counterintuitive that the probability that a penalty is incurred in a system with flexibility level 3 for tactical plans with robust level 2, 3, and 4 is higher compared to those of the systems with lower flexibility levels. The reason for this anomaly is that at the flexibility level 3 where only the fleet size is fixed, the execution optimization model re-optimizes the vehicle routing and cargo routing decisions to minimize the total cost. In certain demand realizations, it may be more economical to outsource certain portion of demand to third party companies than adding additional vehicles route to accommodate such demand. Adding a vehicle trip to a certain arc can be costly because it generally translates to adding more than one trip to the system since the conservation of flow for the vehicles has to be met.

Although the most optimal robust level of a tactical plan for each flexibility level depends on the risk preference of the decision maker, the expected total cost may be used as an evaluating criterion to determine the robust level that performs the best on average. From Table 3.10, a plan with robust level 3 has the lowest expected cost under flexibility level 1, a plan with robust level 2 has the lowest expected cost under flexibility level 2, and plans with robust level 1, 2, 3, and 4 have the lowest expected cost under flexibility level 3. This shows that on average, as the flexibility level of the system increases, the optimal robustness level of the tactical plan decreases. In addition, it is interesting to note that the expected cost for the deterministic tactical plan under a system with flexibility level 3 is lower than the expected cost for all the robust plans under the system with flexibility levels 1 and 2. This means that as the system gets more flexible, a deterministic plan could perform better than a robust plan under a system with lower flexibility level.

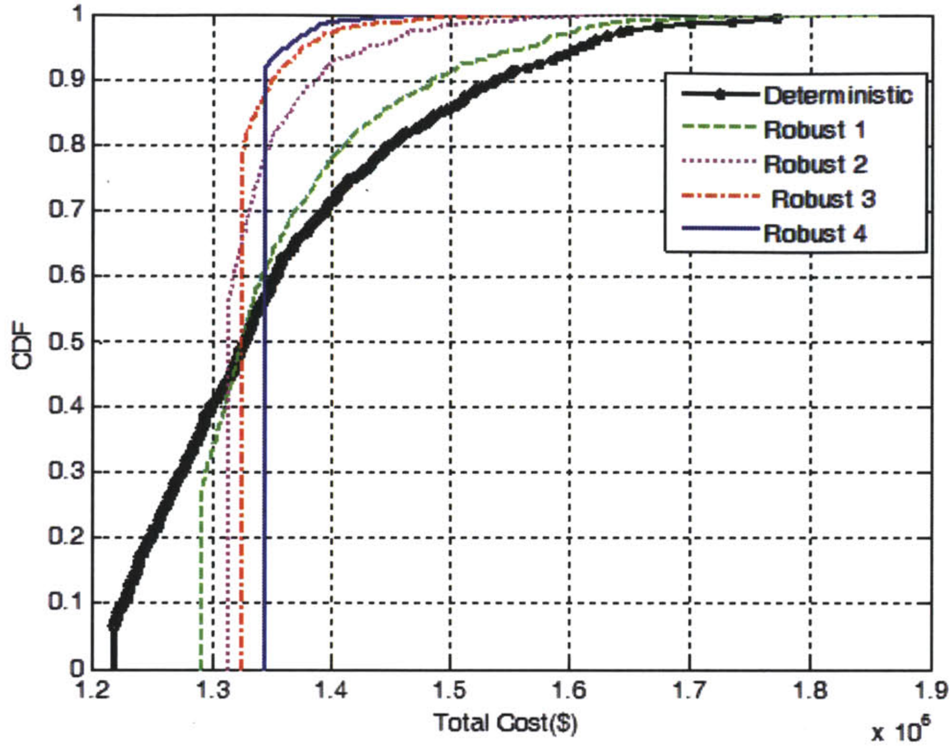

Figure 3.7: CDF of the total cost of the different plans under the P1F1 policy

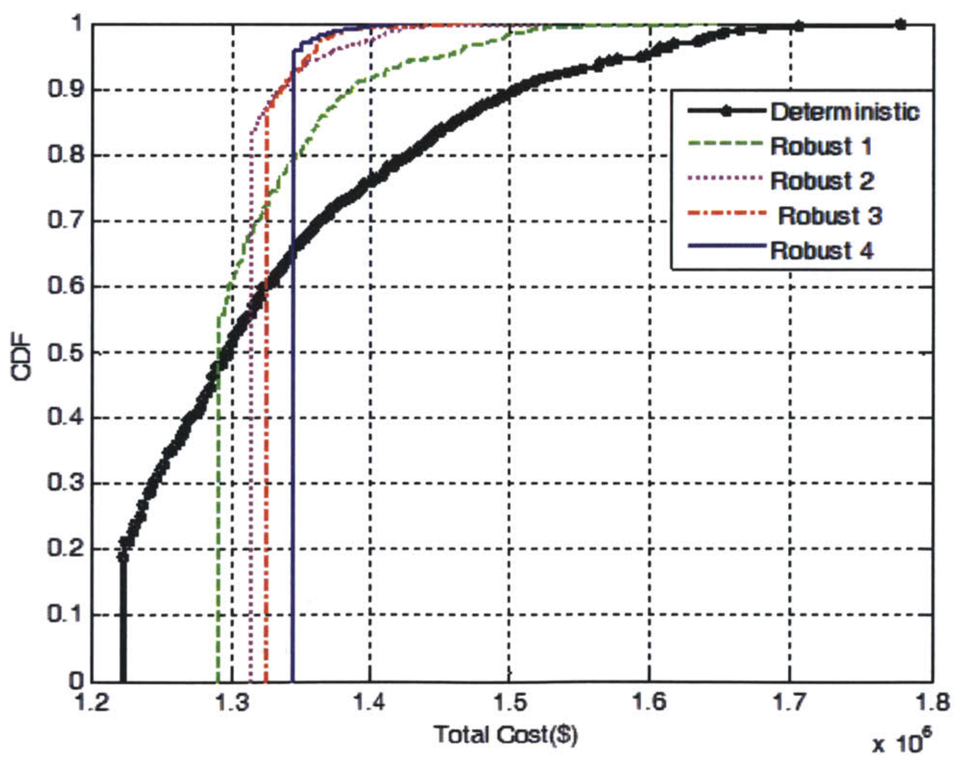

Figure 3.8: CDF of the total cost of the different plans under the P1F2 policy

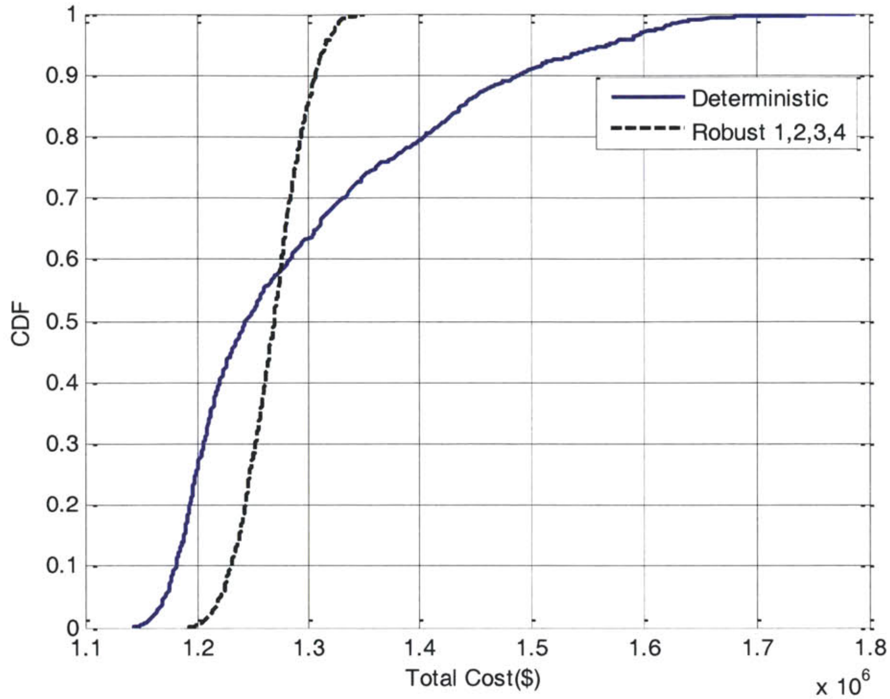

Figure 3.9: CDF of the total cost of the different plans under the P1F3 policy

Robust Level	Fleet Size	P1F1 Policy	P1F2 Policy	P1F3 Policy
0	8	93.8%	81.8%	69.1%
1	9	72.5%	44.2%	21.9%
2	9	44.5%	16.1%	21.9%
3	9	21.6%	12.8%	21.9%
4	9	8%	4.2%	21.9%

Table 3.9: Probability that a penalty is incurred for Recourse Policy 1

Robust Level	Fleet Size	P1F1 Policy	P1F2 Policy	P1F3 Policy
0	8	1,357,300 (122,580)	1,331,700 (117,760)	1,295,600 (126,260)
1	9	1,359,200 (88,826)	1,321,040 (54,437)	1,269,100 (28,857)
2	9	1,336,400 (45,922)	1,320,700 (23,547)	1,269,100 (28,857)
3	9	1,332,800 (25,479)	1,328,490 (14,441)	1,269,100 (28,857)
4	9	1,347,300 (12,292)	1,345,700 (6,791)	1,269,100 (28,857)

Table 3.10: Expected value and standard deviation of the total cost for Recourse Policy 1

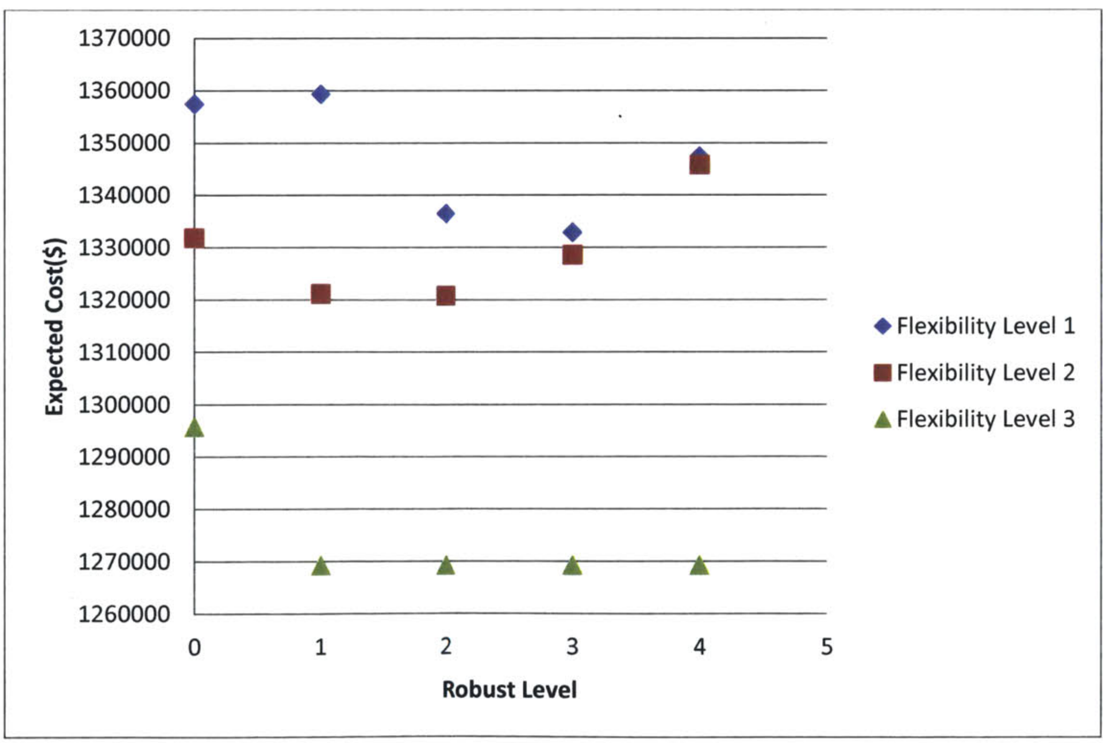

Figure 3.10: Expected cost for Recourse Policy 1 under different levels of robustness and flexibility

3.5.3.2 Results for Recourse Policy 2

Under P2F1, the optimization model (EM_{p2f1}) is used for step 3 and the equation (TC_{p2f1}) is used to calculate the total cost for step 4. Under P2F2, the optimization model (EM_{p2f2}) is used for step 3 and the equation (TC_{p2f2}) is used to calculate the total cost for step 4. Finally, the optimization model (EM_{p2f3}) is used for step 3 and the equation (TC_{p2f3}) is used to calculate the total cost for step 4 under P2F3.

Similar to the results of the recourse policy 1, the cumulative distribution functions for the different robustness level under each flexibility level intersects each other. However, the deterministic tactical plan performs substantially worse under this recourse policy. Although the decision maker may have a hard time determining the most optimal robust level for the tactical plan in each flexibility level, it is clear from the CDF plot that the deterministic tactical plan performs much worse than the other tactical plans with higher degree of robustness. This is especially true for the system with flexibility levels 1 and 2, where the probability for going over $1.72 million is approximately 90% and 80% respectively for the deterministic tactical plan; the probability for going over such figure for the other robust tactical plan is less than 10%. One possible explanation for the poor performance of the deterministic tactical plan is that the penalty cost under this recourse policy is relatively high, so a plan that is not protected against fluctuation in demand values is likely to be charged with a high penalty cost.

One interesting result from the simulation that is worth mentioning is that since the vehicle utilization rate for the tactical plan with robust level 4 is almost 100%, it is not possible to add extra vehicle routes using the existing capacity to accommodate excess demand. As a result, a new vehicle is always required when there is not enough capacity to meet demand, and there is always a high jump in cost even if the excess unmet demand volume is very marginal. Thus, the total cost for the tactical plan with robust level 4 is particularly high for the cases when the allocated capacity is not enough to accommodate the excess demand.

In addition, using the expected total cost as the evaluating criterion, it is interesting to note that flexibility doesn't have much effect on the performance of each of the tactical plan under this recourse policy. As opposed to the previous recourse policy where there is a large drop in the expected total cost as the flexibility level of the system increases to 3, the drop in the expected cost under this recourse policy is very minimal. There is also a very little drop in the expected total cost as the flexibility level of the system increases from 1 to 2 for each of the tactical plan with robust levels 1, 2, 3, and 4.

