University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

12-2013

Optimization models for resource allocation under perturbation.

Dongxue Bridgeman 1986-
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

OPTIMIZATION MODELS FOR RESOURCE ALLOCATION
UNDER PERTURBATION

By

Dongxue Bridgeman

B.E., Beijing University of Posts and Telecommunications, 2008

M.S., University of Louisville, 2009

A Dissertation

Submitted to the Faculty of the

J. B. Speed School of Engineering of the University of Louisville

in Partial Fulfillment of the Requirements for the

Doctor of Philosophy

Department of Industrial Engineering

University of Louisville

Louisville, Kentucky

December 2013

OPTIMIZATION MODELS FOR RESOURCE ALLOCATION UNDER PERTURBATION

By

Dongxue Bridgeman

B.E., Beijing University of Posts and Telecommunications, 2008

M.S., University of Louisville, 2009

A Dissertation Approved on

July 26th, 2013

by the following Dissertation Committee:

Dissertation Director: Dr. Lijian Chen

Dr. Sunderesh Heragu

Dr. Gerald Evans

Dr. Lihui Bai

Dr. Anup Kumar

DEDICATION

This dissertation is dedicated to my parents

Mr. Jing Ma

and

Mrs. Lihuang Yang

who have given me invaluable educational opportunities.

ACKNOWLEDGMENTS

ABSTRACT

OPTIMIZATION MODELS FOR RESOURCE ALLOCATION UNDER

PERTURBATION

Dongxue Bridgeman

July 26th, 2013

Optimization Models for Resource allocation are investing in how to make the best use of available but limited resources in order to achieve the best results. In strategic planning, resource allocation is a plan for using available resources, especially in the near future, to achieve the goals of the future. It is a process of allocating resources during the entire planning horizon and among the various units. Resource allocation plans can be decided by using mathematical programming. In this dissertation, the research has been focused on how to allocate resources in the uncertain environment. The mathematical programming formulations for the resource allocation model under severe uncertainty will be studied. In particular, we will focus on solving the stability issues of the traditional probabilistic model. We propose an approach consisting of solving a sequence of convex robust optimization models with unknown-but-bounded random variables along with the stochastic programming to pursue the allocation performance for the expected overall objective value. Our theoretical results show that the proposed approach can always obtain an equivalent or a better expected revenue with the corresponding allocation, while significantly reducing the risk under perturbations. Although this method requires solving two convex mathematical

programming models, both models are solved within a timely manner thanks to their convex model instances and with effective, and less, computationally demanding algorithms.

With the increasing threats from public health emergencies, such as earthquakes, tornados, pandemic flus, or terrorist attacks, high attention has been paid to the public health response to a pandemic from federal to national level, together with local health departments, and the health-care community. Various organizations cooperate with each other to strengthen the preparedness for the pandemic and disastrous emergencies, thus to improve the public health. The Strategic National Stockpile (SNS) is maintained by the Centers for Disease Control and Prevention (CDC) and the U.S. Department of Health and Human Services (DHHS) for the United States in the event of a shortage of local medical resources or other unanticipated supply problems. The SNS is the United States' national repository of antibiotics, vaccines, chemical antidotes, antitoxins, and other critical medical equipment and supplies. It has the capability to supplement or re-supply local health authorities with necessary materials for relief action within the response time in as little as 12 hours. The pilot study is done with the support of Kentucky SNS to determine the capacity allocation plan for each county in order to maximize the health benefit under various uncertainties, which can never be accurately estimated. We thereby employ a heuristic method named "resource reservation" to suggest the resource allocation plan for Kentucky SNS.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Real-Time Decision-Support System for Health Care and Public Health Protection Project

The research work presented in this dissertation is funded and part of the real-time, decision-support system (RTDSS) for health care and public health protection project. The goal of this project involves developing a RTDSS to help personnel in the health-care and public health (HPH), as well as emergency services sectors (ESS), make real-time decisions relative to allocation and re-allocation of scarce resources in the aftermath of a pandemic influenza or other viral attack.

Dr. Heragu, as the Project Director, says "some studies indicate that up to 40% of the population could be stricken and hospitals could be operating at 50% of their capacity during a pandemic attack. During a time of medical surge when we really need HPH as well as ESS personnel and equipment to be operating at full capacity, the challenge for planners is to allocate the few resources at their disposal in the most efficient manner in response to fast changing conditions on the ground so that a large number of people can be served in as short a time as possible."

1.2. The Problem

The idea of the study is the applying scientific approach to improve the decision making accuracy in order to better prepare for public health emergencies, such as pandemic, especially using scientific approaches to determine how best to design and operate a system, usually under conditions requiring the allocation of scarce resources. Large-scale emergencies are of high-consequence, low-probability (HCLP) events such as H1N1 pandemic that may result in loss of life and severe property damage. In recent years, developing decision-oriented operations research models to improve preparation for and response to major emergencies has drawn more and more attention.

We consider "emergency response" as response to catastrophic and disaster events and do not include daily response of ambulance, police, or fire departments to routine emergency calls. Large quantities of vaccines or other medical supplies are stored in strategic locations around the nation. In the event of a large-scale public health emergency that causes local supplies to run out, these assets will be used to supplement the state and local public health departments. The decision to deploy SNS assets is a collaborative process decided among local, state, and federal officials. It starts at the local level when officials identify a potential or actual problem that they believe will threaten the health of their community.

Traditional public health resource allocation practices and processes, as currently employed for resource allocation during public health emergencies, are based on the project area's population size. As an example, if six million doses total (three million doses of nasal spray vaccine and three million doses of injectable vaccine) are ready for ordering nationally (as of today) and a state has 10 % of the U.S. population, then their allocation for today is 600,000 doses total (300,000 doses of the nasal spray vaccine and 300,000 doses

of injectable vaccine).

Kentucky SNS administrators and strategic planners face the challenge of determining the number of doses of vaccines that will be distributed. To date, there is no tool to help them determine the vaccines distribution under the uncertain scenario. Currently, the resource allocation policy adopted by Kentucky SNS and other local SNS is to distribute by the demands. If the resources on hand are not enough to satisfy the demands, they will distribute the available resources by the population sizes of each required area. Under severe uncertain scenario, the real demands may be very different from the predication. Allocating the scarce resources by the current policy will cause serious losses on the public health.

The basic question of medical resource allocation is how to allocate the right amount of resources to the right patients at the right time to better utilize the available resources, maximize the satisfaction of the public, and minimize the loss of public health. The problem that this research work aims to solve seems really simple, but the allocation decision is hard to make because of the incomplete information regarding the future demand, especially during the pandemic outbreaks. The demands of resources closely depends upon the number of people that are infected or are going to be infected. The number of people who become infected also depends on the strain of virus, the severity of disease caused by the virus, the vulnerability of affected populations, and many other factors. The overall effect of various factors make the prediction of the future demands of vaccines extremely hard.

We propose to apply a mathematical modeling method to improve the accuracy of traditional decision making strategy by taking into consideration the uncertainty of the current estimate of resource demands, and future demands, in order to determine the minimum

needs to prepare for distribution, and to reserve the rest of resource for future use in emergency situations, so that the overall impact of the pandemic influenza can be minimized. Our team also designed and developed a web tool to allow public-health emergency coordinators to easily utilize mathematical models and make vaccine allocation decision.

One basic way to solve the uncertainty problem is to take stochastic demands into account through expected values, thus yielding a deterministic program that can be easily solved. However, the drawback of such an approach is obvious, as it ignores any distributional information about the demands. A common way to attempt to overcome that issue is to repeatedly solve the linear programming during the planning horizon. While such an approach may seem intuitive, it turns out that this strategy may actually backfire. In Cooper (2002), the author shows a counter example where this re-solving linear programming may lead to lower total revenue.

An alternative way to incorporate demand into the model is by formulating stochastic programming models. The stochastic programming models are fundamentally based on the assumption that the random variables' distributional information is known. However, the data are more likely to be perturbed, e.g. incomplete, erroneous, or wrongly estimated data for stochastic models. Although the perturbation could be minor without significantly changing the nominal values of the parameter, the impacts on the optimal solution are not trivial. For example, in Nemirovski et al. (2009), the authors show that a perturbation of 0.1 % - accurate could lead to a violation of up to 450% of the nominal values. This observation is very counter intuitive because the quite small perturbation can make the obtained optimal solution meaningless. Perturbation is not rare because new strains will always come out and any prediction about the demands of a vaccine against a new strain can not be made based on old strains with ease, especially in prediction of vaccine demands during

pandemic outbreaks. In addition, the historical data can not always be available, since there are a lot of factors that will impact the demands of resources. Therefore, this study works to resolve the uncertainties, such as the application of scarce resource allocation during pandemic outbreaks, and will utilize advanced mathematical models to generate reliable allocation plans that optimized for various scenarios.

In this dissertation, typical stochastic programming methods will be investigated along with the techniques for solving these models. Particularly, a unique combined optimization method named "Resource Reservation" will be proposed, which is designed for the situations where one or more parameters are severely uncertain, the distributional pattern of uncertain parameters are not available, or the ranges of uncertain parameters are known. This proposed method is applied to solve the Kentucky SNS vaccines allocation problem under uncertain scenarios since vaccines are the most important resources to be allocated and distributed during the pandemic outbreaks. An intuitive interactive web tool has been developed and implemented and is available to use 24/7. This web tool is designed to help public health coordinators determine the way to distribute vaccines under uncertainty and minimize the consequences of a vaccine shortage. The potential health consequences that may occur under various vaccine shortages can be the number of confirmed cases, hospitalizations, and deaths. In this application, we assume that the consequence of a vaccine shortage is evaluated by the expected unmet need of a vacancies.

1.3. Key Contribution

The contributions of this research work can be summarized as follows.

- Our proposed resource reservation heuristic has a wide application, especially for the complicated optimization problems involving uncertain parameters. In this

research work, we apply our methodology to solve the Kentucky SNS vaccine allocation problem, which helps the public health emergency managers or coordinators make better decision on how to allocate scarce resources during the pandemics, or other types of public health emergencies.

- Only a few analytical models in the literature include the discussion of the effect of parameter of uncertainty on the vaccination policies. Some researchers apply the stochastic programming method with chance constraints to determine the vaccination policies, which requires the complete information of the epidemiological characteristics of the virus. However, in reality, the epidemiological information is not necessarily available before the allocation decision has to be made. Most of the time, only the rough estimation of the demands will be known when the public health emergency occurs. Our proposed method does not require the complete epidemiological information or any distributional information to be exactly known. It is designed for the situation that limited information of disease is known and the estimation of the demands is under severe uncertainty. Since the models are designed for the severe uncertainty situation, the resource reservation heuristic will first determine the "absolutely necessary amount" of resources to release, and if there are resources left, then it will be reserved for the future use.