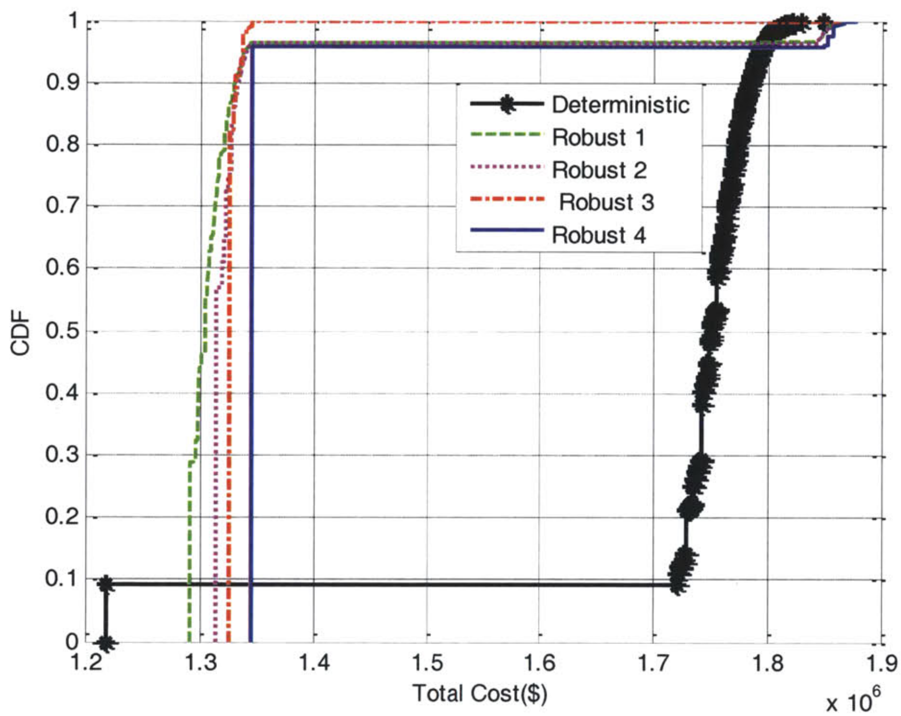

Figure 3.11: CDF of the total cost of the different plans under the P2F1 policy

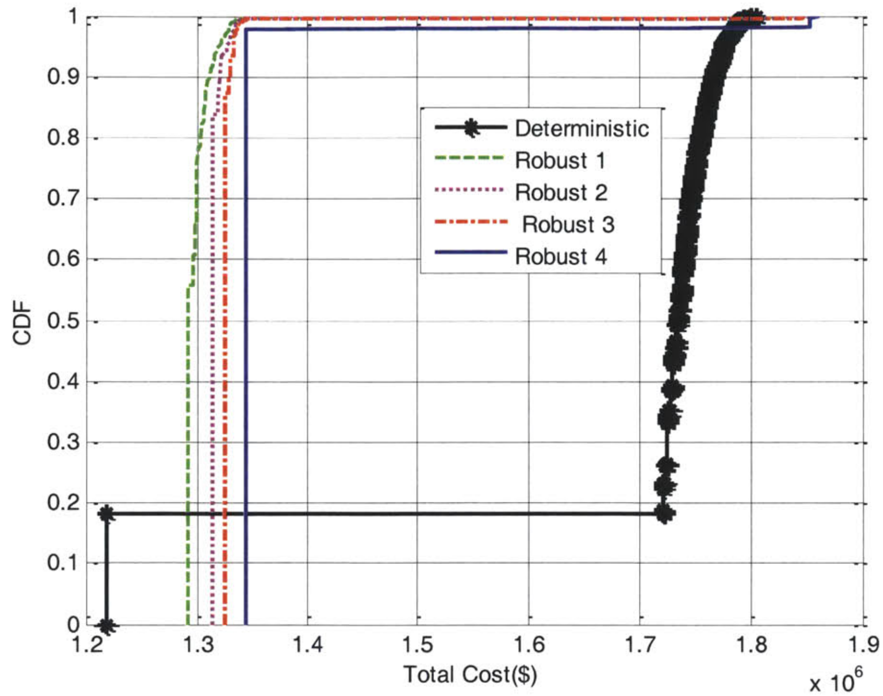

Figure 3.12: CDF of the total cost for the different plans under the P2F2 policy

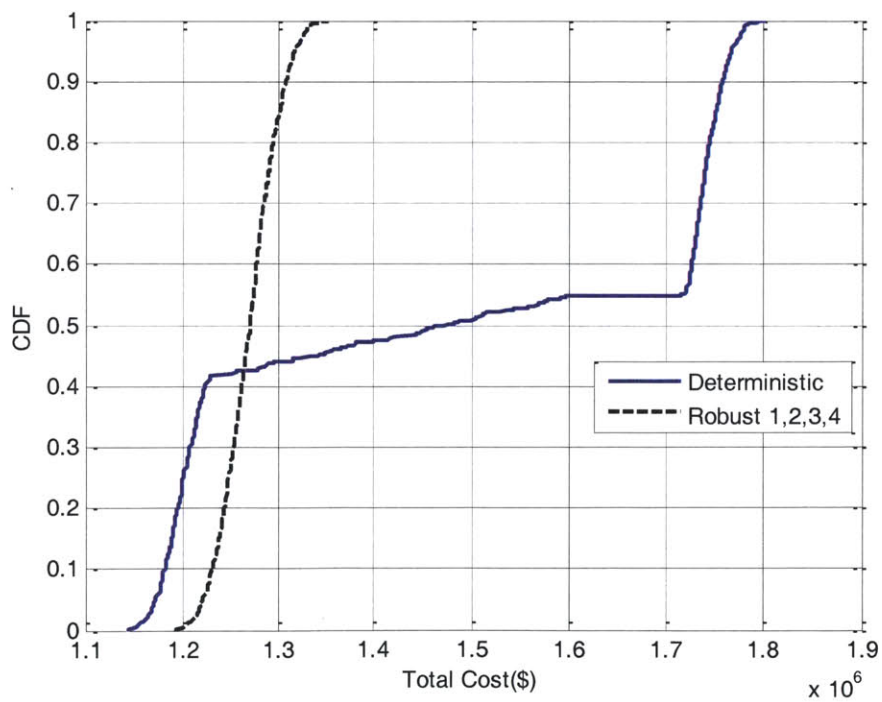

Figure 3.13: CDF of the total cost for the different plans under the P2F3 policy

Robust Level	Fleet Size	P2F1 Policy	P2F2 Policy	P2F3 Policy
0	8	90.8%	81.8%	45.1%
1	9	71.3%	44.2%	0%
2	9	43%	16.1%	0%
3	9	17.7%	12.8%	0%
4	9	7.4%	3%	0%

Table 3.11: Probability that a penalty is incurred for recourse policy 2

Robust Level	Fleet Size	P2F1 Policy	P2F2 Policy	P2F3 Policy
0	8	1,706,900 (157,000)	1,647,700 (203,520)	1,473,500 (260,200)
1	9	1,324,799 (102,750)	1,298,333 (26,478)	1,269,400 (28,887)
2	9	1,337,800 (99,663)	1,316,800 (295,740)	1,269,400 (28,887)
3	9	1,327,100 (14,441)	1,320,500 (28,899)	1,269,400 (28,887)
4	9	1,366,100 (102,790)	1,355,300 (73,003)	1,269,400 (28,887)

Table 3.12: Expected value and standard deviation of the total cost for recourse policy 2

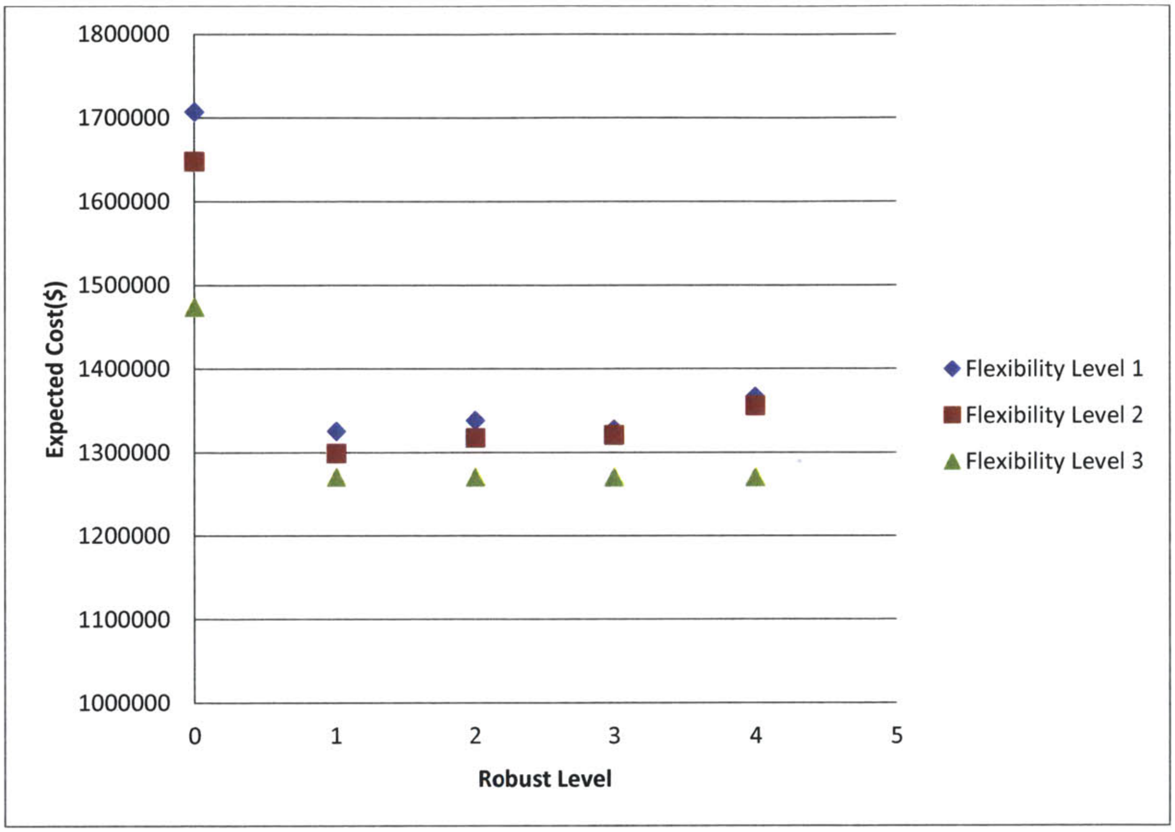

Figure 3.14: Expected total cost for plans with different degree of robustness and execution policies with different levels of flexibility under recourse policy 2

3.5.4 Summary

This section provided several analyses on how the degree of robustness of a tactical plan interacts with flexibility level of an execution policy, and how such interaction affects the performance of the system. A method for addressing the trade-off between the degree of robustness and cost for increasing robustness was first introduced. The tactical plans and execution policies were simulated, and the two major statistical results of interest were the cumulative distribution function and the expected value of the total cost. For each execution policies, as long as the cumulative distribution functions of the total cost for each tactical plan intersect, the appropriate level of robustness to select for the tactical plan depends on the risk preference of the decision maker. A risk-averse decision maker would prefer a tactical plan with more robustness than a risk-loving decision maker. However, the expected value of the total cost may be used as metric for selecting the appropriate level of robustness on average for each execution policies. Using the expected value of the total cost as the evaluating metric, the trade-off between the degree of robustness of a tactical plan and the flexibility level of the system can also be addressed. The results suggested that as the flexibility level of the system increases, the optimal level of robustness required from a tactical plan decreases. Furthermore, the deterministic tactical plans that operate under high flexibility level could perform better than tactical plans with added robustness but operate under low flexibility level.

In general, the appropriate level of robustness in a tactical plan for each execution policies depends primarily on two factors: the cost for adding robustness and the penalty cost. Accordingly, the challenge of finding the trade-off between total cost and risk boils down to finding a balance between the cost for adding robustness and the penalty cost. The cost for adding robustness is influenced by the demand pattern and the cost parameters that include the fixed aircraft usage cost, aircraft routing cost, and cargo routing cost. The penalty cost is influenced by the demand pattern - particularly the demand distribution - and the cost for fixing any infeasibility. To have a better understand of how these two factors, including the cost for adding robustness and the penalty cost, affect the optimal robustness level in a tactical plan, consider the problem setting described in section 3.4.1 and assume that the recourse action is to outsource excess demand to third party companies. Holding all the parameters fixed, consider two scenarios: the penalty cost increases and the aircraft usage cost increases, which consequently increases the cost for adding robustness. In the first scenario, as the penalty cost increases, tactical plans with lower robustness and consequently higher probability of infeasibility would perform poorer as the cost for fixing infeasibilities increases. According, the slope of the non-vertical line in the cumulative distribution function for each of the tactical plan would be less steep, and the optimal level of robustness as measured by the total expected cost is likely to increase. On the other hand, as the cost for adding robustness increases, the cumulative distribution function of the total cost would have the same shape but would be shifted to the right. Accordingly, it is likely that the optimal level of robustness under such scenario would decrease.

Chapter 4
Military Logistics System

This chapter provides background information on the US military logistics system with an emphasis on the operations within the Air Mobility Command.

4.1 United States Transportation Command

The United States Transportation Command (USTRANSCOM) is primarily responsible for managing civilian and military transportation resources to provide transportation services to the Department of Defense in both time of peace and time of war. Demand for services is driven by dynamic global events, and may arise from the need to sustain ongoing military operations, the need for humanitarian support in areas affected by natural disasters, or the need to transport important personnel. These events generate requirements that must be accomplished in a responsive and efficient manner. Providing transportation services at a global scale requires an effective management of different modes of transportation resources including land, sea, and air. Accordingly, the USTRANSCOM is comprised of 3 component commands that include the Air Mobility Command (AMC), Military Sealift Command (MSC), and the Military Surface Deployment & Distribution Command (SDDC). The focus of this thesis is on the airlift planning problem faced by the AMC.

4.1.1 Air Mobility Command

The AMC primarily serve six mission areas including the Channel, Contingency, Special Assignment Airlift Mission (SAAM), Exercises, Training, and Joint Airborne/Air Transportability Training (JA/ATTs). The purpose of each mission area falls into one of the two categories: mobility and training purposes (Koepke, 2004). Exercises, training, and the JA/ATTs missions utilize aircraft primarily for military training purposes, whereas Channel, Contingency, and SAAM missions are for mobilizing military equipment and/or personnel. Planning for mobility missions is the primary focus of this thesis. Channel missions deal with the routine movement of supplies and goods to replenish and sustain various ongoing military operations around the globe. These missions are operated continuously on established routes without delivery dates. Contingency missions deal with the discrete movements of customer requirements between specific locations during specific time windows. Finally, SAAMs are high priority missions that deal with chartering aircraft to military organizations for their exclusive use.

AMC utilizes both organic (military owned) and civilian aircraft to fulfill the requested missions. Organic aircraft and aircrews are organized into aircraft *wing*, which is composed of one or several aircraft squadrons. Each aircraft squadron is located at the aerial port and acts as a home base for a single aircraft type and the associated aircrew. Aerial ports are the nodes in the AMC's network system. They are locations where cargo get processed, loaded/unloaded onto aircraft, and where aircraft take off and land, and may serve as a home base for certain aircraft types. One month prior to the execution month, each wing allocates a certain number of "contract" aircraft and aircrews for mobility purposes and reserves the remaining "fenced" resources for internal use. Each aircraft is characterized by the cargo capacity, maximum flying range, cruising speed, cargo type accommodation, and the minimum ground time (minimum turn-time between missions).

The Commercial Civil Reserve Air Fleet (CRAF) program was designed to facilitate times when there are insufficient military resources to perform all the necessary air transportation requests (Nielsen et al., 2004). Commercial airlines that are participants of the CRAF must commit a certain portion of their resources to military use when the CRAF program is activated during wartime. The participants are not compensated unless the CRAF program is activated but they receive bid points in proportion to the number of committed resource. These bid points can be used to bid for peacetime contracts.

4.2 Planning Process in Air Mobility Command

The unit within the AMC that is responsible for the scheduling and execution of air operations is the Tanker Airlift Control Center (TACC). Requirements of different mission types enter the TACC and are distributed to the corresponding planning units. Within the TACC, the Global Channel Directorate is responsible for the channel missions, the Global Readiness Directorate is responsible for the exercise and contingency missions, and the Current Operations Directorate is responsible for the SAAM missions. The different mission types follow different planning processes. As an example, this section describes the process for contingencies since these represent a large proportion of all missions. The planning cycle for contingency missions spans approximately 6 months and begins at the semi-annual force flow conference where Combatant Commanders (CC) and the USTRANSCOM planners meet to discuss the baseline set of transportation requirements that are expected to emerge and be executed over the course of the next few months. The conference gives the CCs an opportunity to receive feedback on their initial rough drafts of requirements from the USTRANSCOM planners. By the end of the conference, the CC's requirements are refined with respect to the time frame and travel mode to match with the projected capacity and to increase transportation feasibility.

After the force flow conference, contingency requirements are continually refined and adjusted as specifications to old requirements may change and new requirements may enter into the system. Airlift requirements are typically validated by the CCs 21 days before execution. Once requirements are validated (V'ed), the USTRANSCOM must determine and confirm each requirement's mode of transportation, transport asset type, transport asset source (military or civilian carrier), and the projected time frame. Once an airlift requirement is determined to be transportation feasible (T'ed), the USTRANSCOM is committed to transporting the requirement and sends it to the TACC for detailed planning.

The TACC office is responsible for generating a "mission" that specifies the relevant information relating to the movements of each requested requirement. Each planned mission is a schedule that specifies the mission start and end time, priority, aircraft type, wing, and route. The mission planner must ensure that the total mission time, or the total between the Available to Load Date, the earliest date that the mission can start, and the Latest Arrival Date, the latest date that the mission should end, is greater than the total flight time and ground time. The flight time includes all the air time required to make all the requested stops at each airport. The ground time includes all the connection time between successive flights, where various activities such as loading/unloading cargo, refueling, crew rest, and crew change are performed.

The generated missions from the different units within the TACC are filtered by aircraft types and are sent to the corresponding Barrel office which tracks the availability of the "contract" aircraft and aircrews that are allocated at each wing for the TACC to utilize. Based on the mission constraints and the resource availability, the Barrel decides which mission to accept and which wing to operate the accepted missions. Once an aircraft is allocated, the mission planner completes the itinerary with the allocated tail number as well as adding positioning and de-positioning legs to the itinerary. Finally, the Command and Control Directorate or the "Floor" is responsible for executing all the missions.

4.3 Literature Review

This section provides a summary of some of the works that have been done in the literature on airlift planning within the AMC. Smith et al. (2004) proposed a constraint-based scheduling model, the "AMC Allocator", that combines constraint programming and heuristic search to incrementally assign new missions to wings. The model is designed to operate under a continuous and dynamic environment, thereby providing 3 major capabilities: accommodating new high priority missions while minimizing disruptions to previously schedule missions, allowing for easy reconfiguration to accommodate several constraint relaxations to increase the number of supportable missions, and combining missions to minimize non-productive flying time. Furthermore, the AMC Allocator may be utilized in various levels of automated fashion, ranging from a fully manual mode, where the system is simply used for bookkeeping, to semiautomatic mode, where the system is used to explore options, to a fully automatic mode, where the system is used to determine the allocation decisions.

The primary task of the AMC Allocator is to determine a set of new assignments - defined by the mission, wing, start-time, and end time – that minimizes disruption to the existing schedule and maximizes the sum of priorities of the newly assigned missions. To achieve such goal, each new mission is ranked based on the priority and delivery date, and is independently assigned to a wing. For each mission, the assignment procedure proceeds in several sequential processes. A set of candidate wings are first generated based on the specification of the mission. A search operator then generates a set of feasible allocation interval for each candidate wing. Finally, an evaluating module evaluates the different allocation and selects the allocation that minimizes/maximizes a certain metric. Due to the "oversubscribed" nature of the allocation problem, it may not always be possible to find a feasible assignment for a given mission. If no feasible assignment can be made for a mission, the mission is considered un-assignable. In such case, the model allows for the decision makers to explore various constraint relaxation options by configuring the search operator and the evaluation criteria. Examples of such constraint relaxations are mission delays, resource over allocations, alternative aircraft types, and mission combination.

Baker et al. (2012) addressed the airlift allocation problem under demand uncertainty using two stage stochastic programming. The aim of the model is to assign aircraft types to the three major mission areas - Channel, Contingency, and SAAM - over a planning horizon of 30 days. Over the planning horizon, the aircraft types are only allowed to serve requirements from the mission area that they are assigned to. The three major aircraft classes that may be allocated are military aircraft, civilian aircraft on long term lease, and short-notice rental civilian aircraft. The model is formulated using a two stage stochastic programming framework where the decisions on the allocation of missions to military aircraft and civilian aircraft on long term lease are fixed in the first stage, and are influenced by the recourse decisions that are made in response to the different demand realizations in the second stage.

The objective function of the model is to minimize the sum of the cost of allocating commercial aircraft in the first stage and the expected cost that results from the decisions made in the second stage. The constraints in the first stage formulation are the aircraft capacity constraints that limit the number of aircraft of each type that can be allocated at each wing. Such capacity depends on the number of "contract" military aircraft at each wing and the number of contracted civilian aircraft. Given the aircraft allocation decisions from the first stage formulation as an input, the second stage formulation determines a detailed routing of cargo and aircraft subject to a wide variety of operational constraints for each individual scenario that represents a realization of the demand values. The second stage formulation allows for an additional allocation of high cost short notice civilian aircraft to accommodate the excess demand. For each scenario, the formulation minimizes the sum of commercial short term rental cost, late and non-delivery penalties, and the cost of aircraft operations. The set of constraints in the second stage formulation can be categorized into two main classes: mission constraint and aircraft usage constraint. The mission constraints include the demand fulfillment constraint, cargo transshipment flow constraints, and capacity constraints. The aircraft usage constraints contain the mission time constraint, flying time constraint, and aerial fueling constraint.

Over a hundred random second stage scenarios were generated, and the model was solved using Bender's decomposition that is facilitated by the computational power of parallel computing. The stochastic solution outperformed the deterministic solution in all the generated scenarios, but only by a small margin. On average, the stochastic solution performed 3% better than the deterministic solution.

Wu et al. (2009) considered a solution strategy based on a dynamic programming framework for solving the airlift problem. Under such framework, decisions are made sequentially at uniform time intervals over a finite horizon based on a policy. Different policies are defined by the nature of information that is used, and the primary goal of the study is to determine the policy that optimizes a certain metric over the time horizon. Four classes of information were considered in this paper, including the physical state, forecast of exogenous information, forecast of impact of decisions on the future, and expert knowledge. The physical state represents the resources and the demands that are in the system at each time step. The evolution of the state through time is defined by the transition function and is dependent on the state, decision, and exogenous information at each time step. The exogenous information represents the changes in the system due to unforeseen events such as equipment breakdowns, flight delays, or arrival of new requirements that occurs in the period between the previous time step and the current time step.

The Rule-Based policy uses only the physical state of the system to make decisions. Similar to the approach introduced by Smith et al. (2004), this policy matches the first available requirement to the earliest available aircraft and performs an assignment feasibility check. If the assignment is feasible, the requirement gets assigned to the aircraft, otherwise, the next earliest available aircraft is considered. This process repeats until there is a feasible assignment, and this procedure is repeated until all the requirements are assigned. The other class of policy that uses only the physical state to make decision is the Myopic policy. Two variations of Myopic policies were discussed. The first Myopic policy considers one requirement and a list of potential candidate aircraft. The assignment that optimizes the evaluating function is selected. The second Myopic policy matches a list of requirements to a list of aircraft. This translates to an integer optimization problem, where the aim is to determine the set of assignments that maximizes a certain metric based on the evaluating function.

The Rolling Horizon policy incorporates forecast information on the requirements that are expected to materialize and the resources that are projected to be available throughout the planning horizon into the decision making process. For practical reasons, the forecast information is represented as deterministic point forecasts. Accordingly, the Rolling Horizon policy translates to a deterministic optimization problem that considers all the decisions that are expected to be made at each decision epoch from the current time step till the end of the planning horizon based on both the known information and the forecasted information.

The policy that utilizes the forecast of impact of decisions on the future is the Approximate Dynamic Programming policy. Such policy introduces stochasticity into the decision making process by solving the Bellman's equation using an approximating function that captures the contribution that a certain decision has on the current time step and the possible consequences that might occur in the future based on this decision. Finally, the Expert knowledge policy seeks to embed certain good behaviors based on expert knowledge into the solution. This is done by constructing a function that measures the difference between the behavior of the decision made by the model and the behavior that is expected from the expert. By associating a monetary value to this function and adding it to the objective function, the decision maker is able to control the degree to which the model should behave based on the expert knowledge.

An experiment was conducted to evaluate the performance for each of the policies. The problem was set up to manage 6 aircraft types to fulfill a set of requirements that are approximately worth 4 times the available capacity over a horizon of 50 days with 4 hour time interval. The experiment showed that as the number of information classes used in the policy increases, the performance with respect to the total operating cost and throughput rate also increases.

This page intentionally left blank

Chapter 5

Analysis of Robustness and Flexibility in a Military Logistics System

5.1 Overview

The purpose of this chapter is twofold. Firstly, it considers a centralized optimization based approach to the scheduling and airlift allocation process within the AMC. Currently, there is very little visibility between the different planning units within the TACC. The schedulers in each planning unit constructs a schedule for their mission area with no knowledge of the requirements that are in the other planning units, and the Barrel unit only has knowledge of the available capacity at the wing that they are responsible for. Furthermore, aircraft allocation decisions are made in a first-come-first-served policy without using any form of optimization. This chapter considers a planning process that allows a centralized decision maker that has total visibility of all the requirements from every mission area and the availability of the aircraft at all the wings to be responsible for the scheduling and aircraft allocation processes. By aggregating requirements from all the mission areas and considering the capacity of all the aircraft across all the wings, scheduling and aircraft allocation decisions can be determined simultaneously, and may lead to a better solution.

Secondly, this chapter aims to analyze how the interaction between the degree of robustness of a tactical plan and the flexibility level of an execution policy affects the performance of a military airlift system. Similar to what has been presented in chapter 3, the metric for measuring robustness of a plan is the probability of feasibility, and the metric for measuring the performance of the system is the final operating cost. The analysis will be done by first considering a solution approach that assumes no flexibility within the system. Under such setting, a plan is made to take into account of potential future changes prior to the operation, and is executed for a certain time horizon without any adjustments.

The major sources of uncertainty in the system includes aircraft delays due to adverse weather condition and mechanical breakdowns, changes to old requirements, and arrival of new requirements. This chapter limits the scope of the analysis to just demand uncertainty due to new requirements that may materialize during the planning horizon. Such continuous and dynamic nature of demand arrivals is driven by global events. Accordingly, predicting the arrival of new requirement is an extremely difficult task. This motivates the use of the robust optimization framework that does not require any probabilistic assumption or knowledge of the demand uncertainty. The information that is required to produce a robust plan is a point forecast and the upper bound value of the demand volume. Such information will be used to construct an uncertainty set that is motivated by Bertsimas and Sim (2002). Consequently, the uncertainty set will be used by the robust formulation to systematically add cushion to the plan so that it is robust against demand fluctuations. Then by varying the flexibility level of the system, this chapter aims to determine how the optimal level of robustness of the plan and the corresponding performance of the system change.

Due to the scale and complexity of the military logistics system, this chapter will be performing the analysis on small problem instances under several simplifying assumptions that will be described in Section 5.2. In the same section, a deterministic tactical plan that resembles an integrated schedule design and fleet assignment model is introduced. Section 5.3 presents a robust optimization model based on an uncertainty set motivated by the concept of budget of uncertainty (Bertsimas & Sim, 2002) to capture the "all or nothing" nature of the demand. The execution policies that will be used to translate a tactical plan into an executable plan are presented in Section 5.4. Finally, the analysis of the interaction between the flexibility of the system and the robustness of the plan is presented in Section 5.5.

5.2 Deterministic Formulation

Given a network structure defined by (V, E) where V is the set of ports and E is the set of arcs, the centralized planner is responsible for managing a fleet of aircraft to transport a set of requirements from several origins to several destinations subject to various capacity, spatial, and temporal constraints over a certain time horizon, which is discretized into uniform time periods.

Each requirement is characterized by its cargo volume, port of embarkation(POE), port of debarkation (POD), Available to load date (ALD), Earliest Arrival Date (EAD), and Latest Arrival Date (LAD). A requirement is ready to be delivered at the POE after the ALD and must be delivered to the POD within the EAD and LAD time window. This study assumes that no late deliveries are allowed. Aircraft are located at their corresponding home base (i.e. wing), and are available for a certain time frame. Each aircraft is available to operate from the wing starting at time st and must be back at the same wing by time ed. Aircraft of different types are assumed to have the same flying time between ports $(\tau_{i,j})$, but differ in the capacity (cap). Furthermore, flight times are assumed to be known with certainty and ground time operations are assumed to be instantaneous. In the actual context, the fleet sizes are specified by the wings and are beyond the control of the TACC planners. However, to make the analysis more interesting, this study assumes that fleet size decisions are also determined by the model and there is an unlimited number of aircraft for each aircraft type.

The aim of the model is to determine the fleet size at each wing, the cargo routes for each requirement, and the aircraft routes. To limit the size of the problem, aircraft flows are considered at the aircraft type level, and cargos are assigned to arcs in the network but not assigned to any specific aircraft types.

Set:

V	:	Set of all ports
E	:	Set of all arcs
A	:	Set of all aircraft types
R	:	Set of all requirements

Data:

d_r	:	Demand volume for requirement r
poe_r	:	Port of embarkation for requirement r
pod_r	:	Port of debarkation for requirement r
ald_r	:	Available to load date for requirement r
ead_r	:	Earliest arrival date for requirement r

lad_r	:	Latest arrival date for requirement r
w_a	:	Home wing for aircraft type a
st_a	:	Earliest time period that aircraft type a is available for use
ed_a	:	Latest time period that aircraft type a must be back at the wing
cap_a	:	Capacity for aircraft type a
$\tau_{i,j}$	:	Expected flight time between port i and j
$PortCap_i$	:	number of aircraft that port i can accommodate per time period
C_a	:	Fixed cost for using 1 unit of aircraft type a for the planning horizon
$\gamma_{a,i,j}$	:	Cost for aircraft type a to fly from port i to port j
$\beta_{r,i,j}$	:	Cost for moving a unit of requirement r from port i to port j
T		Length of planning horizon

Decision Variables:

$f_{r,i,j,t}$	:	Fraction of demand from requirement r leaving port i for port j at time t
$I_{r,i,t}$	:	Cargo inventory level for requirement r at port i at time t
$x_{a,i,j,t}$	:	Number of aircraft type a leaving port i for port j at time t
$s_{a,i,t}$	:	Aircraft type a inventory level at port i at time t
y_a	:	Number of aircraft type a allocated for the planning horizon

The deterministic formulation $(ALNM)$ is presented below:

Objective Function:

$$minimize \sum_{a\in A} C_a \times y_a + \sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times d_r \times f_{r,i,j,t}$$

The objective function minimizes the sum of the fixed aircraft usage cost, aircraft routing cost, and cargo flow cost.

Cargo Flow Constraint:

$$1 = I_{r,poe_r,ald_r} + \sum_{(poe_r,j)\in E}\sum_{a\in A} f_{r,poe_r,j,ald_r} \quad \forall r \in R$$

The cargo for each requirement r becomes available at time ald_r at port poe_r. At such time unit, the cargo may either be stored at the port or be sent off to another port.

$$I_{r,j,t-1} + \sum_{(i,j)\in E} f_{r,i,j,(t-\tau_{i,j,a})} = I_{r,j,t} + \sum_{(j,i)\in E} f_{r,j,i,t} \qquad \forall j \in V, r \in R, t \in \{rld_r + 1, \dots, T\}$$

The cargo flow constraints for each requirement must be satisfied at each port and time step. Note that the flow and inventory variables for each requirement are 0 prior to time rld_r.

$$I_{r,pod_r,t} = 1 \qquad \forall \mathrm{r} \in \mathrm{R}, \mathrm{t} \in \{\mathrm{lad}_r, \dots, \mathrm{T}\}$$

The cargos for each requirement must be delivered to the corresponding POD by the latest arrival date.

$$I_{r,pod_r,t} = 0 \qquad \forall \mathrm{r} \in \mathrm{R}, \mathrm{t} \in \{0, \dots, \mathrm{ead}_r - 1\}$$

Cargos are not allowed to be delivered before the earliest arrival date.