- The web tool developed by the RTDSS team is convenient for users to utilize the models without any mathematical background and industrial engineering knowledge. The user only needs to provide a few data points to make the model work, such as the amount of available resources, number of new cases, number of cased identified, and period of times, number of time periods during the whole planning horizon, and other general information associated with the outbreaks. The user can also select the projected counties, and then the allocation decision will only be made among those projected counties. The user can also manually reserve the

available doses of resources by selecting "risk level" in the user input interface. In addition, all the default value that we use in the model is changeable by the administers.

- The model can also help the public health emergency managers or administers determine the number of distribution that the current available resources are good for. For example, in the later part of the dissertation, you will see there are no resource to be allocated for some multi-time periods allocation. In this case, the results suggest the number of time periods that the users pre-determined is too high. Based on the results they get, the user can find out how many time periods that the available resources can be best utilized and distributed.

- In the dissertation, we also investigate the effect of an uncertainty parameter "θ" on the conservativeness of the resource allocation plan with an example. Changing this parameter allows the users to adjust the conservativeness of the allocation plan. But there is no certain linear relationship between the conservativeness and the value of "θ" and the feasible range of "θ" varies as the number of selected counties changes. Therefore, we determine to use 1 as the default value of "θ" since it works no matter how many counties that the users select.

- Our method perfectly preserves convexity and overcomes the stability issue of the stochastic programming, therefore it can be efficiently solved by modern interior point solvers on average desktop computers in milliseconds.

1.4. Background on Resource Allocation

Resource allocation is a major problem in many diversified areas, such as portfolio selection, airline seat allocation, network logistics and distribution, inventory management and so on. The best way to allocate resources depends on the nature of the resources, the constraints at hand and the organizations mission. Since the nature of the resource allocation

problem is to seek the best solution among alternatives, quantitative methods like optimization is widely applied to solve the resource allocation problems.

Optimization, in the sense of mathematics, computational science, or management science, refers to the selection of a best solution from some available alternatives. It may use the mathematical expression of a problem to maximize or minimize some function. The alternatives are frequently restricted by constraints on the values of the variables. A simple example might be to find the most efficient transport pattern to carry commodities from the point of supply to the markets, given the volumes of production and demand, together with unit transport costs. Optimization involves designing a system or process to be as good as possible in some defined sense. The defined sense varies from organization to organization. What is optimal for you - with your goals and values - could be suboptimal for the next person. Every performance management paradigm, every mission statement, could point to a different definition of success and therefore to a different way to "optimally allocate resources. A resource optimization model should be based on the objective, decision variables, and constraints. For a certain problem, there may exist particular constraints or restrictions affecting how to properly assign available capacity of resources to achieve the best yield in general. There are plenty of factors affecting the effectiveness of the resource allocation plan, such as the uncertainty of some parameters, the goal that each resource allocation problem is seeking, how it will take to get a solution ,and etc.

Mathematical optimization is one of the most efficient approaches to evaluate and improve the efficiency of resource allocation in the public health emergency event. Mathematical optimization models can explicitly represent the functioning of the system, generally resulting in large linear/nonlinear and integer models with significant numbers of variables and constraints. Optimization applications in public health emergency events address a variety

of issues, such as the cost-effectiveness of selected interventions, and portfolio analysis for policy/decision making besides the efficient use of resources. Public health emergency events are complex, often involving a variety of factors that influence their efficiency. To make these events even more complex is the uncertainty of the factor and incompleteness of the information. Due to the stochastic nature of factors that influence the performance of public health emergency events, the uncertainty based mathematical optimization model may be more suitable for the purpose of representation and analysis.

There are many subfields in optimization in general. In this dissertation, stochastic programming, and robust programming, will be mainly investigated in order to solve resource allocation problems in the public health emergency events. Generally, the stochastic programming studies the case in which some of the constraints, or parameters, depend on random variables. Robust programming is, like stochastic programming, an attempt to capture uncertainty in the data underlying the optimization problem. This is not done by the use of random variables, but instead, the problem is solved taking into account inaccuracies in the input data.

1.5. Background on Strategic National Stockpile

With the increasing threats from large-scale emergencies such as earthquakes, tornados, pandemic flus, terrorist attacks, and so on, emergency supplies units have been set up around the United States to provide necessary materials for relief actions to the affected areas during these large-scale emergencies. One of the units is the Strategic National Stockpile (SNS). Take the pandemic flu as an example: according to the World Health Organization (WHO), influenza rapidly spreads around the world in seasonal epidemics and imposes a considerable economic burden in the form of hospital and other health

care costs and lost productivity. In the United States of America, recent estimates put the cost of influenza epidemics to the economy at U.S. $ 71-167 billion per year. In annual influenza epidemics, 5-15% of the population are affected with upper respiratory tract infections. Hospitalization and deaths mainly occur in high-risk groups, such as the elderly and chronically ill. Although difficult to assess, these annual epidemics are thought to result in between three and five million cases of severe illness and between 250,000 and 50,000 deaths every year around the world.

According to the information on CDC's website, the strategic national stockpiles, as a physical form of surplus capacity, have been created to supplement or re-supply state and local public health agencies in a event of a national emergency anywhere and at any time within the U.S. or its territories. CDC and DHHS is congressionally mandated to maintain the stockpiles, which include vaccines, antitoxins, ventilators, and 12-hour Push Packs.

For fast and flexible response, the SNS assets can be categorized into two types: 12-Hour Push Packs and vendor or SNS, managed inventory. Around 6% of SNS assets are 12-Hour Push Packages which are located strategically around the United States. Each package contains:

- Large amount of prepackaged, individual 10-day regiments for over 300,000 people
- Intravenous drugs and supplies for administration
- Chemical antidotes and related supplies
- Airway management supplies
- Medical/Surgical supplies

These packages can be delivered anywhere in the United States within 12 hours of a federal order to deploy. 4% of the SNS assets are vendor managed, which is government owned

inventory held at vendor warehouses. 90% of the SNS assets are SNS managed inventory, which is owned by government and held in warehouses controlled by SNS. Those vendor or SNS managed inventories mainly consists of vaccines, antiviral drugs, and antitoxin.

The decision to deploy the SNS assets is a collaborative process decided among local, state, and federal officials. The decision starts at the local level, when officials identify a potential or actual problem that they believe will threat the health of their community. The CDC's Emergency Operations Center (EOC) will arrange a telephone conference call that will include representatives from DHHS, SNS, and the requesting state's governor and representatives. The affected state will request the SNS material from the CDC as soon as state officials, in consultations with local officials, determine it is necessary to do so to protect the health of the public. The SNS assets can be requested alone, or as part of an overall request for federal assistance, through the National Emergency Response System.

The requests for the SNS assets, from local to state, can be made through WebEOC, which is a web-enabled incident management tool used to coordinate the use of state assets and provide a direct link from the State EOC to local EOCs. The system is used by all state agencies and local emergency managers to report, track, and respond to incident reports during an emergency, disaster, or catastrophe. Data is entered and viewed on various status boards via WebEOC allowing all authorized federal, state, local, and private agencies simultaneous access. To be more concise, the sequence of events for the request can be summarized as follows:

- the state determines the need for SNS assets
- the governor's office is notified
- the state calls the CDC Emergency Operations Center
- the CDC arranges a telephone conference call

- the governor makes the formal request during the conference call

Once the request is approved for the SNS materials to be deployed, the CDC will send the assets to a location in the requesting state. Once the SNS assets arrive in a state, they are signed for by a state representative where they become the responsibility of the state, and the received SNS material will be stored in a warehouse until it is picked up and sent out to the counties for dispensing. The SNS materials will be allocated, or distributed, to the counties by request, depending on the event. Staff and volunteers in the RSS will pick and palletize orders to be sent to one local health department and one hospital in each county. The local health departments in the state are responsible for dispensing the pharmaceuticals to the population. In a statewide event, the first round of medications from the SNS will be allocated by population. Once significant amounts of medications arrive from Managed Inventory. Health departments and hospitals will have the opportunity to request additional medications and supplies through WebEOC, or by fax.

There are twelve state level SNS sites across the United States which are all managed by the federal authorities; in the mean while, there are also three local level organizations that cooperate with the SNS, which are the RSSs and PODs. The two candidate sites serving as warehouses in the state of Kentucky are responsible for receiving, staging, and storing those medical supplies. These are called RSSs. There are fifteen health care planning regions and fifty health department districts in Kentucky. The fifty health department districts are serving one hundred and twenty counties in Kentucky. The fifty health department districts are defined as Regional Distribution Nodes (RDNs). Additionally, there are two hundred and fifty points of dispense (PODs) locations to cover all the population in Kentucky. Each pod is responsible for certain areas which are categorized by the geographical distribution. From the introduction above, it is a multi-stage distribution system that consists of one federal strategic national stockpiles, two RSSs, fifty RDNs and two hundred and fifty PODs.

The research work in this dissertation is mainly focused and designed for the Kentucky SNS. The models and platform that we design for the Kentucky SNS can be applied to other states only with minor changes, since a great level of similarity in operations is shared by the states.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews the applications of work in operations research in recent years which is devoted to health care resource allocation. This chapter is separated into three sessions. The first section summaries the historical research work that has been done by OR/MS researchers on typical health-care system issues with applicable methodologies and application provided for illustration. The second chapter reviews what other researchers have done to solve vaccine allocation problems during pandemics. Last but not least, as mentioned in Chapter 1, uncertainty will be considered in decision-making on vaccine allocation, so a review of literatures on various types of optimization models with uncertain parameters will also be presented.

2.1. Applications of Operations Research in Health-care Delivery

This session will summarize the historical research work that have been done by many operations researchers. It will cover the categorization of typical health-care system issues that OR/MS research is suitable for, and provide applicable OR/MS methodologies and application for each topic. Pierskalla and Brailer (1994) reviews the body of work in operations research in health-care delivery before 1994. Important reservations have been pointed out about the characteristics of health-care system. Although many problems faced by operation researchers in health-care are not inartistically different from problems in other industries, many other are quite unique due to certain characteristics of health care delivery systems. Some of these are the possibilities of death or low quality of remaining life, the difficulty in measuring quality and value of outcomes, the sharing of decisions

among several decision makers (physicians, nurse and administrators), third party payment mechanisms for diagnoses and treatments, and the concept of health-care access as a right of citizens in society. Additionally, Pierskalla and Brailer (1994) summarize and separate health-care problems into three general categories, where the knowledge of OR/MS can be used. First is the system design and planning, which deals with large allocation decisions, both at the policy level and at the operational level. Second is management of operations, which examines monitoring and control methodologies at the operational and tactical levels. Third is medical management, which involves patient disease detection and treatment at the policy, and the patient levels. Several main topics that each category frequently deal with are also provided below, along with applicable OR/MS methodologies and applications.

(1) System Planning and Strategy

(a) Planning and strategy: typical problems of planning and strategy are optimal clustering problems (the decision maker attempts to partition the set service so that some objective is optimized) or resource allocation problems (the central authority must plan for and allocate scarce resources to regions or districts).

In the early eighties, operations research has been already applied to solve system design and planning problems in different aspects of health-care fields. At the national level, Rizakou et al. (1991) developed a spreadsheet model based decision support system, and AIDSPlAN to plan for the resources needed for HIV/AIDS related services in the United Kingdom. The model incorporated demand forecasting by patient categories, care protocols and resource and budget needs for central and local planners in the British National Health Service. It can also be used with the planning efforts of earlier math

programming balance of care (BOC) models incorporated in micro-computer software to determine resource needs and allocations for entire health service regions and districts.

Pezzella et al. (1981) provides an example of health service districting. In this paper, the authors developed a model by which local health departments in Italy are assigned into regional health structures. Dimensions of the analysis include an analysis of the demand for health services based on demographic, socio-economic, and geographical information, interviews with experts, surveys of special disease populations, and an analysis of available hospital service, which are considered in the proposed mathematical model for optimal districting. The authors note previous attempts to perform such rationalization by set partitioning, generalized assignment models, location and allocation models, and other linear programs. The authors use a linear program which has two objectives: to minimize the average distance of individuals from the nearest center, which improves access, and to minimize the deviation between proposed and existing districting, which improves poetical acceptability.

In our RTDSS project, our team also designed a few applications regarding the system planning. For example, the system planning model in the RTDSS web tool, is used to allocate critical medical or surgical resources among the Kentucky hospital network. Another application called the patient allocation model is developed to improve the patient acceptability depending on the distance and the availability of resources.

(b) Demand forecasting: aggregate demand forecasting and daily prediction is important in improving the efficiency of capital resource use in health-care. It is a fundamental input to many other analysis in health-care operations research. Health-care operations research tends to employ existing and well known operations research algorithms from other industrials to perform forecasting.

Qualitative approaches such as historical analysis are frequently used in health-care industry. This technique, while inexpensive and easy to use, ignores environmental changes and does not support the analysis of new venture decisions. Another qualitative approach Delphi technique is also used in demand forecasting. Future predictions are extracted from experts and the process is repeated until a consensus emerges. The Dephi technique is susceptible to ideological biases which make it difficult to use in settings where predictions may not conform to standard views of system operation. For more details on the review and application of these techniques, please refer to Harrington (1977).

There are many other qualitative techniques that have been employed in health-care service forecasting as well. For example, least squares regression analysis has been used to determine the demand for pre-hospital care in order to make ambulance staffing decisions in Kamenetzky et al. (1982). Kao and Tung (1980) employ an auto-regressive, integrated moving average (ARIMA) time series model to forecast demand for inpatient services.

(c) Location selection: Location of health-care facilities and services has much in common with the location aspects of many types of facilities or services which have a geographically dispersed customer base, and where there is a need to be close enough to customers for ease of access and/or speed of access, as well as a need for low cost of sitting and operations. The difference of health-care facilities and services location problem is health-care facilities and services may be subject to public control laws and in the case of emergencies, vehicle locations where there are maximum response time requirements, may need to balance closeness to customers and facilities. The SNS logistics design model in RTDSS web tool is designed to determine the location of RRSs, RDNS, and PODs.

(d) Capacity planning: this topic relates to those decisions regarding the appropriate levels, equipment and personnel for some demand. In health-care, capacity planning usually focuses on decisions such as total bed capacity, surgical system capacity, bed capacity allocation to different services, capital equipment capacities, ancillary service capacity, and factors with which affect capacity utilization such as patient flow, staffing levels, and staff skill mix. The common goal of capacity planning problems is to seek ways to increase the productivity of existing assets and improve service quality. Many of the studies drive the capacity planning process with an underlying queueing system and use simulation to obtain solutions.

(2) Management of operations

(a) Management information systems (MIS): MIS's have the potential to improve decision-making responsiveness, quality of patient care, and productivity. Quality of care improvements can result from reduced waiting time for

physicians' orders to be carried out and results to become available, elimination of unnecessary services, reduction of errors, and increased provider and patient satisfaction. Productivity is improved by appropriate data availability for decisions about staffing patterns, scheduling, use of equipment and materials, and by elimination of redundant systems and actions. More details about the benefits of MIS in health-care can be found in Moidu and Wigertz (1989).

(b) Patient Scheduling: Patient Scheduling can be a very effective way to match the demand with the supply of services available. The objective of most patient scheduling problems is to optimize the satisfaction of patients and physicians as well as the utilization of facilities. Many different methodologies are proposed involving queueing models as represented by Markov and semi-Markov processes, mathematical programming, heuristic and expert systems, and simulation.

(c) Workforce planning and scheduling: The management of human resources is a major activity in health-care organization. Like many service organizations, the ability to match staffing resources to a fluctuating demand directly affects operating efficiency and the quality of service. The development of innovative approaches to the organization and management of nursing and other human approaches to the organization and management of nursing and other human resources holds great promise for further cost savings in the delivery of health services.

(3) Medical management

(a) Screening for diseases, which are tests which detect diseases, have resulted in advances in medical diagnosis and disease detection. Individual screening and mass screening are different modeling problems because the objectives of the decision makers frequently differ and the constraints and parameters affecting the decisions may vary. Operations researchers have been attracted to this area due to its importance. Current work focuses primarily on descriptive modeling of disease processes, progression, and causal factors. In the case of contagious diseases, very complex simultaneous differential equation and/or statistical models are frequently used to describe the growth, maturity, and decline of various specific or general types of diseases. Another branch of research focuses on maintainability. The applications include a large variety of decision models such as mathematical programming, Markov decision, and simulation. However, more work is needed on the linkages to the epidemiology literature and to the construction and evaluation of models which can aid in decisions that are often based on sparse or incomplete data and incomplete knowledge of the disease etiology and propagation.

(b) Clinical decision-making:In this area, decision analysis can be used to aid in the structuring of medical decisions. Performance improvement can also be used to address the concerns about quality of care and practice efficiency. Last but not least, the use of operations research techniques and cost effective analysis (CEA) can be used to analyze health-care policies and those policies that affect large populations.

So far, we have reviewed the major categories of health-care problems and many operations researchers' contribution towards solving problems by the use of knowledge of operations research. Since the purpose of this dissertation is to explore the resource allocation

and/or reallocation for large-scale emergency response, the next session will more focus on operations research methodologies and applications to address resource allocation and/or reallocation for large-scale emergency.

2.2. Vaccination Strategies under Parameter Uncertainty for Emergency Response

Vaccination is one of the primary strategies used by public health authorities to control human infectious disease. The major resources to be allocated and/or reallocated in the SNS are vaccines. So we have decided to make solving vaccine allocation for pandemic problem as our primary goal. In this section, we will review a series of recent research work that have been done by other OR/MS researchers on the similar issue, which can definitely help us deepen our understanding of the characteristics of our problem and thus provide more effective solutions.

Mathematical models have long played a major role in identifying and evaluating strategies in order to allocate resources to guarantee maximum effectiveness of vaccination in controlling infectious disease outbreaks. There are three primary modeling approaches that have been used in this effort - deterministic analytical models, stochastic analytical models, and computer simulations.

Very few of these analytical models include discussion of the effect of parameter of uncertainty on the vaccination policies identified and/or evaluated but this uncertainty can have major consequences. For example, Tanner et al. (2008) presents a stochastic programming framework for finding the optimal vaccination policy for controlling infectious disease epidemics under parameter under uncertainty. They formulated the problem to seek the minimum cost vaccination policy under a chance-constraint. The chance-constraint requires the

21

probability that $R_x \leq 1$ be greater than some parameter α, where R_x is the post-vaccination reproduction number. They also show how to formulate the problem in two additional cases: a) Finding the optimal vaccination policy when vaccination supply is limited and b) a cost benefit scenario. The class of epidemic models for which this method can be used is also described. Epidemiological models are often formulated as a series of compartments corresponding to different disease states, e.g.: susceptible, exposed, infectious, recovered, etc. The models are then referred to by the series of capital letters that corresponds to the compartments within the basic model structure. For example, an SIR model considers individuals to be either suspectable (S), infectious(I), or recovered(R), and to progress through the stages in that order;

Longini Jr et al. (1978) shows that the optimal allocation of vaccines derived from their influenza model is highly sensitive to both the epidemiological characteristics of the virus and to the choice of the objective function used in the optimization process. Similar conclusions have been drawn by Bansal et al. (2006), who finds that optimal strategy depends critically on the viral transmission level (reproductive rate) of the virus. They present a comparative analysis of two classes of suggested vaccination strategies: mortality-based strategies that target high-risk populations and morbidity-based strategies that target high-prevalence populations. Applying the methods of contact network epidemiology to a model of disease transmission in a large urban population, they assume that vaccine supplies are limited and then evaluate the efficacy of these strategies across a wide range of viral transmission rates and for two different age-specific mortality distributions.

Dushoff et al. (2007) developed a simple mathematical model of an epidemic that includes assortative mixing between groups of hosts. They evaluate the impact of different vaccine allocation strategies across a wide range of parameter values. With this model they

demonstrate that the optimal vaccination strategy is extremely sensitive to the associativity of population mixing, as well as the reproductive number of the disease in each group.

Clancy and Green (2007) use a Bayesian-decision theoretic approach and a general stochastic SIR (susceptible → infective → removed) model for the spread of infection through a closed population under parameter uncertainty. They also mention that in reality, uncertainty over parameter values exists. So they consider the effect upon the optimal policy of changes in parameter estimates, and of explicitly taking into account parameter uncertainty via a Bayesian decision-theoretic framework.