Aircraft Flow Constraint:

$$y_a = s_{a,w_a,st_a} + \sum_{(w_a,j)\in E} x_{a,w_a,j,st_a} \qquad \forall a \in A$$

The supply for each aircraft type a is determined by the integer variable y_a. Each aircraft type a becomes available for use starting at time st_a at port w_a. At such time step, the aircraft may fly to another port or stay idle at the wing.

$$s_{a,j,t-1} + \sum_{(i,j)\in E} x_{a,i,j,(t-\tau_{i,j,a})} = s_{a,j,t} + \sum_{(j,i)\in E} x_{a,j,i,t} \qquad \forall j \in V, a \in A, t \in \{st_a + 1, \dots, T\}$$

The aircraft flow constraints for each aircraft type must be satisfied at each port and time step. Similar to the cargo flow variables, the flow and inventory variables for each aircraft type are 0 prior to st_a.

$$s_{a,w_a,t} = y_a \qquad \forall \mathrm{t} \in \{et_a, \dots, T\}$$

The aircraft must be back at their corresponding wing by the end of the availability time window. This is enforced by setting the inventory level for each aircraft type at the corresponding wing to the initial allocated number of aircraft for all time period from et_a(time when aircraft are due) till the end of the planning horizon period.

$$x_{a,i,j,t} = 0 \quad a \in A, t \in \{0, \dots, st_a - 1\} \cup \{et_a, \dots, T\}$$

Aircraft flow variables are forced to be 0 outside the availability time window.

Capacity Constraint:

$$\sum_{r\in R} \bar{d}_r \times f_{r,i,j,t} \le \sum_{a\in A} cap_a \times x_{a,i,j,t} \qquad \forall\, (i,j) \in E, t \in \{0, \dots, T\} \qquad (5.2.1)$$

The total cargo flow assigned to each arc at each time step must be less than the capacity.

Port Capacity:

$$\sum_{a\in A}\left(s_{a,j,t-1} + \sum_{(i,j)\in E} x_{a,i,j,(t-\tau_{i,j,a})}\right) \leq PortCap_j \qquad \forall j \in V, t \in \{0,\dots,T\}$$

At each time step, the total number of aircraft at each port must be less than the port's capacity.

Non-negativity Constraints:

$$f_{r,i,j,t} \geq 0\,, \qquad \forall\,(i,j) \in E, r \in R, t \in \{0,\dots,T\}$$

$$I_{r,i,t} \geq 0 \quad \forall\, i \in V, r \in R, t \in \{0,\dots,T\}$$

$$y_a \geq 0, y_a \in \mathbb{Z} \qquad \forall \mathrm{a} \in A$$

$$s_{a,i,t} \geq 0, s_{a,i,t} \in \mathbb{Z} \quad \forall \mathrm{a} \in A, i \in V, t \in \{0,\dots,T\}$$

5.3 Robust Formulation

This section presents a robust formulation for addressing demand uncertainty in the airlift planning problem. As mentioned in the previous section, this study only considers the demand uncertainty due to the arrivals of new requirements over the time horizon. The set of requirements R that was introduced is further categorized into a set of requirements R_k that is known with certainty and a set of requirements R_u that is uncertain. The uncertain requirements are requirements that are not in the system prior to execution but may appear during the operation. Given that a requirement appears, the information that is available is the forecast value $\bar{d}$ and an upper bound value $\bar{d} + \hat{d}$. Although it is not known if the requirement from set R_k will materialize during operation, it is assumed that specifications regarding the time window, port of embarkation, and port of debarkation are known with certainty.

To accommodate the "zero or nothing" nature of the uncertainty in the demand, an uncertainty set motivated by the concept of budget of uncertainty (Bertsimas & Sim, 2002) is introduced. Suppose that there are m uncertain requirements, the uncertainty set in (3.3.1) assumes that these m requirements will materialize in the system, and a Γ parameter is used to adjust the robustness level of the solution. However, in this setting, it is possible that less than m requirements will appear in the system, and as a result, the uncertainty set in (3.3.1) may produce a result that is too conservative. Furthermore, it does not provide much flexibility in terms of how solutions with varying degrees of robustness can be generated. Thus, a slight variation to the uncertainty set in (3.3.1) is considered in (5.3.1).

$$U^d(\beta,\Gamma) = \left\{ \begin{array}{c} \tilde{d}_i = \bar{d}_i s_i + \hat{d}_i z_i, \quad i \in R_u \\ \sum_{i \in R_u} s_i \leq \beta \\ \sum_{i \in R_u} z_i \leq \Gamma \\ 0 \leq s_i \leq 1, i \in R_u \\ 0 \leq z_i \leq s_i, i \in R_u \end{array} \right\} \quad (5.3.1)$$

The β and Γ parameters can be adjusted to control the robustness level of the solution. A solution obtained from a robust formulation using this uncertainty set will be deterministically feasible if up to β requirements materialize during the operation and the demand values of at most Γ requirements out of these β are at their maximum values while the demand values of the remaining $\beta - \Gamma$ requirements are at most at their nominal values.

Once the uncertainty set is established, the robust formulation constructs a solution that is always feasible for realizations that belong to the set. For this particular problem, the robust formulation differs from the deterministic formulation only in the objective function and the capacity constraint.

The capacity constraint from (5.2.1) can be expressed as:

(5.3.2):

$$\sum_{r \in R_k} d_r \times f_{r,i,j,t} + \sum_{r \in R_u} \tilde{d}_r \times f_{r,i,j,t} \leq \sum_{a \in A} cap_a \times x_{a,i,j,t}$$

$$\forall\, (i,j) \in E, t \in 1..T, \tilde{d} \in U^d(\beta,\Gamma)$$

<u>Theorem 5.1:</u>

Constraint (5.3.2) can be re-expressed as:

$$\sum_{r \in R_n} (d_r \times f_{r,i,j,t}) + \beta \times p_{i,j,t} + \Gamma \times q_{i,j,t} + \sum_{r \in R_u} m_{r,i,j,t} \leq \sum_{a \in A} cap_a \times x_{a,i,j,t}$$

$$\forall (i,j) \in E, t \in \{0..T\}$$

$$p_{i,j,t} + m_{r,i,j,t} - n_{r,i,j,t} \geq \bar{d}_r f_{r,i,j,t} \qquad \forall (i,j) \in E, t \in \{0..T\}, r \in R_u$$

$$q_{i,j,t} + n_{r,i,j,t} \geq \hat{d}_r f_{r,i,j,t} \qquad \forall (i,j) \in E, t \in \{0..T\}, r \in R_u$$

$$p_{i,j,t} \geq 0, q_{i,j,t} \geq 0 \qquad \forall (i,j) \in E, t \in t \in \{0..T\}$$

$$m_{r,i,j,t} \geq 0, n_{r,i,j,t} \geq 0 \quad \forall (i,j) \in E, t \in t \in \{0..T\}, r \in R_u$$

Proof:

Constraint (5.3.2) can be expressed as:

$$\sum_{r\in R_k} d_r \times f_{r,i,j,t} + \underset{\tilde{d}\,\in\, U^d(\beta,\Gamma)}{maximize} \sum_{r\in R_u} \tilde{d}_r \times f_{r,i,j,t} \le \sum_{a\in A} cap_a \times x_{a,i,j,t}$$

$$\forall\,(i,j) \in E, t \in \{0..T\}$$

For a given set of $(i,j) \in E, t \in \{0..T\}$, and $f_{r,i,j,t}$, consider the auxiliary optimization problem:

$$\underset{\tilde{d}\,\in\, U^d(\beta,\Gamma)}{maximize} \sum_{r\in R_u} \tilde{d}_r \times f_{r,i,j,t}$$

This equals to:

$$maximize \sum_{r\in R_u} \left(\bar{d}_i s_i + \hat{d}_i z_i\right) \times f_{r,i,j,t}$$

st:

$$\sum_{i\in R_u} s_i \le \beta$$

$$\sum_{i\in R_u} z_i \le \Gamma$$

$$0 \le s_i \le 1 \;\; \forall i \in R_u$$

$$0 \le z_i \le s_i \;\; \forall i \in R_u$$

Consider the dual:

$$minimize \;\; \beta \times p + \Gamma \times q + \sum_{r\in R_u} m_r$$

st:

$$p + m_r - n_r \ge \bar{d}_r f_{r,i,j,t} \qquad \forall r \in R_u$$

$$q + n_r \ge \hat{d}_r f_{r,i,j,t} \qquad \forall r \in R_u$$

$$p \ge 0, q \ge 0$$

$$m_r \ge 0\,, n_r \ge 0 \;\; \forall r \in R_u$$

Thus, (5.3.2) can be re-expressed as:

(CRC):

$$\sum_{r\in R_n}(d_r \times f_{r,i,j,t}) + \beta \times p_{i,j,t} + \Gamma \times q_{i,j,t} + \sum_{r\in R_u} m_{r,i,j,t} \le \sum_{a\in A} cap_a \times x_{a,i,j,t}$$

$$\forall\ (i,j) \in E, t \in \{0..T\}$$

$$p_{i,j,t} + m_{r,i,j,t} - n_{r,i,j,t} \ge \bar{d}_r f_{r,i,j,t} \qquad \forall(i,j) \in E, t \in \{0..T\}, r \in R_u$$

$$q_{i,j,t} + n_{r,i,j,t} \ge \hat{d}_r f_{r,i,j,t} \qquad \forall(i,j) \in E, t \in \{0..T\}, r \in R_u$$

$$p_{i,j,t} \ge 0, q_{i,j,t} \ge 0 \qquad \forall(i,j) \in E, t \in \{0..T\}$$

$$m_{r,i,j,t} \ge 0, n_{r,i,j,t} \ge 0 \quad \forall(i,j) \in E, t \in \{0..T\}, r \in R_u$$

Now consider the objective function(OBR) of the robust formulation:

$$minimize \left(\sum_{a\in A_C} C_a \times y_a + \sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R_k}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times d_r \times f_{r,i,j,t} + \underset{\tilde{d} \in U^d(\beta,\Gamma)}{maximize} \sum_{r\in R_u}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times \tilde{d}_r \times f_{r,i,j,t} \right)$$

<u>Theorem 5.2:</u>

(OBR) can be re-expressed as:

$$minimize \left(\begin{array}{c} \sum_{a\in A_C} C_a \times y_a + \sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R_k}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times d_r \times f_{r,i,j,t} + \\ \beta \times \varphi + \Gamma \times \omega + \sum_{r\in R_u} \vartheta_r \end{array} \right)$$

$$\varphi + \vartheta_r - \pi_r \ge \bar{d}_r \times \left(\sum_{(i,j)\in E}\sum_{t} \beta_{r,i,j} \times f_{r,i,j,t} \right) \qquad \forall r \in R_u$$

$$\omega + \pi_r \ge \hat{d}_r \times \left(\sum_{(i,j)\in E}\sum_{t} \beta_{r,i,j} \times f_{r,i,j,t} \right) \qquad \forall r \in R_u$$

$$\varphi \ge 0, \omega \ge 0$$

$$\vartheta_r \geq 0, \pi_r \geq 0 \quad \forall r \in R_u$$

Proof:

Consider the auxiliary optimization problem:

$$\underset{\tilde{d} \in U^d(\beta,\Gamma)}{maximize} \sum_{r \in R_u} \sum_{(i,j) \in E} \sum_{t=0}^{T} \beta_{r,i,j} \times \tilde{d}_r \times f_{r,i,j,t}$$

This equals to:

$$maximize \sum_{r \in R_u} \left(\bar{d}_i s_i + \hat{d}_i z_i\right) \times \left(\sum_{(i,j) \in E} \sum_{t=0}^{T} \beta_{r,i,j} \times f_{r,i,j,t}\right)$$

st:

$$\sum_{i \in R_u} s_i \leq \beta$$

$$\sum_{i \in R_u} z_i \leq \Gamma$$

$$0 \leq s_i \leq 1 \quad \forall i \in R_u$$

$$0 \leq z_i \leq s_i \quad \forall i \in R_u$$

Consider the dual:

$$minimize \ \beta \times \varphi + \Gamma \times \omega + \sum_{r \in R_u} \vartheta_r$$

st:

$$\varphi + \vartheta_r - \pi_r \geq \bar{d}_r \times \left(\sum_{(i,j) \in E} \sum_{t=0}^{T} \beta_{r,i,j} \times f_{r,i,j,t}\right) \qquad \forall r \in R_u$$

$$\omega + \pi_r \geq \hat{d}_r \times \left(\sum_{(i,j) \in E} \sum_{t=0}^{T} \beta_{r,i,j} \times f_{r,i,j,t}\right) \qquad \forall r \in R_u$$

$$\varphi \geq 0, \omega \geq 0$$

$$\vartheta_r \geq 0, \pi_r \geq 0 \quad \forall r \in R_u$$

Thus the (OBR) can be re-expressed as:

$OBRC$:

$$\text{minimize} \left(\begin{array}{c} \sum_{a\in A_C} C_a \times y_a + \sum_{a\in A} \sum_{(i,j)\in E} \sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R_k} \sum_{(i,j)\in E} \sum_{t=0}^{T} \beta_{r,i,j} \times d_r \times f_{r,i,j,t} + \\ \beta \times \varphi + \Gamma \times \omega + \sum_{r\in R_u} \vartheta_r \end{array} \right)$$

$$\varphi + \vartheta_r - \pi_r \geq \bar{d}_r \times \left(\sum_{(i,j)\in E} \sum_{t=0}^{T} \beta_{r,i,j} \times f_{r,i,j,t} \right) \qquad \forall r \in R_u$$

$$\omega + \pi_r \geq \hat{d}_r \times \left(\sum_{(i,j)\in E} \sum_{t=0}^{T} \beta_{r,i,j} \times f_{r,i,j,t} \right) \qquad \forall r \in R_u$$

$$\varphi \geq 0, \omega \geq 0$$

$$\vartheta_r \geq 0\, , \pi_r \geq 0 \quad \forall r \in R_u$$

The full formulation of the robust counterpart can be obtained by using the deterministic formulation $(ALNM)$ as a base and replacing the objective function and constraint (5.2.1) with $(OBRC)$ and (CRC) respectively.

The uncertainty set (3.3.1) provides an easy intuitive mechanism for adjusting and comparing the robustness level of a solution – the higher the Γ value, the higher the robustness of the solution. On the other hand, comparing the robustness levels for solutions with different β and Γ values is not as trivial for the formulation that uses the uncertainty set (5.3.1). To reiterate the intuition for how the β and Γ parameters affect the performance of the system, a solution obtained from a robust formulation using this uncertainty set will be deterministically feasible if up to β requirements materialize during the operation and the demand values of at most Γ requirements out of these β are at their maximum values while the demand values of the remaining $\beta - \Gamma$ requirements are at most at their nominal values. For a fixed value of β, a plan with a higher Γ value is at least as robust as a plan with lower Γ values. Similarly, given a fixed Γ value, a plan with higher β value is at least as robust as a plan with lower β values. However, a plan with higher β value in general is not necessarily more robust than a plan with a lower β

value. For instance, consider a plan $P1$ which was constructed using $(\beta = 1, \Gamma = 1)$ and a plan $P2$ which was constructed using $(\beta = 2, \Gamma = 0)$. In this case, the robustness levels of the two plans can't be relatively compared. Either $P1$ or $P2$ can be more robust depending on the problem structure.

Consider a simple example where there are 3 nodes as shown in Figure 5.15, and there are 2 uncertain requirements #1 and #2 in the system. Suppose requirement #1 has a demand requirement of 20 units with an upper bound of 25, and requirement #2 has a demand requirement of 25 and an upper bound 30. Furthermore, assume that both requirements are available for delivery at time period 0 and must be delivered to the POD by period 1. For simplicity, ignore the aircraft flow constraints for the moment and assume that the allocated capacity can take any value that minimizes the cost.

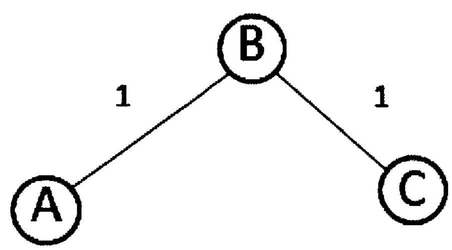

Figure 5.15: Undirected network with 3 nodes. The numbers on the arc represents the travel times.

In the first scenario, let port A and port B be the POE and POD for requirement #1 respectively, and let port C and port B be the POE and POD for requirement #2 respectively. The robust model constructs a solution such that a plan is deterministically feasible if at most β uncertain requirements materialize, the demand values of at most Γ out of the β requirements that materialize are at their maximum value, and the demand values of the other $\beta - \Gamma$ requirements are less than or equal to their expected values. Applying this interpretation to this particular setting, $P1$ is deterministically feasible if at most one requirement flows through each at a certain time step and the demand for such requirement is at its maximum value. Accordingly, $P1$ would allocate 25 units of capacity from port A to port B, and 30 units of capacity from port B to port C. On the other hand, plan 2 is deterministically feasible if at most 2 requirements flow through an arc but none of the requirements are allowed to go beyond their expected values. For this particular example, since it is only possible for 1 requirement to flow from A to B and from C to

B, it is suffice for the solution to allocate 20 units to arc A-B and 25 units to arc C-B to be deterministically feasible if at most 2 requirements were to materialize. In this case, $P1$ allocates more capacity than $P2$ and is relatively more robust.

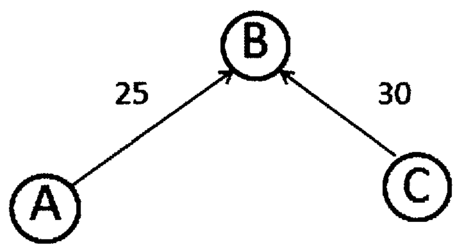

Figure 5.16: Allocation for plan *P*1 under the first scenario

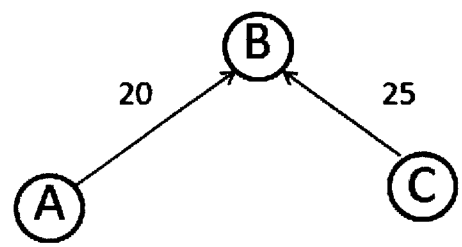

Figure 5.17: Allocation for plan *P*2 under the first scenario

Now consider a second scenario where the POE and POD for both requirements #1 and #2 are at port A and port B respectively. Under such setting, $P1$ would allocate 30 units to arc A-B and would guarantee feasibility if at most one requirement appears and is at its maximum value. On the other hand, plan 2 would allocate 45 units since it must be protected against the worst instance which is when 2 requirements with demand equaling their expected values appear. Thus under this scenario, $P1$ allocates more capacity than $P2$ and is relatively more robust.

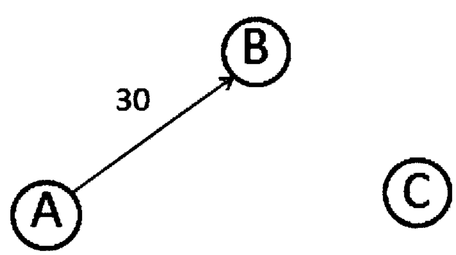

Figure 5.18 Allocation for plan *P*1 under the second scenario

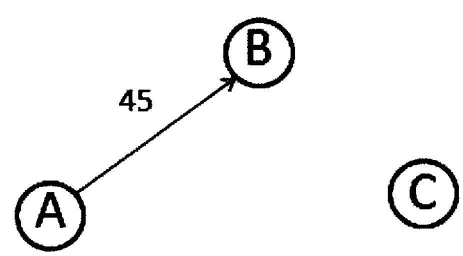

Figure 5.19 Allocation for plan *P*2 under the second scenario

5.4 Execution Policies

This section introduces the execution policies that will be used to evaluate the performance of the plans. The only recourse action that will be considered in this chapter is to outsource excess demand to third party companies at a certain penalty cost. Similar to what have been presented in chapter 3, three levels of systemic flexibility are considered:

- Level 1: All decisions are fixed
- Level 2: Cargo routes are allowed to be re-optimized after the demand values are realized.
- Level 3: Aircraft and cargo routes are allowed to be re-optimized after the demand values are realized.

Let $\dddot{d}_r$ be the demand realization for each requirement r, $pen_{i,j}$ be the cost for outsourcing one unit of demand to a third part company, and set R contain all the realized requirements. The details of each execution policies are presented below.

<u>Policy 1 Flexibility 1 (P1F1):</u>

All decisions from the plan are fixed. The total operating cost is calculated by summing up the fixed aircraft usage cost, aircraft routing cost, cargo routing cost, and the penalty cost.

<u>Input from tactical planning model:</u>

y_a Number of aircraft type a allocated for the planning horizon

$x_{a,i,j,t}$ Number of aircraft type a leaving port i for port j at time t

$f_{r,i,j,t}$ Fraction of demand from requirement r leaving port i for port j at time t

The penalty cost can be calculated using:

$$z_{p1f1} = \sum_{(i,j)\in E} \sum_{t=0}^{T} pen_{i,j} \times maximum\left(\sum_{r\in R} \dddot{d}_r \times f_{r,i,j,t} - \sum_{a\in A} cap_a \times x_{a,i,j,t}, 0\right)$$

(TC_{p1f1}):

$$Total\ Cost = \sum_{a\in A} C_a \times y_a + \sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j,t} + z_{p1f1}$$

Policy 1 Flexibility 2 (P1F2):

Cargo routing decisions are allowed to be re-optimized after the demand values are realized.

Input from tactical planning model:

y_a Number of aircraft type a allocated for the planning horizon

$x_{a,i,j,t}$ Number of aircraft type a leaving port i for port j at time t

Decision variable:

$f_{r,i,j,t}$ Fraction of demand from requirement r leaving port i for port j at time t

$\rho_{i,j,t}$ Unit of unmet demand for trips leaving port i for port j at time t

$$z_{p1f2} = minimize \sum_{r\in R}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times \ddot{d}_r \times f_{r,i,j,t} + \sum_{(i,j)\in E}\sum_{t=0}^{T} pen_{i,j} \times \rho_{i,j,t}$$

st:

$$1 = I_{r,poe_r,ald_r} + \sum_{(poe_r,j)\in E}\sum_{a\in A} f_{r,poe_r,j,ald_r} \quad \forall r \in R$$

$$I_{r,j,t-1} + \sum_{(i,j)\in E} f_{r,i,j,(t-\tau_{i,j,a})} = I_{r,j,t} + \sum_{(j,i)\in E} f_{r,j,i,t} \quad \forall j \in port, r \in R, t \in \{rld_r + 1, \dots, T\}$$

$$\sum_{r\in R} \ddot{d}_r \times f_{r,i,j,t} \le \sum_{a\in A} cap_a \times x_{a,i,j,t} + \rho_{i,j,t} \quad \forall\ (i,j) \in E, t \in 1..T$$

$$I_{r,pod_r,t} = 1 \quad \forall \mathrm{r} \in \mathrm{R}, \mathrm{t} \in \{\mathrm{lad}_r, \dots, \mathrm{T}\}$$

$$I_{r,pod_r,t} = 0 \quad \forall \mathrm{r} \in \mathrm{R}, \mathrm{t} \in \{0, \dots, \mathrm{ead}_r - 1\}$$

$$f_{r,i,j,t} \ge 0 \quad \forall\ (i,j) \in E, r \in R, t \in \{0, \dots, T\}$$

$$\rho_{i,j,t} \quad \forall\ (i,j) \in E, t \in \{0, \dots, T\}$$

(TC_{p1f2}):

$$Total\ Cost = \sum_{a\in A} C_a \times y_a + \sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + z_{p1f2}$$

Policy 1 Flexibility 3(P1F3):

Under the highest flexibility level, the aircraft routings and the cargo routings are allowed to be re-optimized. The formulation of the execution model under this policy follows closely to $(ALNM)$. The difference is in the objective function.

Input from tactical planning model:

y_a Number of aircraft type a allocated for the planning horizon

Decision variable:

$f_{r,i,j,t}$ Fraction of demand from requirement r leaving port i for port j at time t

$x_{a,i,j,t}$ Number of aircraft type a leaving port i for port j at time t

$\rho_{i,j,t}$ Unit of unmet demand for trips leaving port i for port j at time t

$$z_{p1f3} = minimize \left(\sum_{a\in A}\sum_{(i,j)\in E}\sum_{t=0}^{T} \gamma_{a,i,j} \times x_{a,i,j,t} + \sum_{r\in R}\sum_{(i,j)\in E}\sum_{t=0}^{T} \beta_{r,i,j} \times \dddot{d}_r \times f_{r,i,j,t} + \sum_{(i,j)\in E}\sum_{t=0}^{T} pen_{i,j} \times \rho_{i,j,t} \right)$$

st:

$$\sum_{r\in R} \dddot{d}_r \times f_{r,i,j,t} \le \sum_{a\in A} cap_a \times x_{a,i,j,t} + \rho_{i,j,t} \quad \forall\, (i,j) \in E, t \in 1..T$$

Other constraints in $(ALNM)$ excluding (5.2.1)

(TC_{p1f3}):

$$Total\ Cost = \sum_{a \in A} C_a \times y_a + z_{p1f3}$$

5.5 Experiment Analysis

This section brings together the different components that were discussed in section 5.3 and 5.4 to analyze the performance of the system under different flexibility levels and different robustness of a plan. The analysis will be conducted under the simple undirected network configuration shown in Figure 5.20, where the flying times between the nodes are assumed to be 1 time step, and the wing ports are assumed to be located at ports A and B. Each aircraft type has a capacity of 10 units and is assumed to be available for the whole planning horizon which is set to be 10 periods. This means that the aircraft could be utilized in any fashion over the time horizon but must be back at the corresponding wing by the end of the time period. The summary of the parameters used in the study are summarized in Table 5.13, and the specifications of the requirements that are known with certainty are summarized in Table 5.14. For simplicity, the study also assumes that there is no cargo flow cost, implying that an empty flight costs the same as a loaded flight, and that there is no limit on the number of aircraft that can be accommodated at each port at any given time.

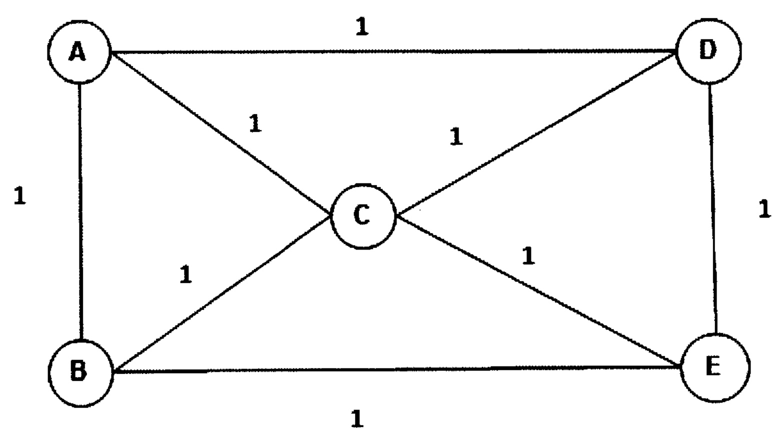

Figure 5.20: Network configuration used in the analysis. Wings are located at ports A and B. The numbers on the arcs represent the flight times.

Fixed Cost	Aircraft Routing Cost	Aircraft Capacity	Planning Horizon Period
$10,000	$1,000 per 1 hour	10	10

Table 5.13: Parameter values used in the experiment

Requirement	POE	POD	Demand	ALD	EAD	LAD
1	A	C	105	2	2	8
2	A	D	116	1	1	7
3	A	E	84	2	4	8
4	B	C	93	1	3	7
5	B	D	102	2	4	8
6	B	E	136	3	3	9

Table 5.14: Requirements that are known with certainty

The study considers 3 test instances of different levels of uncertainties involved in the system. On top of the 6 requirements shown in Table 5.14, test instance 1 has 3 additional uncertain requirements, test instance 2 has 4 additional uncertain requirements, and test instance 3 has 6 additional uncertain requirements. The details of the uncertain requirements, including the probability that each requirement will materialize during operation, for each of the test instance are shown in Table 5.15, Table 5.17, and Table 5.19. Given that a requirement appears, the demand follows a triangular distribution with lower bound, mean, and upper bound given by $\bar{d} - \hat{d}, \bar{d}$, and $\bar{d} + \hat{d}$ respectively, where $\hat{d}$ is the maximum deviation. The optimization results for the selected robust levels for each test instances are presented in Table 5.16, Table 5.18, and Table 5.20. The simulation procedure ($SimProc$) from chapter 3 is used to evaluate the performances for each of the robust plans under the 3 execution policies. A penalty cost of $1,000 per unit of demand per hour is incurred.

Requirement	POE	POD	Demand	Maximum Deviation	ALD	EAD	LAD	Probability
7	A	D	100	15	3	4	8	0.5
8	B	E	75	10	2	3	7	0.8
9	C	D	84	8	2	3	7	0.3

Table 5.15: Uncertain requirements for demand pattern 1

Beta	Gamma	Fleet size at wing A	Fleet size at wing B	Cost ($)
0	0	5	13	327,000
1	0	7	16	412,000
1	1	6	18	427,000
2	0	10	14	433,000
2	1	12	13	447,000

Table 5.16: Optimization results for demand pattern 1

Requirement	POE	POD	Demand	Maximum Deviation	ALD	EAD	LAD	Probability
7	A	E	80	15	2	3	8	0.5
8	C	E	95	20	2	2	9	0.8
9	B	C	85	15	2	3	9	0.7
10	B	D	70	10	1	4	8	0.3

Table 5.17: Uncertain requirements for demand pattern 2

Beta	Gamma	Fleet size at wing A	Fleet size at wing B	Cost ($)
2	1	8	18	473,000
2	2	8	18	476,000
3	0	8	18	479,000
3	1	9	18	499,000
3	2	10	17	502,000

Table 5.18: Optimization results demand pattern 2

Requirement	POE	POD	Demand	Maximum Deviation	ALD	EAD	LAD	Probability
7	A	E	80	15	2	3	8	0.5
8	C	E	95	20	2	2	9	0.8
9	A	D	75	10	1	3	8	0.5
10	B	C	85	15	2	3	9	0.7
11	B	D	70	10	1	4	8	0.3
12	B	E	90	15	2	3	9	0.6

Table 5.19: Uncertain requirements for demand pattern 3

Beta	Gamma	Fleet size at wing A	Fleet size at wing B	Cost ($)
3	2	11	19	551,000
3	3	13	18	565,000
4	1	8	23	572,000
4	2	11	21	588,000

Table 5.20: Optimization results for demand pattern 3

5.5.1 Simulation Results for Demand Pattern 1

The cumulative distribution function of the total operating cost for the plans with different levels of robustness under the 3 execution policies are shown in Figure 5.21, Figure 5.22, and Figure 5.23. The summary of the statistics from the simulation are presented in Table 5.21, Table 5.22, and Figure 5.24. Under flexibility level 1, it is clear from Figure 5.21 that the plan with $(\beta = 0, \Gamma = 0)$ performs relatively worse than the other plans. Note that the $(\beta = 0, \Gamma = 0)$ plan was constructed without considering the uncertain requirements, and corresponds to a plan with no added robustness. Accordingly, it is expected that such plan will perform poorly under a system with no flexibility. As the flexibility level of the system increases to level 1, the performance of the $(\beta = 0, \Gamma = 0)$ plan performs substantially better as suggested by the 40% decrease in the expected total cost. On the other hand, the decreases in the expected total cost for the other plans are all below 5%. Finally, the improvement in the performance of the system for each plan as the flexibility level increases from level 2 to level 3 is very marginal, with an average decrease in the expected total cost across the different robustness level of 4%.