2.3. Introduction of Optimization Models under Uncertainty

From the review in 2.2, it is very important to take into account of the uncertainty in vaccine allocation because the optimal solution is extremely sensitive to the epidemiological characteristics of the virus. Why data uncertainty deserves special treatment can be answered by this example. The optimal solutions of linear programming problems may become severely infeasible if the nominal data is slightly changed. This phenomenon has been demonstrated by studying `PILLOT4` from the well-known `NETLIB` collection, see Ben-Tal and Nemirovski (2000). It is a linear programming with 1000 variables and 410 constraints; One of the

constraints (#372) is

$$a^T x = 15.79081 x_{826} - 8.598819 x_{827} - 1.88789 x_{828} 1.362417 x_{829} \qquad (2.1)$$

$$1.526049 x_{830} 0.031883 x_{849} 28.725555 x_{850} 10.792065 x_{851}$$

$$0.19004 x_{852} 2.757176 x_{853} 12.290832 x_{854} + 717.562256 x_{855}$$

$$0.057865 x_{856} 3.785417 x_{857} 78.30661 x_{858} 122.163055 x_{859}$$

$$6.46609 x_{860} 0.48371 x_{861} 0.615264 x_{862} 1.353783 x_{863}$$

$$84.644257 x_{864} 122.459045 x_{865} 43.15593 x_{866} 1.712592 x_{870}$$

$$0.401597 x_{871} + x_{880} 0.946049 x_{898} 0.946049 x_{916} \geq b = 23.387405$$

The related nonzero coordinates of the optimal solutions x^* of the problem, as reported by CPLEX , are as follows:

$x^*_{826} = 255.6112787181108$ $x^*_{827} = 6240.488912232100$ $x^*_{828} = 3624.613324098961$

$x^*_{829} = 18.20205065283259$ $x^*_{849} = 174397.0389573037$ $x^*_{870} = 14250.00176680900$

$x^*_{871} = 25910.00731692178$ $x^*_{880} = 104958.3199274139$

Most of the coefficients in (2.1) are "ugly reals". It is natural to assume that the " ugly coefficients" are in fact uncertain. Ben-Tal and Nemirovski (2000) proved that quite small (just 0.1%) perturbations of "obviously uncertain" data coefficients can make the "normal" optimal solution x^* heavily infeasible and thus practically meaningless.

The data of real-world optimization problems most likely are uncertain. The sources of uncertainty may be from measurement errors or estimation errors. In real-world applications of optimization, one cannot ignore the possibility that even a small uncertainty in the data can make the nominal optimal solution to the problem completely meaningless from

a practical viewpoint. Therefore, in optimization, there exists a real need of a methodology capable of dealing with uncertain data. Two major mathematical methods, stochastic programming and robust optimization, can address the data uncertainty issues will be introduced in this chapter. In addition, we propose a novel methodology that is a combination of two major methods, which can address the optimization problem with uncertain data overcoming the drawbacks of each method. We apply our proposed methodology on the vaccine allocation optimization during the pandemic. Vaccine allocation policy assuming the epidemiological characteristics are not accurate or uncertain, is a very possible scenario that could happen when any pandemic disease occurs. Our proposed methodology is not only good on vaccine allocation optimization but also applicable on other complicated uncertainty optimization problems with similar characteristics. This chapter reviews the theory and methodology that have been developed to cope with the complexity of optimization problems under uncertainty. We summarize the evolution of various techniques used to solve problems under uncertainty, discuss and contrast the advantages and disadvantages of the classical recourse-based stochastic programming, and robust optimization. The disadvantages of each method are the motivation for us in designing our proposed methodology.

2.3.1. Stochastic programming

In mathematical programming, there are a variety of ways to deal with the problems of uncertainty. The data in the objective $f(x)$ and in the convex constraints $g_i(x), i = 1, \ldots, m$ associated with a convex programming problem:

$$\min\{f(x)|g_i(x) \leq 0, i = 1, \ldots, m\} \tag{2.2}$$

are uncertain in most real problems. The easy solution is to replace the "uncertainties" by nominal values (e.g., expected values) and solve the model thereafter. This approach has

been shown to be inferior to stochastic solutions through numerous papers (see Birge and Louveaux (1997a), Ruszczynski and Shapiro (2003), Wets (1996) and references therein). The classical stochastic approaches are stochastic programming and robust optimization. The difference between these two approaches is in the way that the constraints are treated.

The most obvious merit of stochastic programming is that uncertainty is constructed explicitly as part of the model. It uses random variables as representatives of uncertainty. Since the basic premise of stochastic programming is that the probability of random variables is known, the random variables can be generated by their underlying probability. Stochastic programming treats mainly soft constraints where the constraints may be violated with certain penalty. In the stochastic programming recourse models, for example, the decision variables are partitioned into two sets. The first is a set of variables that are decided prior to the realization of uncertain events. The second is a set of recourse variables which represent the optimal solution corresponding to the first stage decision and realized uncertainty. Other penalty methods are the "Scenario Optimization" and the "Entropic Penalty methods" (see Ben-Tal (1985), Rockafellar and Wets (1991) respectively). It is fair to say that stochastic programming solves a relaxation of the constraints (see Kall and Wallace (1994), and references therein).

The purpose of stochastic programming is to minimize the expected objective subject to some constraints, i.e.

$$\min_{x \in X} \mathbb{E}[g(x, \xi)] \tag{2.3}$$

where $x \in X \subset \mathbb{R}^n$ is the feasible region and $g : X \times \Omega \to \mathbb{R}$ is convex of x under every realization of ξ. However, it ignores higher moments of the distribution as risks. Therefore, the decides of model (2.3) is always risk neutral.

The first stochastic model was formulated in Dantzig (1955), and the model was named as the linear programming under uncertainty. The stochastic models can be divided into the following categories according to the complexity of the problems: linear, nonlinear, two stage stochastic programming problems, and multistage stochastic programming.

A classic stochastic programming problem "the farmer's problem" (Birge and Louveaux (1997b)) will be presented in order to demonstrate how to model uncertainty.

A Farmer's Problem: A farmer raised wheat, corn, and sugar beets on 500 acres of land. Before the planting season he wants to decide how much land to devote to each crop. At least 200 tons of wheat and 240 tons of corn are needed for cattle feed, which can be purchased from a wholesaler if not raised on the farm. Any grain in excess of the cattle feed requirement can be sold at $170 and $150 per ton of wheat and corn, respectively. The wholesaler sells the grain for 40% more (namely $238 and $210 per ton, respectively.) Up to 6000 tons of suger beets can be sold for $36 per ton; any additional amounts can be sold for $10/ton. Crop yields are uncertain, depending upon weather conditions during the growing season (see Table 2.1). Three scenarios have been identified ("good", "fair", and "bad"), each equally likely. (In this data, only the yields are scenario-dependent, while in reality the purchase prices and sales revenues from grain would be higher in year with poor yield, etc.)

Table 2.1. Yields under different weather condition

Scenario	Wheat yield (tons/acre)	Corn yield (tons/acre)	Beet yield (tons/acre)
Good	3	3.6	24
Fair	2.5	3	20
Bad	2	2.4	16

The general stochastic linear programming model is:

$$\min \quad cx + \sum_{k=1}^{K} p_k q_k y_k$$

Subject to

$$T_k x + W y_k = h_k, k = 1, \ldots, K; \tag{2.4}$$

In this example, only T_k varies by scenario, while the cost vector q_k and the right-hand-side h_k are fixed. Decision variables are

First stage:

x_1 = acres of land planted in wheat

x_2 = acres of land planted in corn

x_3 = acres of land planted in beets

Second stage:

w_1 = tons of wheat sold

w_2 = tons of corn sold

w_3 = tons of beets sold at \$36/T

w_4 = tons of beets sold at \$10T

y_1 = tons of wheat purchased

y_2 = tons of corn purchased

The stochastic decision problem is

$$\min \quad 150x_1 + 230x_2 + 260x_3 + 1/3 \sum_{k=1}^{3} Q_k(x) \tag{2.5}$$

$$\text{subject to} \quad x_1 + x_2 + x_3 \leq 500$$

$$x_j \geq 0, j = 1, 2, 3$$

28

Where $Q_i(x)$ is the optimal solution of the second stage (recourse) problem after the scenario has been determined, given that the first stage variables x have been selected.

$$Q_1(x) = \min \quad -170w_1 - 150w_2 - 36w_3 - 10w_4 + 238y_1 + 210y_2 \qquad (2.6)$$

$$\text{s.t.} \quad y_1 - w_1 \geq 200 - 3x_1$$

$$y_2 - w_2 \geq 240 - 3.6x_2$$

$$w_3 + w_4 \leq 24x_3$$

$$y_1 \geq 0, y_2 \geq 0, w_1 \geq 0, w_2 \geq 0, 0 \leq w_3 \leq 6000, w_4 \geq 0$$

$$Q_2(x) = \min \quad -170w_1 - 150w_2 - 36w_3 - 10w_4 + 238y_1 + 210y_2 \qquad (2.7)$$

$$\text{s.t.} \quad y_1 - w_1 \geq 200 - 2.5x_1$$

$$y_2 - w_2 \geq 240 - 3x_2$$

$$w_3 + w_4 \leq 20x_3$$

$$y_1 \geq 0, y_2 \geq 0, w_1 \geq 0, w_2 \geq 0, 0 \leq w_3 \leq 6000, w_4 \geq 0$$

$$Q_3(x) = \min \quad -170w_1 - 150w_2 - 36w_3 - 10w_4 + 238y_1 + 210y_2 \tag{2.8}$$

$$\text{s.t.} \quad y_1 - w_1 \geq 200 - 2x_1$$

$$y_2 - w_2 \geq 240 - 2.4x_2$$

$$w_3 + w_4 \leq 16x_3$$

$$y_1 \geq 0, y_2 \geq 0, w_1 \geq 0, w_2 \geq 0, 0 \leq w_3 \leq 6000, w_4 \geq 0$$

To formulate a problem in a consistent way, a number of fundamental assumptions should be made. By χ we denote the space of decision variables. And in most cases χ can be identified with a finite dimensional vector space $\mathbb{R}^n$. It is assumed that there is a given set $X \subset \chi$ of feasible decision and an objective function $F(x, \omega)$ of decision vector $x \subset X$ and random element ω. We consider ω as an element of a sample space Ω equipped with a sigma algebra F. In some cases, the random data is formed by a finite number of parameters. Consequently, the objective function is given in the form $F(x, \omega) := V(x, \xi(\omega))$, where $\xi(\omega)$ is a finite dimensional random vector and $V(x, \xi)$ is a function of two vector variables x and ξ.

$$\min F(x, \omega) \tag{2.9}$$

where $F(x, \omega)$ subject to $x \in X$ depends on ω, which does not make much sense. It means for different realizations of the random parameters, one would get different optimal solutions. But we are not looking to investigate the effects of different realizations on solutions.

A way to reduce the effects of different realizations is to optimize the objective function on average. And the new form of (2.9) is

$$\min_{x \in X} \{f(x) := \mathbb{E}[F(x, \omega)]\} \tag{2.10}$$

Since the probability distribution P is known on $(\Omega, \mathcal{F})$ as we assumed, the expectation of $f(x)$ is easy to get by definition.

$$\mathbb{E}[F(x, \omega)] = \int_{\Omega} F(x, \omega) dP(\omega) \tag{2.11}$$

For convenience, we use $f(x)$ to represent the expectation or expected value function.

Since we already have the objective function, the next step is to find out how to model the constraints. We have assumed that a given set $X \in \chi$ is a feasible set of decisions. For instance, the feasible set X can be written in a standard mathematical programming formulations as follows:

$$X := \{x \in X_0 : g_i(x) \leq 0, i = 1, \ldots, m\}, \tag{2.12}$$

When uncertainty is considered in the constrains, the constraints can be expressed in this formulation

$$G_i(x, \omega) \leq 0, i = 1, \ldots, m, \tag{2.13}$$

As previously discussed, stochastic programming treats mainly soft constraints, where the constraints may be violated with certain penalty. With the help of the mathematical model, this idea can be demonstrated more clearly. By that, we mean some values of χ may satisfy (2.13) for some ω and violate these conditions for other ω. It is unlikely to require that constraints (2.13) hold for all $\omega \in \Omega$.

There are several approaches available to include $G_i(x, \omega)$ into the formulation of constraints. And the most popular one is to consider the expected values as the way to build the objective function.

$$g_i(x) := \mathbb{E}[G_i(x, \omega)], i = 1, \ldots, m, \tag{2.14}$$

To reconcile the problem, as it varies, caused by uncertainty, the stochastic programming petitioners model the problem into two distinct stages. At the first stage, before a realization of the corresponding random variables become known, one chooses the first stage decision variables to optimize the expected value of an objective problem which in turn is the optimal value of the second stage optimization problem.

The two-stage stochastic linear program can be written as follows:

$$\min_{x} \quad c^T x + \mathbb{E}[Q(x, \xi(\omega))]$$

$$\text{s.t.} \quad Ax = b, x \geq 0, \tag{2.15}$$

where $Q(x, \xi)$ is the optimal value of the second stage problem

$$\min_{y} \quad q^T y$$

$$\text{s.t.} \quad Tx + Wy = h, \quad y \geq 0. \tag{2.16}$$

Here x and y are vectors of the first and second stage decision variables, respectively. $\xi = \xi(\omega)$ is already defined earlier. In this case, since the solution of the second stage problem depends on the data $\xi := (q, h, T, W)$ and some or all of the elements can be random, $\xi(\omega)$ is a random vector. The expectation in (2.16) is supposed to be known because it is depends on $\xi(\omega)$. The matrices T and W are called the technology and recource matrices,

respectively. If the W is fixed, the two-stage stochastic problem is called the problem with fixed recourse. It is worthwhile to note at this point that the problem (2.16) is a particular case of the stochastic programming problem.

As it was discussed above, the essence of two stage modeling is that the decision variables are partitioned into two sets. The value of the first vector $x \in \chi$ has to be decided before any realization of the unknown quantities. The observed value are summarized in the data vector $\xi = \xi(\omega)$. The value of the second part, y, can be chosen after the realization of ξ becomes known and generally depends on the realization of ξ and the choice of x. And y represents the optimal solution corresponding to the first stage decision and the realized uncertainty. As a result, at the first stage one has to solve the expectation optimization problem

$$\min_{x \in X} \mathbb{E}[F(x, \omega)] \tag{2.17}$$

And in the case of two stage linear problem,

$$F(x, \omega) := c^T x + Q(x, \xi(\omega)) \tag{2.18}$$

with $Q(x, \xi(\omega))$ being the optimal value of the second stage optimization problem, which totally depends on decision variable y.

There is another point to be made: the stochastic programming solves a relaxation of the constraints and the uncertainties' distributional information is assumed to be known. As in the example (2.15), (2.16), we may relax the expectation problem (2.17) by allowing the first stage decision variables to depend on the random data and then to correct that by some penalties built in constraints.

If a comparison is made between the solution of stochastic programming and its corresponding deterministic counterpart, it is noticeable that the solution from stochastic programming is well-hedged because of the built-in flexibility in models dealing with uncertainty. Another important observation is the sequencing of decisions and observations are important. In constructing a stochastic model, it is not enough to just specify the decision variables. The modeling person must also construct the model in such a way that prevents decisions that anticipate future uncertainty.

Last but not least, by solving the stochastic models, one can find the optimal solution on average. However, for a particular realization of uncertainty event, on a particular time period, the objective value might be vary different from the corresponding expected value. This is due to the variance in the uncertainty. So the solution from stochastic programming might not be the best for some scenarios.

2.3.2. Robust Optimization

The perturbation could be minor without significantly changing the nominal values of the parameter, however, the impacts on the optimal solution are not ignorable. Thus, before applying stochastic models for any real-world problems, it is always good to validate that the input data is reliable and perturbation will not happen in the future.

The Robust Optimization (RO) method has become more and more popular since the early 2000s. The main reason for its popularity is that the RO model is a decision support approach without assuming known distributional information. It generates a series of solutions that are progressively less sensitive to realization of the real input data. This method

seeks the tradeoff between optimality and feasibility. So it is a fitful method for some realistic applications (see Ben-Tal and Nemirovski (1999) and Ben-Tal and Nemirovski (2000)).

Consider the two-stage linear program (2.15), (2.16) In that problem the optimal value $Q(x, \xi(\omega))$ of the second stage problem is optimized on average. It is not impossible that for a particular realization ξ of the random data $\xi(\omega)$, the corresponding value $Q(x, \xi)$ can be quite different from the expected value $\mathbb{E}[Q(x, \xi(\omega))]$.

An "unlucky" realization of $\xi(\omega)$ may have disastrous consequences for the user of stochastic programming. Consider a situation: A *XYZ* company is thinking about investing in the stock market. Suppose there are two stocks that they are pretty interested in. The first stock has the same possibility of earning 1 dollar or losing 1 dollar per share with an even chance. The second stock may earn the company \$1001 or lose \$1000 per share with the same possibility. The expectation of the first stock is \$0 and the expectation of the second is greater than \$0. According to the expected-value criterion, the company will choose the second stock. But in fact, it is quite dangerous for the company to invest on the second one at the risk of losing 1000 per share. This example tells us why it is necessary to investigate trade-offs between means and variances of costs (or profit).

In order to avoid such disastrous consequences one may try be more conservative and to reach a compromise between the average(i.e, the mean) and a risk associated with variability of $Q(x, \xi)$. A model which integrates means and variance was introduced by Markowitz in Markowitz (1952) as a weighted mean-risk criterion. It seems natural to add the term $k\mathbb{V}[Q(x, \xi)]$ to the object of the optimization problem, where the coefficient $k \geq 0$ represents a compromise between the expectation and variability of the objective. However, this

will destroy the convexity and the second stage optimality of the target problem.

There are two types of robust models in literature. In the 1990s, some researchers placed the penalty caused by uncertainties in the objective, which is generally a Lagrangian relaxation. This method seeks the tradeoff between optimality and feasibility (see Mulvey et al. (1995)). This method is the same as SP in the way to deal with constraints, which treats mainly soft constraints.

Another RO model is called the "robust counterparts" (Ben-Tal and Nemirovski (1999, 2000), Ben-Tal (1998)), in which the parameters are modeled within a computational tractable structure such as an ellipsoid or an interaction of finite many ellipsoids. Therefore, the resulting robust counterpart would be efficiently solved by interior method without demanding computationally which it is essential for practical implementations. This approach treat uncertainty with hard constraints, which means the constraints must be satisfied whatever is the realization of the input data. In the remaining part of the proposal, RO model is automatically referred to the robust counterpart method.

In order to better understand the properties of RO models, we will introduce the general formation of robust counterpart in this part. The RO model is constructed specially for the uncertain linear programming (LP) problem. Suppose the data A, b within the LP model

$$\min \quad \{c^T x \quad | \quad Ax \geq b\} \tag{2.19}$$

are uncertain.

Denoted by $\mathcal{U}$ is the "uncertainty set", which includes all the realization of the data (A, b) that satisfies the hard constraints. Therefore, the feasibility of a vector x can be interpreted

as

$$Ax \geq b \quad \forall (A,b) \in \mathcal{U}. \tag{2.20}$$

As a result, the problem defined by

$$\min\{c^T x| \quad Ax \geq b \quad \forall (A,b) \in \mathcal{U}\} \tag{2.21}$$

is called the robust counterpart of the uncertain LP problem in Ben-Tal and Nemirovski (1999) and a vector x^* is called a robust solution of the uncertain problem.

Readers may have a question: how to construct the ellipsoidal "uncertainty set". Now, we are going to discuss where the ellipsoidal "uncertainty set" is from. In general, there is no uniform method that is applicable to every problem. It depends on the particular situation we meet. If we are given several primary scenarios of the data, we can construct the uncertainty ellipsoid as the minimal volume ellipsoid that contains these scenarios. There is another source of the ellipsoid uncertainties which is from statistical considerations. A simple portfolio problem will be presented in order to explain how robust optimization works. Portfolio problem as stated by Dantzig and Infanger (1993) is shown below.

Portfolio Selection Problem: The portfolio selection is the problem of allocating capital over a number of available assents in order to maximize the "return" on the investment while minimizing the "risk". For example, \$1 is used to invest at the beginning of the year in a portfolio comprised of n shares. The end-of-the-year return per \$1 invested in share i is $p_i \geq 0$. At the end of the year, you sell the portfolio. The goal is to determine the amount x_i to be invested in share i, $i = 1, \ldots, n$, so as to maximize the end-of-the-year portfolio value $\sum_{i=1}^{n} p_i x_i$.

When the quantities p_i are known in advance, the situation is modeled by the following simple LP program:

$$\max \{\sum_{i=1}^{n} p_i x_i | \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0\} \tag{2.22}$$

and the optimal solution is evident: we should invest all we have in the "most promising"(with the largest p_i) share. Assume all the coefficients p_i are distinct and $p_{(n)}$ is denoted as the largest coefficient. And the solution will the $x_{(n)}$ corresponding to $p_{(n)}$ should be equal to 1 and the others will be assigned as 0.

Mostly likely, the coefficients p_i are uncertain in reality. Assume that we know the "nominal" values of these coefficients $p_i^*, p_1^* \leq p_2^* \leq \dots p_n^n$ and bounds $\sigma_i \leq p_i^*$ such that the actual values of p_i are within the "uncertainty intervals" $\Delta_i = [p_i^* - \sigma_i, p_i^* + \sigma_i]$. Moreover, it is necessary to assume that p_i are of statistical nature and that they are *iid* in Δ symmetrically with respect to the nominal value p_i^*.