Using the expected total cost as the criteria for selecting the optimal robust level for each flexibility level, the $(\beta = 2, \Gamma = 0)$ plan performs best for flexibility level 1 and level 2, and the $(\beta = 1, \Gamma = 0)$ plan performs best for flexibility level 3. On average, the optimal expected cost decreased by approximately 0.3% as the flexibility level increases from level 1 to level 2, 4.99% as the flexibility level increases from level 2 to level 3, and 5.3% as the flexibility level increases from level 1 to level 3.

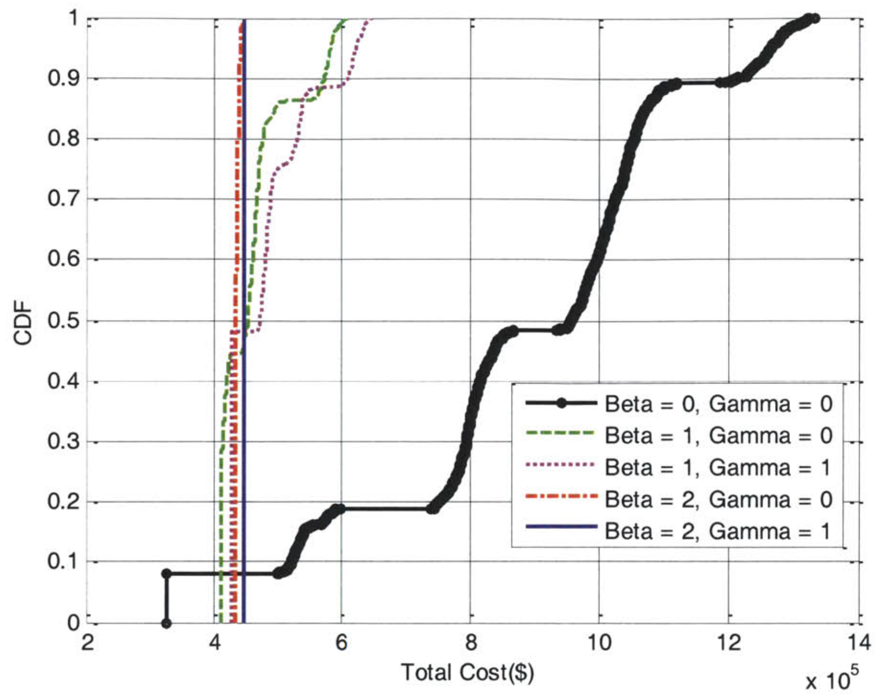

Figure 5.21: CDF of the total cost for the plans under P1F1 policy for demand pattern 1

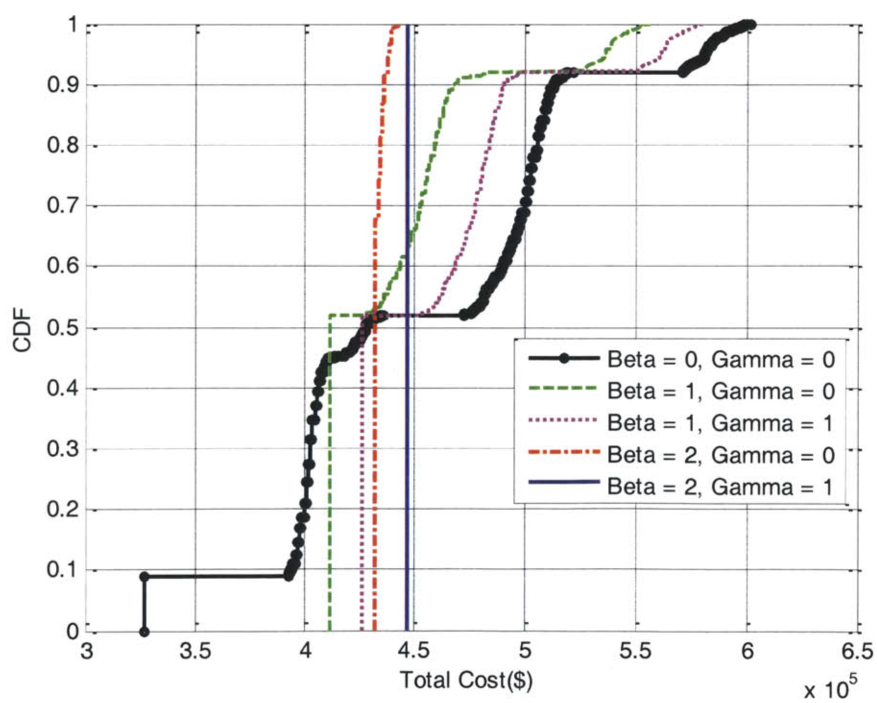

Figure 5.22: CDF of the total cost for the plans under P1F2 policy for demand pattern 1

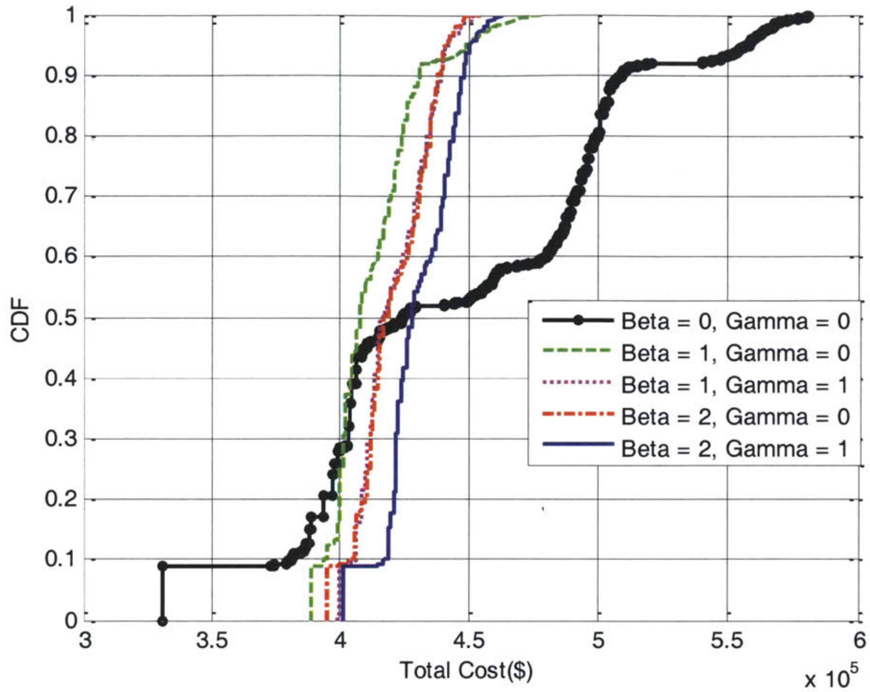

Figure 5.23: CDF of the total cost for the plans under P1F3 policy for demand pattern 1

Beta	Gamma	P1F1 Policy	P1F2 Policy	P1F3 Policy
0	0	875,600 (252,310)	450,643 (68,269)	443,020 (62,953)
1	0	458,260 (54,566)	438,667 (36,104)	412,406 (16,985)
1	1	478,730 (64,299)	458,269 (39,900)	420,810 (13,289)
2	0	435,430 (2,926)	434,104 (2,005)	420,680 (13,944)
2	1	447,000 (0)	447,000 (0)	430,318 (14,083)

Table 5.21: Expected value and standard deviation of total cost for the plans and execution policies for demand pattern 1

Beta	Gamma	% decrease in expected cost from P1F1 to P1F2	% decrease in expected cost from P1F2 to P1F3
0	0	48.5%	1.7%
1	0	4.3%	5.6%
1	1	4.3%	8%
2	0	0.3%	3%
2 ·	1	0%	3.7%

Table 5.22: Decrease in expected cost as the flexibility level of the system increases for demand pattern 1

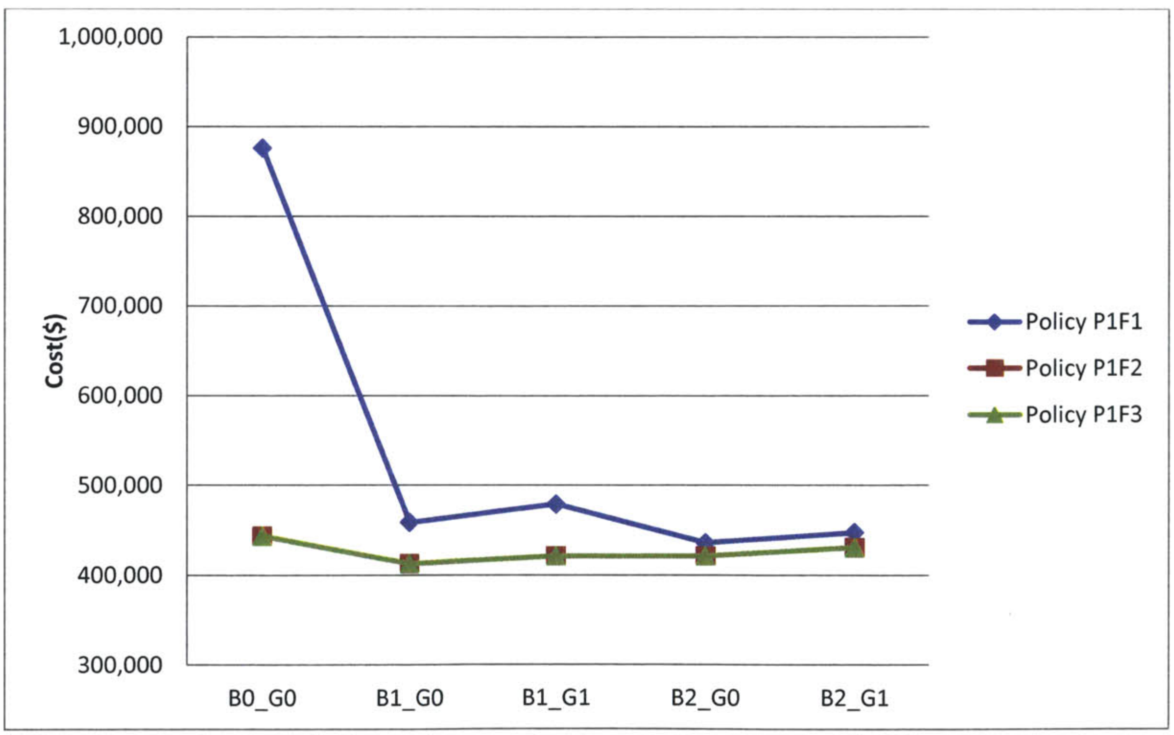

Figure 5.24: Plot of expected total cost for the plans with different degrees of robustness and execution policies with different levels of flexibility for demand pattern 1

5.5.2 Simulation Results for Demand Pattern 2

The cumulative distribution function of the total operation cost for the plans with different levels of robustness under the 3 execution policies are shown in Figure 5.25, Figure 5.26, and Figure 5.27. The summary of the statistics from the simulation are presented in Table 5.23, Table 5.24, and Figure 5.28 . Looking at Figure 5.27, the CDF graph of the $(\beta = 2, \Gamma = 0)$ plan under flexibility level 3 is always to the right of the CDF graph of the other plans, suggesting that the $(\beta = 2, \Gamma = 0)$ plan is being stochastically dominated (i.e. the probability of being below a certain cost is always lower for the $(\beta = 2, \Gamma = 0)$ plan). Under such case, it is clear that the $(\beta = 2, \Gamma = 0)$ plan performs worse than that of the other plans.

The average decrease in the expected total cost across the plans with different robustness is 0.86% as the flexibility level increases from level 1 to level 2, and 7.8% as the flexibility level increases from level 2 to level 3. Using the expected total cost as the criteria for selecting the optimal robust level for each flexibility level, the $(\beta = 3, \Gamma = 0)$ plan performs best for all the 3 flexibility levels. On average, the optimal expected cost decreased by 0.9% as the flexibility level increases from level 1 to level 2, 5.8% as the flexibility level increases from level 2 to level 3, and 6.6% as the flexibility level increases from level 1 to level 3.

Beta	Gamma	P1F1 Policy	P1F2 Policy	P1F3 Policy
2	1	506,240 (43,819)	498,453 (36,984)	453,444 (19,468)
2	2	499,670 (34,169)	490,503 (25,815)	453,444 (19,468)
3	0	485,690 (76,127)	481,290 (4,390)	453,444 (19,468)
3	1	499,960 (3,265)	499,003 (94)	453,444 (19,468)
3	2	502,000 (0)	502,000 (0)	464,110 (19,206)

Table 5.23: Expected value and standard deviation of total cost for the plans and execution policies for demand pattern 2

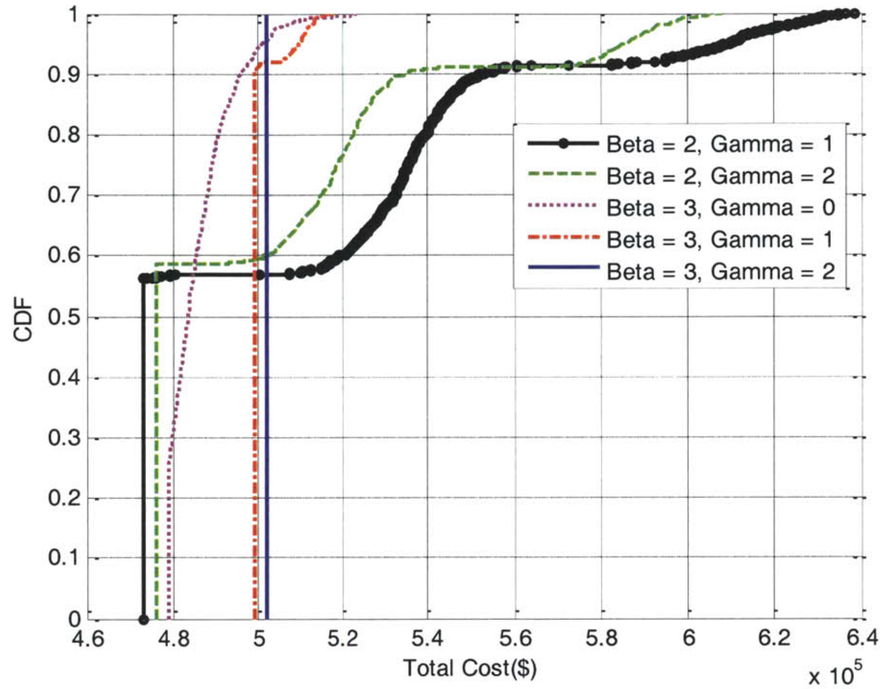

Figure 5.25: CDF of the total cost for the plans under P1F1 policy for demand pattern 2