If this problem is solved by the stochastic method, the model can also be represented by (2.22) with p_i replaced by their "nominal" values p_i^*; the "nominal" solution is simply to invest all the money in the "most promising" share. This policy will generate a random yield x^{nom} with the expected value

$$\mathbb{E}\{x^{nom}\} = p_n^* \tag{2.23}$$

In contrast to the stochastic method, if a robust counterpart approach is applied on this problem, the corresponding robust optimal policy would invest everything in the shares with the largest worst-case return $p_i^* - \sigma_i$. The most important element of this approach is to construct an uncertainty set. The most straightforward answer is to construct a box. Each point of the box is a possible value of the uncertain coefficients p. For the simple

portfolio problem, the uncertainty set can be constructed as follows:

$$\mathcal{U}^{\theta} = \{p \in \mathbb{R}^n | \sum_{i=1}^{n} \sigma_i^{-2}(p_i - p_i^*)^2 \leq \theta^2\}. \tag{2.24}$$

The parameter θ is used to describe the decision maker's risk attitude; the larger θ is, the more risk averse s/he is. There are several important scenarios needed to be payed attention;

- when $\theta = 0$, $\mathcal{U}^{\theta}$ shrinks into the singleton $\mathcal{U}^0 = \{p^*\}$ which is the nominal data;
- when $\theta = 1$, $\mathcal{U}^1$ is the smallest volume ellipsoid containing the box.
- When $\theta = \sqrt{n}$, $\mathcal{U}^{\theta}$ is the largest volume ellipsoid containing the box.

The robust counterpart of (2.22) with respect to the uncertainty set $\mathcal{U}^{\theta}$:

$$\max\{\sum_{i=1}^{n} p_i^* x_i - \theta V^{1/2}(x) | \sum_{i=1}^{n} x_i = 1\} \tag{2.25}$$

where $V(x) = \sum_{i=1}^{n} \sigma_i^2 x_i^2$.

Problem (2.25) can be deemed as the mathematical model of the famous Markovitz portfolio selection model, although in his classical model $\sqrt{V(x)}$ is taken placed by Vx. As we mentioned in the beginning (2.3.2), robust optimization is a stable approach built with robust risk mechanism compared with stochastic programming. Due to the comprise of risk control mechanism, the optimal solution generated by the robust optimization approach may be not as optimal as the one from stochastic programming.

2.3.3. Concerns on modeling uncertainty

Modeling uncertainty using the above methods will inevitably encounter computational tractability issues. By July 2013, the fastest supercomputer ever built was a system called "Tianhe-2" or "Milky way 2", and deployed at the National Supercomputer Center in

Guangzho, China. This system is capable of executing 54.9 petaflops or (54.9×10^{24}) Floating point Operations Per Second (FLOPS)[1] When solving programs with n variables, the computer is actually solving an $n \times n$ Newton system of linear equations. With standard techniques, the cost will be $O(n^3)$. [2] Certainly, when special structures, e.g. banded, and/or sparse, are identified, some factorization method can be applied to significantly reduce the computational overhead. For example, the Cholesky factorization is a commonly-used approach which has been available in the Matlab suite (see Matlab user's manual).

Typical linear programming models in commercial scales always lead to a sparse Newton system and as a result, modern supercomputers are able to solve commercial linear programs with up to 10^6 variables (see Nemirovski (2000)). Typical nonlinear convex programs, however, usually lead to dense Newton systems and the supercomputer "Tianhe-2" is only capable of solving a nonlinear convex programming with up to 10^8 variables. Of course, thanks to the long-term trend of computing hardware, as embodied in Moore's law, the increasing trend on the maximum number of variables in a tractable dense Newton system is not expected to stop for another decade at least (see Kanellos (2009)).

Another issue we are concerned about is the stability of stochastic programming models under uncertainty. When the underlying distribution associated with random variables and the perturbed or contaminated constraints are encountered, the quality of the solution becomes questionable. The current research focuses on the Lipschitz continuity property on the optimal values (see Shapiro (2006) and references therein) instead of practical implementation because of two reasons. Firstly, the real distribution is not available for decision until revealed afterward; secondly, the Lipschitz constant is extremely difficult to evaluate

---

[1]Although FLOPS is a relevant measurement, it is used in most optimization literature.

[2]In computer science, the big O notation denotes the time complexity. The time complexity of an algorithm quantifies the amount of time taken by an algorithm to run as a function of the length of the string representing the input.

and thereby constructing a tight bound becomes highly impossible.

However, the negative impacts of stability issues on RO models are minimum because RO uncertainty is modeled as unknown-but-bounded variables without assuming underlying distributional information. In some literature, the RO model is described as a quantitative approach to control loss under the worst scenario. It is true that the RO solution cannot optimize the expected overall objective value compared to stochastic programming. But the solution is not as bad as it sounds. We need to point out that the term "worst scenario" does not mean that all the parameters are as "bad" as they could be. The RO model is modeled by an uncertainty set, which is to reflect the fact that the extremely worst values will not simultaneously happen. Rather than minimizing the first moment with distributional information, RO model would yield stable, feasible solution which is suitable for practical implementations.

In most real-world resource allocation problems, the decision maker will most likely prefer a plan that generates better object value and well controlled risks. Unfortunately, neither stochastic programming nor robust optimization will yield such a risk averse and a stable decision under uncertainty with incomplete and/or erroneous distributional information. By the previous introduction, the RO cannot generate the best optimized objective value, but the method is risk averse and the negative impacts of unwanted perturbations are minimized for the RO model, because it does not require the distributional information. On the other hand, the SP is risk neutral and not perturbation-proof, but it can generate optimized objective value.

So in response to those concerns and challenges, we propose a combined method, named as Resource Reservation (RR) for resource allocation problems under uncertainty with perturbation. First of all, this combined method is tractable and practical, so that it can be applied to well diversified uncertainty modeling applications. This method itself will enrich the body of knowledge on uncertainty modeling and solution techniques to facilitate decision making. Secondly, the essence of this method is only allocating a required amount of resources by the RO model, instead of allocating all the resources at once. And then the SP model will decide how to assign the required amount of resources to each region or client. This combined method contains multiple iterations. Within each iteration, we will solve both the RO and SP for different purposes. The RO model is to determine the amount of resources to allocate prior to possessing accurate information. By taking the amount of resource from the RO model, the SP model provides a plan that is used to optimize the objective value. The iteration will keep going until all the resources are depleted.

2.3.4. Comparison of the SP and RO methods

In some circumstances, robust optimization has some advantages over alternative approaches for dealing with uncertainty and is generally more applicable.

Stochastic programming, however, optimizes only the first moment of the distribution of the objective value of ξ. It ignores higher moments of the distribution, and the decision maker's attitude towards risks. The aim of stochastic programming is to optimize the expected value at each scenario through adjusting the control variables such as resource variables. Large changes may occur in ξ among different scenarios, but their expected value will still be optimal.

The term "perturbed stochastic programming" is a form of stochastic programming with perturbed, contaminated constraints, and incorrect estimation of the underlying distribution. Current research frontiers for perturbed stochastic programming problems are stability theorems and the minimax optimization model. Although they are insightful for theoretical research, they are unable to provide decision support for business scale problems. In typical applications of uncertainty modeling, the decision makers demand proactive actions which are to dynamically adjust the decision yielded by the perturbed stochastic programming problem.

Robust optimization models will minimize the higher moments as well, such as the variance of ξ. So the solution provided by the RO model will be a little conservative compared to stochastic programming with little or no adjustment of the control variables. Here we borrowed an example from Mulvey et al. (1995) to illustrate the difference in the real-world problem. For example, the SP and RO models are applied to personnel planning problems. An SP solution will design a workforce that can be adjusted (by hiring or layoffs) to meet demand at the least expected cost. In contrast to the SP model, the RO model will design a workforce that will need few adjustments to cope with demand for all scenarios. However, the total expected cost will be higher than that of the SP solution.

Stability issues will occur in the SP model when the input data is incomplete, erroneous, or perturbed. There are two types of stability issues, which are the consistency issue or qualitative stability issue, and the quantitative stability issue respectively in literature. The causes of consistency issue are statistical estimation on structural parameter, i.e. fixed parameter which are not subject to changes, and the perturbation/noisy data input. The structural parameters are largely coefficients which are used to describe the feasible region.

Since the structural data is estimated by various statistical methods, preserving consistency has been a major challenge. Almost surely, the sampling methods, based on the law of large numbers, are dominating techniques. The concept of ε-convergence has been extremely useful in theoretical analysis. The consistency issue has been addressed in many articles (see Pflug et al. (1998), Shapiro (1989, 1990), Vogel (1992)).

Another major stability issue is called quantitative stability in literature and it is caused by the perturbation on underlying probability measure. In uncertainty modeling, stability problems are generally caused by the following reason: The measure of ξ on the sample space $(\Omega, \mathcal{F})$ is $\mathbb{Q}$ rather than previously thought $\mathbb{P}$. If the true probability $\mathbb{Q}$ is known, one can get bounds of the error due to approximating $\mathbb{Q}$ by another measure $\mathbb{P}$. The idea is simple: choose in $\mathcal{F}$ a suitable metric d that metrizes, at least locally, the weak convergence. The quantitative stability mostly refers to continuity properties of the optimal objective value and a suitable distance, such as ζ-structure distance in Römisch (1993), Römisch and Schultz (1996), Römisch and Wets (2006) between the underlying and real probability measures. If the ζ-structure distance can be defined as follow:

$$d_{\mathcal{F}}(\mathbb{P}, \mathbb{Q}) = sup_{f \in \mathcal{F}} | \int_{\Omega} f(\xi) d\mathbb{P}(\xi) - \int_{\Omega} f(\xi) d\mathbb{Q}(\xi) | \qquad (2.26)$$

where f is measurable function in $\mathcal{F}$ and $\xi \in (\Omega, \mathcal{F}, \cdot)$. Typical metrics are, for instance, bounded Lipschitz metric and Kantorovich metric. For a two stage stochastic programming model, we have,

$$|v(\mathbb{P}) - v(\mathbb{Q})| \leq L \times dist(\mathbb{P}, \mathbb{Q}) \qquad (2.27)$$

where $v(\cdot)$ is the optimal values of the SP model under corresponding distribution $\mathbb{P}$ and $\mathbb{Q}$. Theoretically, the gap between $v(\mathbb{P})$ and $v(\mathbb{Q})$ can be adjusted by changing Lipschitz

metric. But in practice, it is not easy to do for the following reasons: first, the real distribution $\mathbb{Q}$ is not available for decision until the event occurs in reality. Secondly, the Lipschitz metric is extremely difficult to evaluate and thereby constructs a tight bound becomes almost impossible.

Compared to the SP model, the stability issue does not affect the effectiveness of the RO model because the RO method does not depend on the distributional information. Instead, it is modeled by unknown-but-bounded variables without assuming underlying distributional information. To be more specific, the uncertainty set, which we already introduce earlier, is modeled by an ellipsoid, which has little sense to do with the support of the distribution of ξ.

The RO model has a better control of the variance of solutions than the SP model. The variance of solutions has been ignored for a long time until Markowitz raised the issue and attracted researchers' attention in the application of portfolio management. Hence, he also won the Noble Prize. The RO model can address the issue directly.

Another aspect that distinguishes a RO model from a SP model is the way to deal with constraints. The SP model aims to find the decision variables x such that for each realized scenario, a control variable y is possible that satisfies the constrains. It is even allowed to have complete recourse in the SP models. By which, we mean a feasible solution y exists for all scenarios and for any value of x that satisfies the control constraints. If there does not exist a feasible pair (x, y) for each scenario, the SP model will be declared infeasible. However, the RO model will not allow this to happen. By the first type of RO models, the scenario-based penalty approach by Mulvey, Vanderbei and Zenios in Mulvey et al. (1995), the RO model will find a solution that violates the constrains with minimal amount.

However, by the other type of RO models, "robust counterparts" developed by Ben-Tal and Nemirovski in Ben-Tal and Nemirovski (1999), is not allowed to violate the constrains at all.

When the data is incomplete, erroneous, or perturbed, there will be another distinction from the RO model to the SP model. For the SP model, stability issues will occur due to the incomplete, erroneous or perturbed data. The "decision environment", that we normally deal with, has uncertainties in some parameters, which are important for modeling the resource allocation systems. Hence, motivated by the stability issues of the SP model and wide applications of uncertain systems, we focus on investigating the optimal allocation policy when the model parameters and the underlying distributional information are perturbed.

CHAPTER 3

RESOURCE RESERVATION MODELS

The idea of resource reservation methodology is inspired by the following factors:

- Data uncertainty in real-world problem cannot be ignored (see Ben-Tal and Ne-mirovski (2000)). Even a small uncertainty in the data can make the nominal optimal solution to the problem completely meaningless in practice. Therefore, there is a real need of a methodology capable of generating a robust solution that is immunized against the effect of data uncertainty.

- Two popular methods that are capable of dealing with uncertain data have their own drawbacks, which limit their applicability. SP models are fundamentally based on the assumption that the random variables' distribution information is known. However, the data in reality are more likely to be perturbed. In contrast, RO models do not need to assume known distribution information. When the data are incomplete, erroneous, there will be consistency and stability issues reported in the literature of stochastic programming.

Motivated by the real desire from real-world problems and the applicability of the SP and RO models, we propose a novel combined approach "Resource Reservation" (RR) to solve stability issues for vaccine allocation optimization. This method is designed for dealing with wrong estimation or massive perturbation and thereby resolve the stability issues raised among the SP models. The negative impacts of the unwanted perturbations are min-imized for the RO model because the RO uncertainty is modeled as unknown-but-bounded variables without unrealistic assumptions. RO model is aimed to yield stable, feasible solu-tion for the practical implementations. However, the RO solution usually cannot optimize

the expected objective. In our opinion, when the government agency is preventing the loss of life and optimizing the benefits from the vaccines allocation under uncertainty, neither RO nor SP models will be solely qualified. Hence, we propose the RR method for vaccines allocation optimization. It is a combination of the SP and RO models with multiple iterations. The RO part of RR method will effectively control the risk whereas the stochastic part will optimize the first moment of the objective.

The basic concept of the RR approach is fairly straightforward: when the critical factor such as the amount of demands that affect the allocation plan is quite uncertain and changing all the time, people should be more conservative in deciding how to allocate the resources. In order to avoid the scenario that allocates all the resources and regrets it later, the RR approach suggests that we only allocate an absolutely necessary amount of resources. The resources not being allocated can be treated as safety stock in logistics in case there will be uncertain demands in the future. The absolutely necessary amount of resources will be determined by the RO approach with the unknown-but-bounded random variables and risk attitude θ as input information. Since the RO model can minimize loss under the worst scenario, the corresponding decision from the RO model tends to be conservative with less amount of resource consumptions than the current available resources. Therefore, the RO model is suited to make a rough allocation plan prior to possessing accurate information. The result from the RO model will feed the SP model, which basically includes the total amount of resources to be allocated. By taking the amount of resources from the RO model, the SP model provides a plan on to optimally assign those resources to each request, with best expected objective.

It is encouraging to know that similar research topics have been incorporated in computer science and network communication (see Resource ReSerVation Protocol, RFC 2205). For

example, in the multimedia network communication, the CPU, memory, and flash space are reserved by proactive actions in order to assure that the application with high priorities are executed smoothly. Although these terminologies are based on different methodologies, the fundamental similarities are not deniable and their successful implementations ensure the prospect of our RR method.

3.1. Resource Reservation Heuristic

We propose the RR model to tackle the stability issue of the SP models under unwanted perturbations. The perturbation can easily invalid the perviously obtained optimal solution. In order to avoid the later regret, we combine the RO and SP models to formulate the RR method on the planning horizon $t = 1, \ldots, T$. At time t, we will solve the RO model first to obtain the amount of resource for allocation. Then, we will allocate such number of resources by the SP model to optimize the first moment of the objective. We will repeatedly solve RO and SP with updated, hopefully more accurate demand estimation.

Before proceeding to introduce the RR model in details, it is helpful to define some assumptions to assure that the later discussions are more meaningful.

(1) The feasible region $X := \{x \in R_+^n \,|\, Ax \leq b\}$ where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ is not empty.

(2) The objective $g(x, \xi) : X \times \Omega \to \mathbb{R}$ is convex of x for every ξ and $\mathbb{E}[g(x, \xi)] < \infty$.

(3) The available resource vector b is deterministic and completely known.

(4) ξ's distribution is unknown but estimable. The underlying distribution in $\mathbb{P}$ and the observed distribution is $\mathbb{Q}$.

First of all, we establish the deterministic linear programming to allocate resources to vaccines planning. Consider a situation in which n counties in Kentucky must be supplied

with doses of vaccines A due to a pandemic. Each of the *n* is to be supplied from a central warehouse at different points in time. The following assumptions hold:

(1) Doses of medicine can not be transported from one county to another, but they can be held over from one time period to the next in the same county.

(2) Half of the people who required the vaccines at the time of the first delivery will die if they do not receive the medicine at that time; the other half can wait until the second delivery.

(3) Demands from each county is uncertain but we assume they follow a uniform distribution with upper and lower limits.

(4) The available resources held in the central warehouse is known and fixed at the beginning of the distribution.

The deterministic linear programming model is as follow:

$$\min \sum_{i=1}^{n} \max\{\frac{1}{2}(d_i - x_i), 0\} \tag{3.1}$$

$$s.t. \ \sum x_i \leq c$$

$$x_i \geq 0$$

Where *i* denotes the ith county out of n counties, $i = 1, \ldots, n$. d_i denotes the demand of the ith county. x_i denotes the amount of resources that is allocated to the ith county. The objective function is to minimize the loss of life. We compare the difference between demands and allocated resources and if the difference is larger than 0 which means there are people who can not get vaccines. Half of the people who do not get vaccines will die. Since this is a deterministic linear programming model, it assumes the demands are accurate.

We can remove the max by re-writing the model in the following form.

$$\min \sum_{i=1}^{n} [\frac{1}{2}(d_i - x_i)] \tag{3.2}$$

$$\sum_{i=1}^{n} x_i \leq c$$

$$\frac{1}{2}(d_i - x_i) \geq 0$$

$$x \geq 0$$

We adopt the standard formulation of stochastic programming as follows:

$$\min \ f_0(x) + \mathbb{E}[g(x, \xi)] \tag{3.3}$$

$$s.t. \quad Ax \leq b$$

in order to simplify the notation. In the model above, the random variable $\xi \in (\Omega, \mathcal{F}, \mathbb{P})$ where $\mathbb{P}$ is the underlying but rough estimation of distributional information of ξ. $A \in \mathbb{R}^{m \times n}$, and functions $g : \mathbb{R}^n \times \omega \to \mathbb{R}$, $f_0(x) : \mathbb{R}^n \to \mathbb{R}$ are convex with the respect to x. We are interested in the model performance under the real distribution of ξ in another measure $\mathbb{Q}$. It is also helpful to define that:

(1) $v(\mathbb{P}) := \inf \{ f_0(x) + \mathbb{E}[g(x, \xi)] : Ax \leq b \}$

(2) $X_\varepsilon^*(\mathbb{P}) := \{ Ax \leq b : f_0(x) + \mathbb{E}[g(x, \xi)] \leq v(\mathbb{P}) + \varepsilon \}$

(3) $X^*(\mathbb{P}) := X_0^*(\mathbb{P}) = \{ Ax \leq b : f_0(x) + \mathbb{E}[g(x, \xi)] \}$ where $\varepsilon \geq 0$

In (2.3.3), we brought up the stability issues with stochastic programming, and in (2.3.4) the current research, with respect to the stability problem of stochastic programming, is reviewed. ζ-structure distance can be found in (2.26). In Rachev and Römisch (2002), and Römisch (2003), the authors study the behavior of stochastic programming problems in case of the underlying probability distribution is perturbed. They establish stability results for two-stage and chance-constrained models. One possible useful quantitative stability

result, with respect to weak convergence of probability measure, plays a very important role in justifying the stochastic programming as part of our proposed metrology.