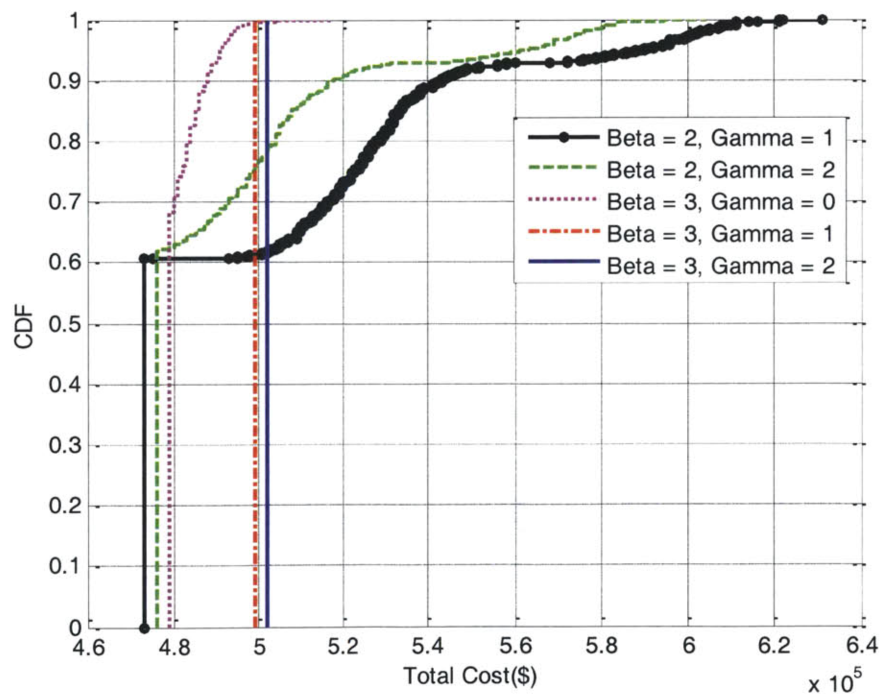

Figure 5.26: CDF of the total cost for the plans under P1F2 policy for demand patter 2

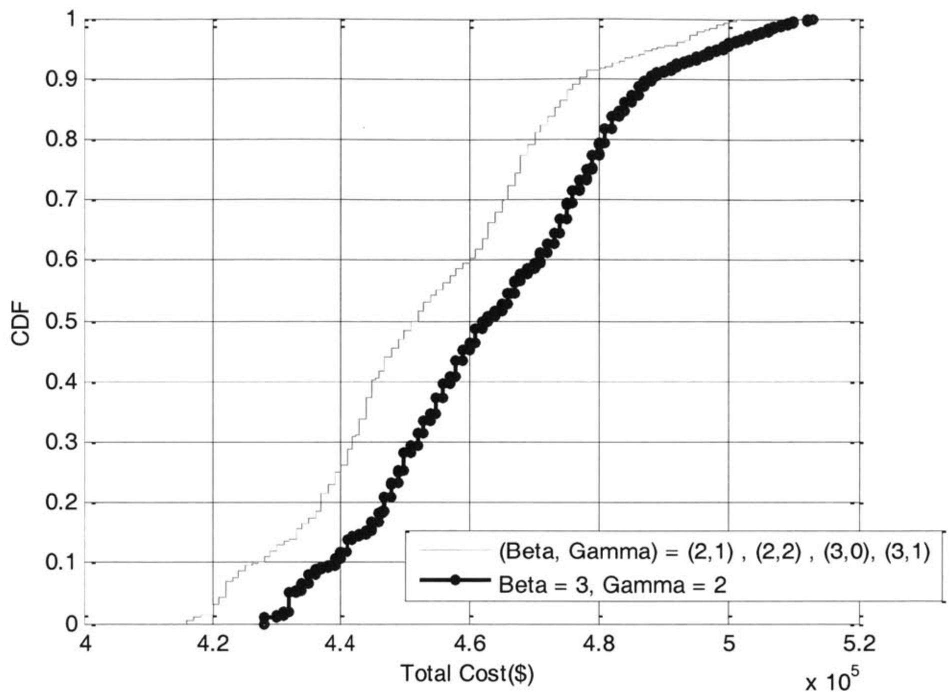

Figure 5.27: CDF of the total cost for the plans under P1F3 policy for demand pattern 2

Beta	Gamma	% decrease in expected cost from P1F1 to P1F2	% decrease in expected cost from P1F2 to P1F3
2	1	1.5%	9%
2	2	1.8%	7.5%
3	0	0.9%	5.7%
3	1	0.2%	9.1%
3	2	0%	7.5%

Table 5.24: Decrease in expected cost as the flexibility level of the system increases for demand pattern 2

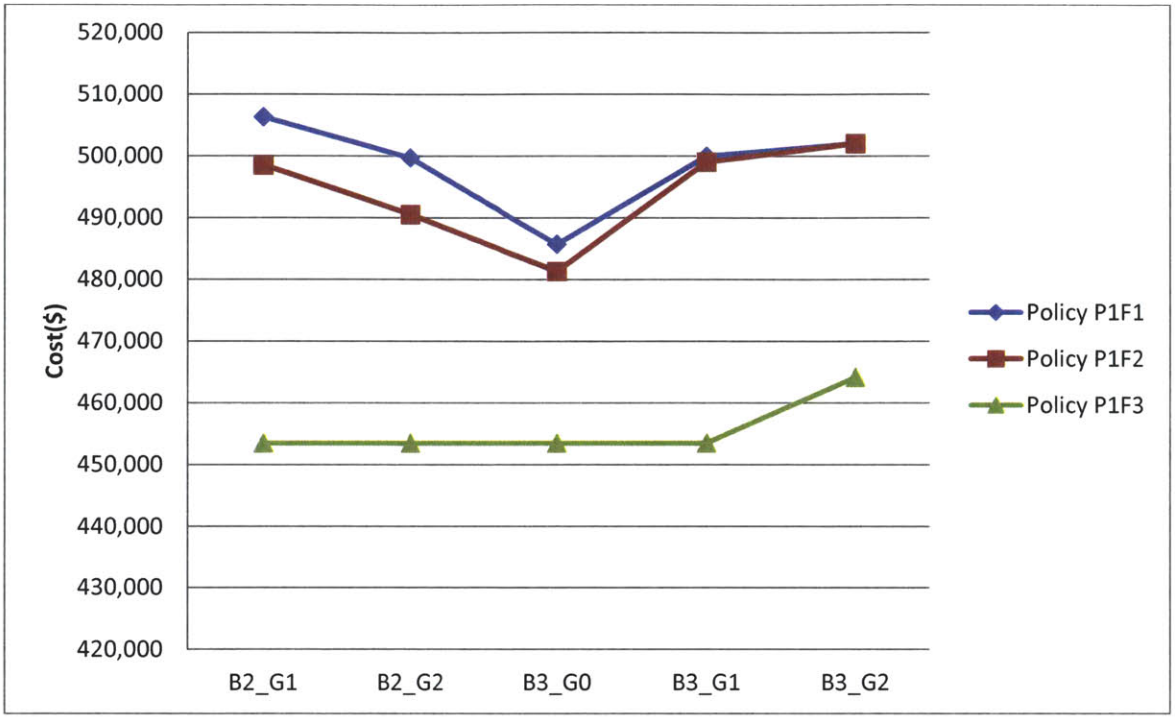

Figure 5.28: Plot of expected total cost for the plans with different degrees of robustness and execution policies with different levels of flexibility for demand pattern 2

5.4.3 Simulation Results for Demand Pattern 3

The cumulative distribution function of the total operation cost for the plans with different levels of robustness under the 3 execution policies are shown in Figure 5.29, Figure 5.30, and Figure 5.31. The summary of the statistics from the simulation are presented in Table 5.25, Table 5.26, and Figure 5.32. Similar to the result from experiment 2, the $(\beta = 3, \Gamma = 2)$ plan stochastically dominates the other plans under flexibility level 3. The average decrease in the expected total cost across the different plans is 0.7% as the flexibility level increases from level 1 to level 2, and 9.1% as the flexibility level increases from level 2 to level 3. Using the expected total cost as the criteria for selecting the optimal robust level for each flexibility level, the $(\beta = 4, \Gamma = 1)$ plan performs best for flexibility levels 1 and 2, and the $(\beta = 3, \Gamma = 2)$ plan performs best for flexibility level 3. On average the optimal expected cost decreased by 0.8% as the flexibility level increases from level 1 to level 2, 10.1% as the flexibility level increases from level 2 to level 3, and 10.8% as the flexibility level increases from level 1 to level 3.

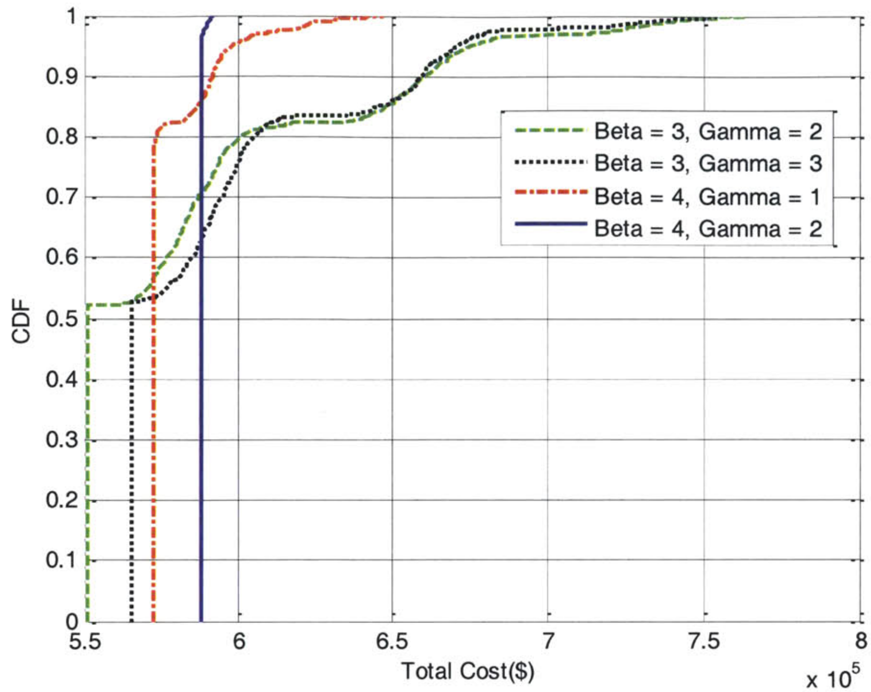

Figure 5.29: CDF of the total cost for the plans under P1F1 policy for demand pattern 3

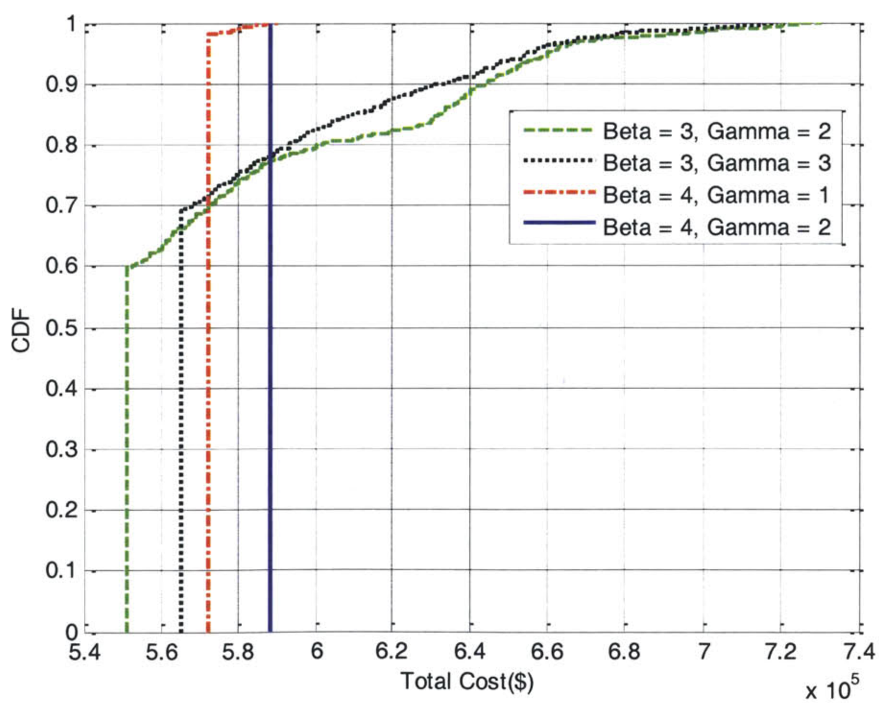

Figure 5.30: CDF of the total cost for the plans under P1F2 policy for demand pattern 3

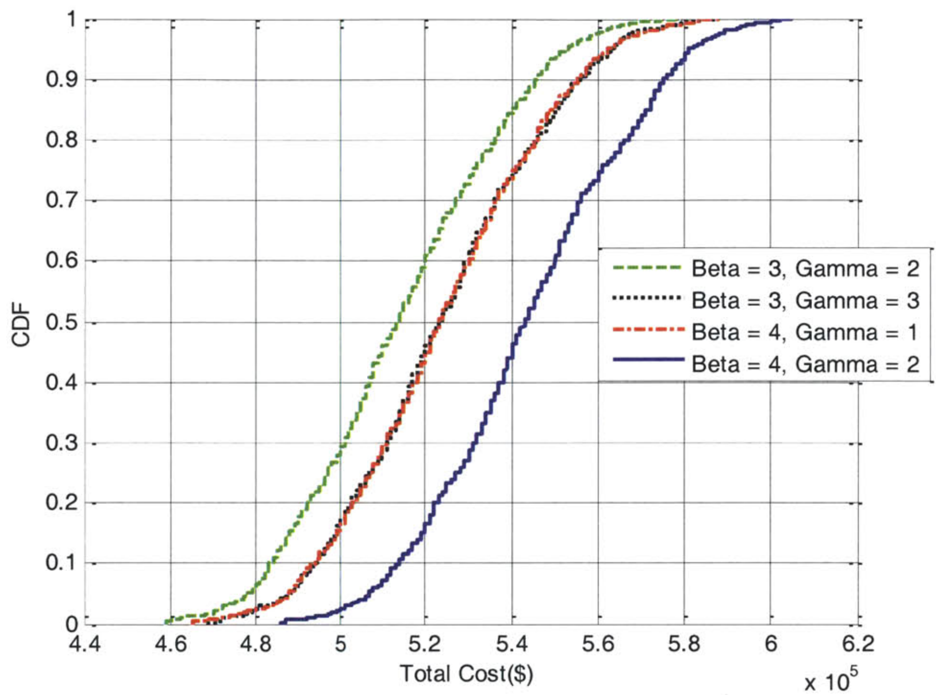

Figure 5.31: CDF of the total cost for the plans under P1F3 policy for demand pattern 3

Beta	Gamma	P1F1 Policy	P1F2 Policy	P1F3 Policy
3	2	582,810 (46,774)	575,167 (40,331)	514,200 (23,346)
3	3	590,810 (39,191)	580,791 (31,288)	524,140 (23,252)
4	1	576,520 (11,274)	572,185 (1,468)	524,150 (23,241)
4	2	588,050 (329)	588,000 (0)	544,114 (23,269)

Table 5.25 Expected value and standard deviation of total cost for the plans and execution policies for demand pattern 3

Beta	Gamma	% decrease in expected cost from P1F1 to P1F2	% decrease in expected cost from P1F2 to P1F3
3	2	1.3%	10.6%
3	3	1.7%	9.75
4	1	0.8%	8.4%
4	2	0.008%	7.5%

Table 5.26: Decrease in expected cost as the flexibility level of the system increases for demand pattern 3

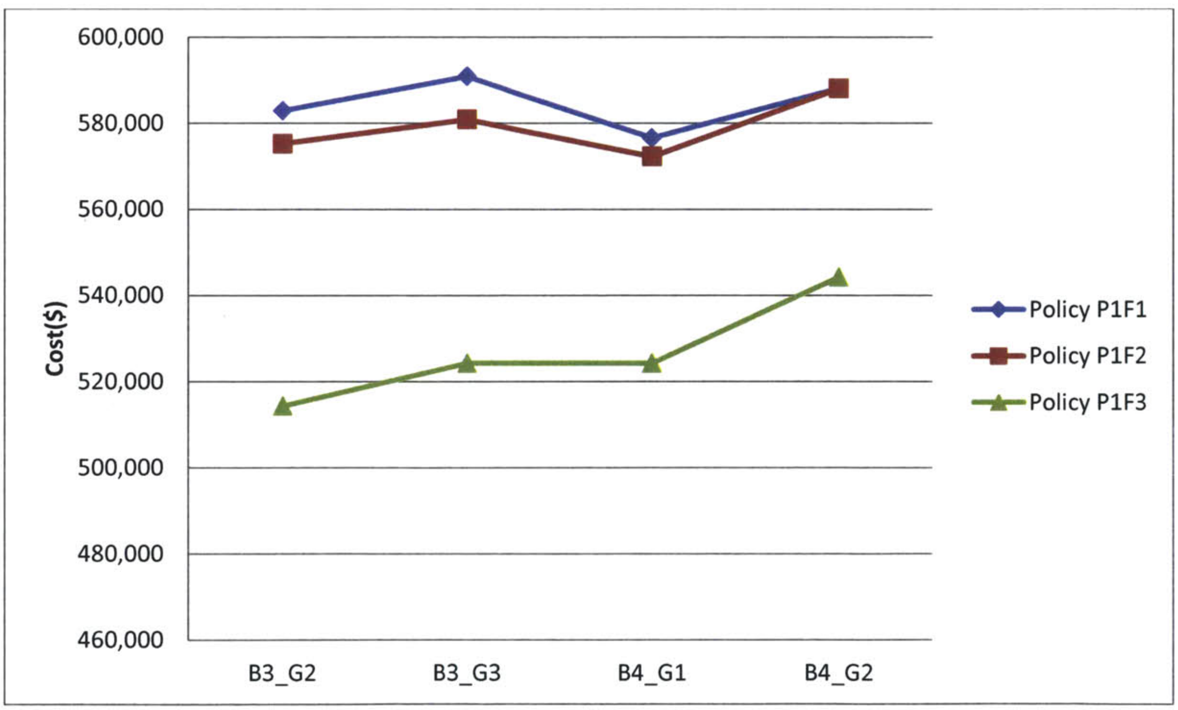

Figure 5.32: Plot of expected total cost for the plans with different degrees of robustness and execution policies with different levels of flexibility for demand pattern 3

5.4.4 Summary

This section provided an analysis of how the interaction between the degree of robustness of a tactical plan and the flexibility level of an execution policy affects the performance of a military airlift system. Tactical plans and execution policies were simulated under three demand settings with varying level of uncertainty, which is defined as the portion of the uncertain requirements to the certain requirements. In the first demand setting where three out of the nine requirements are uncertain, the tactical plan that didn't take into account for the uncertain requirements expectedly performed poorly under the execution policy with no flexibility. However, this tactical plan performed comparatively well with respect to the other tactical plans with added robustness as the flexibility level of the system increases. Using the expected total cost as the evaluating metric, the optimal degree of robustness decreases as the flexibility level increases. For test instances two and three, the performance of every tactical plan increases as the flexibility level increases from level one to three level or level two to level three.

The key findings from these 3 experiments are summarized in Table 5.27. For each test instance with different uncertainty level, the table shows the relative decrease in the optimal expected total cost as the flexibility level of the system increases. Note that the optimal expected total cost for each test instance and flexibility level is determined based on the selected set of robust level, and may not necessarily be the global optimum.

Uncertainty Level of Demand Pattern	% decrease in optimal expected cost from flexibility level 1 to 2	% decrease in optimal expected cost from flexibility level 2 to 3	% decrease in optimal expected cost from flexibility level 1 to 3
1	0.3%	4.99%	5.3%
2	0.9%	5.9%	6.6%
3	0.8%	10.1%	10.8%

Table 5.27: Decrease in the optimal expected cost as the flexibility level increases for each demand pattern

The first finding is that increasing the flexibility level of the system from level 2 to level 3 has much more positive impact on the overall performance of the system than increasing the flexibility level of the system from level 1 to level 2. There were less than 1% decrease in the optimal expected total cost when the flexibility level increase from 1 to 2 for all the 3 test instances but approximately 5% or more when the flexibility level increases from 2 to 3.

In addition, it is interesting to note that as the level of the uncertainty increases, increasing the flexibility level of the system has a greater positive impact on the performance of the system. When the flexibility level increases from level 1 to level 2, the percentage decrease in the optimal expected cost increases from 0.3% for the first test instance to 0.9% for the second test and drops down marginally to 0.8% for the third test instance. For the cases when the flexibility level increases from 2 to 3 or from 1 to 3, the percentage decrease in the optimal expected cost increases as the uncertainty level of the test instance increases.

This page intentionally left blank

Chapter 6

Conclusions

6.1 Summary

This thesis provided an investigation of the different dimensions of incorporating uncertainty into the planning process of a freight transportation system. It first introduced the concepts of robustness and flexibility as means to address uncertainty in Chapter 2. Two types of flexibility were introduced, including *systemic* and *operational* flexibility. As decision makers transition from tactical plan to executional plan, systemic flexibility refers to the degree of freedom to which decisions can be adjusted, changed, or made at the execution level. Given the systemic flexibility level that determines the dynamics of the decision making process, operational flexibility provides the decision makers with options to respond and adapt to changes. Robustness was defined as a metric used to measure the performance of a plan. A plan is considered to be robust if it is able to operate well under various operating conditions without requiring costly adjustments. Planning models are typically formulated based on some assumed systemic flexibility level, as certain output decisions are designed to be fixed. Given the systemic flexibility level, the aim of these models is to create plans that are robust based on certain robustness metric. Two broad classes of methods that are used to make plans robust were introduced. These two classes of approach differ primarily in the utilization of future knowledge. The aim of the first modeling approach that utilizes future knowledge is to create plans that perform well under a wide variety of scenarios without breaking down, i.e. having a high probability of feasibility. On the other hand, the general aim of the second modeling approach that doesn't explicitly incorporate future knowledge into the model is to provide decision makers with options to minimize recovery cost or the impact once disruptions occur.

Chapter 3 introduced a framework for analyzing the effects of the interactions between tactical plans with various degrees of robustness and execution policies with various levels of systemic flexibility on the performance of a transportation distribution system. The study was conducted under a generalized hypothetical distribution planning problem with demand uncertainty where the major decision variables are the fleet size, vehicle routes, and cargo routes. A robust tactical planning model based on the concept of budget of uncertainty (Bertsimas & Sim, 2002) was first introduced to capture the variability in the demand volumes. Tactical plans with different levels of robustness can be obtained by adjusting the input parameter that is used to control the conservative level of the solution. Once the tactical planning model was introduced, six execution policies based on two sets of recourse actions and three levels of flexibility were introduced. The two recourse actions that were considered are outsourcing excess demand to third party companies at a high per cargo unit penalty cost and renting high cost additional vehicle to accommodate the excess demand. Execution policies with various degrees of flexibility were achieved by adjusting the degrees to which the three decisions at tactical level are allowed to change at the execution level. A simulation procedure was then constructed to analyze how the interactions of tactical plans and execution policies affect the performance of the system. Tactical plans with different degrees of robustness and execution policies with three levels of flexibility were simulated for each recourse policies. In each of the three execution policies, which differ in the flexibility level, under the first recourse action of outsourcing excess demand to third party companies, the cumulative distribution functions of the total cost for each tactical plan intersect, suggesting that the optimal level of robustness depends on the risk preference of the decision maker. However, using the expected final cost as the evaluating metric, the result showed that as the flexibility level of the system increases, the optimal robustness level of the tactical plan decreases. The simulation results for the second recourse policy of renting additional high cost vehicle showed that the deterministic tactical plan performs substantially worse than the tactical plans with added robustness. In addition, using the expected total cost as the evaluating criterion, it is interesting to note that increasing the flexibility level doesn't have as much effect on the performance of each of the tactical plan under this recourse policy as compared to the previous recourse policy.

The framework introduced in Chapter 3 was then applied to the military logistics context. Chapter 4 provided background information on the United States Transportation Command (USTRANSCOM), specifically the Air Mobility Command (AMC) and the planning process of the US military aircraft. Chapter 5 then applied the framework introduced in Chapter 3 to a simplified network model motivated by the AMC's scheduling processes. Although there are multiple sources of uncertainty in the system, the study limited the scope to address uncertainty that stems from arrivals of new customer requirements. To accommodate the "all or nothing" nature of the uncertain demand, an uncertainty set that is motivated by the concept of budget of uncertainty (Bertsimas & Sim, 2002) was introduced. The robust formulation for the airlift scheduling problem was then presented. The only recourse action considered in this chapter is outsourcing excess demand to third party companies at certain per unit penalty cost. Three experiments with different levels of uncertainty – measured by the portion of the uncertain requirements to the certain requirements – were then conducted. For each experiment, a simulation was executed to evaluate the performance of the system under tactical plans with various degrees of robustness and execution policies with various degree of flexibility. From the simulation results, several key observations can be made. Using the total expected cost as the evaluating metric, the first observation is that increasing the flexibility level of a system with the fleet size and aircraft routing decisions fixed to a system with only the fleet size decisions fixed has much greater positive impact on the overall performance than increasing the flexibility level of a system with all the decisions fixed to a system with the fleet size and aircraft routes fixed. In addition, as the level of the uncertainty increases, increasing the flexibility level has a greater positive impact on the overall performance of the system.

6.2 Future Research

The study that was presented in Chapter 5 only considered the arrival of new requirements as a source of demand uncertainty. An extension would be to also consider the demand uncertainty that is caused by changes to the old requirements. This may require constructing two separate uncertainty sets to capture the difference in the nature of the two types of demand uncertainty. Furthermore, the study only considered small problem instances and provided a proof of concept for applying the robust optimization to transportation planning. Consequently, a possible future research topic is to come up with decomposition methods to solve large problem instances.

Furthermore, the analysis in this study of how the interaction between the degree of robustness of a tactical plan and the flexibility level of an execution policy affects the performance of the system relied heavily on a major simplifying assumption that customer demands are realized simultaneously at the same time. This simplifying assumption facilitated the two stage decision making process, where tactical decisions are made before and execution decisions are made after the realization of the demand values. In practice, the flow of information is very dynamic and uncertainties are realized sequentially. Accordingly, one area of future research topic is to relax the assumption that all uncertainties within the system are realized at the same time and consider a multi-stage decision making process.

A multi-stage decision making process introduces several interesting research topics. One topic would be to determine the appropriate dynamics of the decision making process that determines when and how often the plan should be adjusted or re-optimized and how much of the plan should be changed each time. For instance, a purely proactive approach may construct a plan to withstand all changes that occur in the planning horizon without requiring any adjustments. On the other extreme, a purely reactive approach may re-optimize a plan every time an uncertain parameter is realized. The challenge would be to determine the right balance between these two approaches for different environment with different uncertainty level and pattern of information flow.

To address such issue, several factors may need to be considered. One of the factors is the threshold that determines if a plan should be modified or left to recover on its own once a disruption occurs. Furthermore, once a decision has been made to modify a plan, another consideration that needs to be taken into account is the question of how much of the plan should be changed. For instance, a minor scale disruption may require no or little modification based on a simple policy but a larger scale disruption may require more corrective efforts. It would be a great challenge to determine when and what corrective actions or rules to apply to a plan for different disruption scenarios. These decisions would also depend on the level of cushion that is present in the plan. As a result, all these decisions - how much cushion to add to the plan, how often the plan should be adjusted, and how much of the plan should be changed - are extremely interdependent. It would be really interesting to investigate more into how these different factors interact with each other and how they affect the performance of the system under different environments.

Bibliography

[1] Ageeva, Y. (2000). Approaches to incorporating robustness into airline scheduling. Master's thesis, Massachusetts Institute of Technology.

[2] Ahmadbeygi, S., Cohn, A., & Lapp, M. (2009). Decreasing airline delay propagation by re-allocating scheduled slack. *IIE Transactions*, 42, 478-489

[3] Baker, S., Palekar, U., Gupta, G., Kale, L., Langer, A., Surina, M., & Venkataraman, R. (2012). *Parallel Computing for Dod Airlift Allocation*. MITRE Technical Report.

[4] Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, 25, 1-13.

[5] Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769-805

[6] Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99, 351-376.

[7] Bertsimas, D., & Sim, M. (2002). The price of robustness. *Informs*, 52(1), 35-53.

[8] Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming*, 98, 49-71.

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

[10] Crainic, T. (1998). *A survey of optimization models for long-haul freight transportation.* Centre for Research on Transportation = Centre de recherche sur les transports (CRT).

[11] Crainic, T.G. (2000). Service network design in freight transportation. *European Journal of Operational Research,* 122, 272-288.

[12] Erera, A., Hewitt, M., Savelsbergh, M., & Zhang, Y. (2012). Improved load plan design through integer programming based local search. *Transportation Science*.

[13] Hoff, A., Lium, A.G., Lokketangen, A., & Crainic, T.G. (2010) A metaheuristic for stochastic service network design. *Journal of Heuristics*, 16(5), 653-679.

[14] Kang, L.S. (2004). Degradable Airline Scheduling: An Approach to Improve Operational Robustness and Differentiate Service Quality. Doctoral thesis, Massachusetts Institute of Technology.

[15] Koepke, C. G. (2004). Multi-mission optimized re-planning in air mobility command's channel route execution. Master's thesis, Massachusetts Institute of Technology.

[16] Koepke, C. G., Armacost, A. P., Barnhart, C., & Kolitz, S. E. (2008). An integer programming approach to support the US Air Force's air mobility network. *Computers & Operations Research*, *35*(6), 1771-1788.

[17] Kouvelis, P., & Yu, G. (1997). *Robust discrete optimization and its application.* Norwell, MA: Kluwer Academic Publishers.

[18] Lan, S., Clarke, J.P., & Barnhart C. (2006). Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Science*, 40(1), 15-28.

[19] List, G.F., Wood, B., Nozick, L.K., Turnquist, M.A., Jones, D.A., Kjeldgaard, E.A., & Lawton, C.R. (2003). Robust optimization for fleet planning under uncertainty. *Transportation Research Part E*, 39, 209-227.

[20] Lium, A. G., Crainic, T. G., & Wallace, S. W. (2009). A study of demand stochasticity in service network design. *Transportation Science*, *43*(2), 144-157.

[21] Morton, D.P., Salmeron, J., & Wood, R.K. (2002). A stochastic program for optimizing military sealift subject to attack. *Stochastic Programming e-Print Series*, http://www.speps.info

[22] Mudchanatongsuk, S., Ordonez, F., & Liu, J. (2008). Robust solutions for network design under transportation cost and demand uncertainty. *Operational Research Society*, 59(5), 652-662.

[23] Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations research*, *43*(2), 264-281.

[24] Nielsen, C. A., Armacost, A. P., Barnhart, C., & Kolitz, S. E. (2004). Network design formulations for scheduling US Air Force channel route missions. *Mathematical and Computer Modeling*, *39*(6), 925-943.

[25] Rosenberger, J., Johnson, E.L., & Nemhauser, G.L. (2004). A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles. *Transportation Science*, 38, 357–368.

[26] Smith S. (2009). *CVAO Optimization Survey: An Analysis of the US Transportation Command's Joint Deployment and Distribution Enterprise from an Optimization Perspective*. Report for Air force Research Laboratory, The Robotics Institute, Carnegie Mellon University.

[27] Shebalov, S., & Klabjan, D. (2006). Robust Airline Crew Pairing: Move-up Crews. *Transportation Science*, 40, 300–312.

[28] Smith, B.C., & Johnson, E.L. (2006). Robust Airline Fleet Assignment: Imposing Station Purity Using Station Decomposition. *Transportation Science*, 40, 497-516

[29] Smith, S.F., Becker, M.A., & Kramer, L.M. (2004). Continuous management of airlift and tanker resources: A constraint-based approach. *Mathematical and Computer Modeling: An International Journal*, 39(6-8), 581-598.

[30] Soyster, A.L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21, 1154-1157.

[31] Wu, T.T., Powell, W.B., & Whisman, A. (2009). The optimizing-simulator: An illustration using the military airlift problem. *ACM Transactions on Modeling and Computer Simulation*, 19(3), 14.