Theorem 1. *Rachev and Römisch (2002) Let the set $\{x : Ax \leq b\}$ be non-empty. Let a sequence of probability measures $\{\mathbb{P}_n\}$ is weakly convergent to $\{\mathbb{Q}\}$ and*

$$\lim_{n \to \infty} d_{\mathcal{F}}(\mathbb{P}_n - \mathbb{Q}) \tag{3.4}$$

Then the sequence $\nu(\mathbb{P}_n)$ converges to $\nu(\mathbb{P})$ and

$$\lim_{n \to \infty} \sup_{x \in X^*(\mathbb{P}_n)} \left[\inf_{y \in X^*(\mathbb{Q})} \{\|x - y\|\} \right] = 0 \tag{3.5}$$

Consider $\mathbb{P}$ the first estimation on ξ for the stochastic programming model (3.3). As more information is revealed, the estimation of ξ becomes more meaningful. Thus, we assume that the sequence of underlying probability measures $\mathbb{P}_n$, which are the estimation of ξ with continuous "learning", converges weakly to the real distribution $\mathbb{Q}$. By Theorem 1, the sequence of optimal values $\nu(\mathbb{P}_n)$ will converges to $\nu(\mathbb{Q})$, i.e.

$$\mathbb{P}_n \to \omega\mathbb{P} \quad and \quad \nu(\mathbb{P}_n) \to \nu(\mathbb{Q}) \tag{3.6}$$

and

$$\lim_{n \to \infty} \sup_{x \in X^*(\mathbb{P}_n)} \left[\inf_{y \in X^*(\mathbb{Q})} \{\|x - y\|\} \right] = 0 \tag{3.7}$$

The decision variables are the amount of resources allocated to each selected county x_i, $i = 1, \ldots, n$. The demands for the resources are random variables ζ_{ik}. Thus, the SP model as part of the RR heuristic is

$$\min_x \{ \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{K} \mathbb{E}[(x_{ik}, \zeta_{ik})] : \sum_{i=1}^{n} x_{ik} \leq c, x_{ik} \leq \zeta_{ik} \} \tag{3.8}$$

where $x_{ik} \geq 0, \quad i = 1, \ldots, n$. c denotes the amount of available resources. ζ is the uncertain demand from uniform distributions with upper and lower limits. $k = 1, \ldots, K$ are the scenarios.

The model (3.8) will provide a resource allocation plan. The typical solution to model (3.8) is to allocate all of the available resources. However, during a pandemic disease, the demands of resources are hard to predict and we don't want to take the risk of allocating all the resources in the beginning and have nothing left for the future demands. Moreover, the estimation of ζ could be very rough if it is not worse. The underlying distribution of (3.8) could be substantially different in reality and consequently, the obtained solution becomes highly problematic. In order to solve the above issues, we introduce another uncertainty optimization method called "Robust Optimization" which can solve the stability issue of the stochastic programming and also it doesn't require the estimation of distributional information of ζ. In some articles, the RO model is describe as a quantitative approach to control loss under the worst scenario and an RO model might seem overly conservative because of the term, the worst scenario. We need to emphasize that the worst scenario does not mean the parameters are all as "bad" as they could be. RO models take into account an uncertainty set, which reflects the fact that the extremely worst values will not simultaneously occur.

Rather than modeling the demands on the estimation of distribution information, we model the demands as unknown-but-bounded random variables, i.e. $di \geq 0$ such that $d_i \in [d_i^* - \Delta d_i, d_i^* + \Delta d_i]$ where $\Delta d_i \geq 0$ is the range of corresponding demand. We assume that d_i^* is distinct and we define $d := [d_1; \ldots; d_i; \ldots; d_n]$ as a $n \times 1$ dimensional demand vector. The

realized demand is modeled as

$$d_i = d_i^* + u_i \Delta d_i, \quad i = 1, \ldots, n, \quad where \quad \|u\| \leq \theta \tag{3.9}$$

where $u := [u_1; \ldots; u_i; \ldots; u_n]$ and $\| \cdot \|$ is the norm of u. Using l^2 norm, the set for possible demands is an ellipsoid,

$$u^\theta = \left\{ d | \sum_{i=1}^{n} \frac{(d_i - d_i^*)^2}{\theta^2 \Delta d_i^2} \leq 1 \right\} \tag{3.10}$$

where d_i^* and Δd_i are known scalars. The transformation is designed to model the perturbation by unknown-but-bounded random variables. The scale of uncertainty is modeled by a Euclidean ball in $\mathbb{R}^n$ with radius of θ.

In order to remove the uncertainty from the objective to the feasible region, we can construct the RO model into the following form.

$$\min_{y,t} \quad t \tag{3.11}$$

$$s.t. \quad \sum_{i=1}^{n} y_i \leq c \tag{3.12}$$

$$\frac{1}{2} \sum_{i=1}^{n} (-y_i + d_i) \leq t \tag{3.13}$$

$$d_i - y_i \geq 0 \tag{3.14}$$

$$y_i \geq 0 \tag{3.15}$$

where $i = 1, \ldots, n$ and y_i are the decision variables.

Let $d^* := (d_1^*, d_2^*, \ldots, d_n^*)'$ and $\Delta d := (\Delta d_1, \Delta d_2, \ldots, \Delta d_n)$. When incorporating the unknown-but-bounded random variables,

$$
d = \begin{pmatrix} d_1^* \\ d_2^* \\ \vdots \\ d_n^* \end{pmatrix} + \begin{pmatrix} \Delta d_1 & 0 & 0 & \ldots & 0 \\ 0 & \Delta d_2 & 0 & \ldots & 0 \\ & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \Delta d_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \tag{3.16}
$$

Plug (3.9) into (3.13), then we have

$$
\frac{1}{2} \sum_{i=1}^n (-y_i + d_i) \le t \tag{3.17}
$$

$$
\iff \frac{1}{2} \sum_{i=1}^n (-y_i + d_i^* + \mathcal{U}_i \Delta d_i) \le t \tag{3.18}
$$

Theorem 2. *The deterministic constraint* (3.18) *can be re-written into a conic quadratic constraint*

$$
\frac{1}{2}[\theta \sqrt{\sum_{i=1}^n \Delta d_i^2} - \sum_{i=1}^n (y_i - d_i^*)] \le t \tag{3.19}
$$

PROOF.

$$
\frac{1}{2} \sum_{i=1}^n (-y_i + d_i^* + \mathcal{U}_i \Delta d_i) \le t \tag{3.20}
$$

$$
\iff \max_{\|\mathcal{U}\|_2 \le \theta} \sum_{i=1}^n (\mathcal{U}_i \Delta d_i) \le 2t + \sum_{i=1}^n (y_i - d_i^*) \tag{3.21}
$$

$$
\iff \theta \sqrt{\sum_{i=1}^n \Delta d_i^2} \le 2t + \sum_{i=1}^n (y_i - d_i^*) \tag{3.22}
$$

$$
\iff \frac{1}{2}[\theta \sqrt{\sum_{i=1}^n \Delta d_i^2} - \sum_{i=1}^n (y_i - d_i^*)] \le t \tag{3.23}
$$

□

55

For the constraint (3.14), such that $i = 1, \ldots, n$, the following constraint must hold under the "worst-case" scenario.

$$\frac{1}{2}(y_i - d_i) \leq 0 = \frac{1}{2}(y_i - d_i^* - \mathcal{U}_i \Delta d_i) \leq 0 \tag{3.24}$$

where $\|\mathcal{U}\| \leq \theta$.

Reorganize the equation (3.24) under the "worst-case" scenario

$$- \mathcal{U}_i \Delta d_i \leq \frac{1}{2}(-y_i + d_i^*) \tag{3.25}$$

$$\Longleftrightarrow \max_{\|\mathcal{U}\| \leq \theta} - \mathcal{U}_i \Delta d_i \leq \frac{1}{2}(-y_i + d_i^*) \tag{3.26}$$

$$\Longleftrightarrow \theta \Delta d_i \leq \frac{1}{2}(-y_i + d_i^*) \tag{3.27}$$

$$\Longleftrightarrow \theta \Delta d_i + \frac{1}{2}y_i \leq \frac{1}{2}d_i^* \tag{3.28}$$

As an instant result, we move the constraint (3.19) back into the objective, and we place constraints back into the model.

Corollary 1. *The robust optimization version of model* (3.11) *is*

$$\min \frac{1}{2}[\theta \sqrt{\sum_{i=1}^{n} \Delta d_i^2} - \sum_{i=1}^{n}(y_i - d_i^*)] \tag{3.29}$$

$$s.t. \sum_{i=1}^{n} y_i \leq b \tag{3.30}$$

$$\theta \Delta d_i + \frac{1}{2}y_i \leq \frac{1}{2}d_i^* \tag{3.31}$$

$$y_i \geq 0 \tag{3.32}$$

where $i = 1, 2, \ldots, n$.

56

There is a fundamental difference between model (3.11) and (3.29). In model (3.11), all of the parameters are constants and under no uncertainty and perturbations are incorporated. On the other hand, model (3.29) allows the structural parameters to change with the range such as $d_i^* \pm \Delta d_i$. Therefore, the model (3.29)'s solution is a robust solution with uncertainty and perturbation fully considered.

When introducing the robust optimization technique, we claim that it is designed to minimize the loss under the worst-case scenario. This statement might give us the impression that the solution of model (3.29) is overly conservative. Actually the conservativeness can be adjusted by changing the value of θ. As previously discussed, the uncertainly is modeled by being based on the definition of uncertainty set. The RO model is designed to provide a robust solution under some less favorable scenarios, which are modeled by the uncertainty set $\|\mathcal{U}\| \leq \theta$, an Euclidean ball with a radius of θ. By adjusting θ, we can either increase or decrease the perturbation and the corresponding results will be adapted to the adjustment. To be more specific, When $\theta = 0$, the RO model becomes a linear programming with nominal parameters. When $\theta = 1$, the worst-case scenarios are defined on the hull of the uncertainty set, i.e., the ellipsoid. The uncertainty set becomes the largest volume ellipsoid contained in $\mathcal{B} := \{d||d_i - d_i^*| \leq \Delta d_i^*\}$. When $\theta = \sqrt{n}$, the uncertainty set enlarges to cover all the possible scenarios and the uncertainty becomes the minimum volume ellipsoid contained in $\mathcal{B}$. Thus, the value of θ can be understood as a trade-off between less risk averse and more risk averse. When the volume of the ellipsoid is larger, it means the feasible region of the solution is larger. For example, for two uncertainty sets $\mathcal{U}_1, \mathcal{U}_2$ with the value of θ at θ_1, θ_2 respectively such that $\theta_1 > \theta_2$. Therefore, $\mathcal{U}_1$ covers more events than $\mathcal{U}_2$ and the worst-case scenario associated with $\mathcal{U}_2$ will be "less worse" than the worst-case of $\mathcal{U}_1$.

Figure 3.1 from the numerical results of the case study proves that the changes of the value of θ will directly affects the risk attitude of the solutions. The value of θ in the example is fixed at 3, 2.5 and 1 respectively and other factors keep the same. The graph demonstrates one of our conclusions that adjusting the value of θ will change the conservativeness of the results. As the value of θ grows, the feasibility of the uncertainty set expands as well. The more "worst-case" scenarios will be incorporated by the model, the more conservative the solutions are likely to become. As we can see in Figure 3.1, when the amount of the available doses is fixed, the model tends to reserve more resource when the value of θ grows bigger. In practice, we prefer $\theta \in [0.5, 1.5]$ to remove the concern of risk sensitiveness.

Figure 3.1. Sensitivity analysis: available doses vs reserved resource/available doses vs θ.
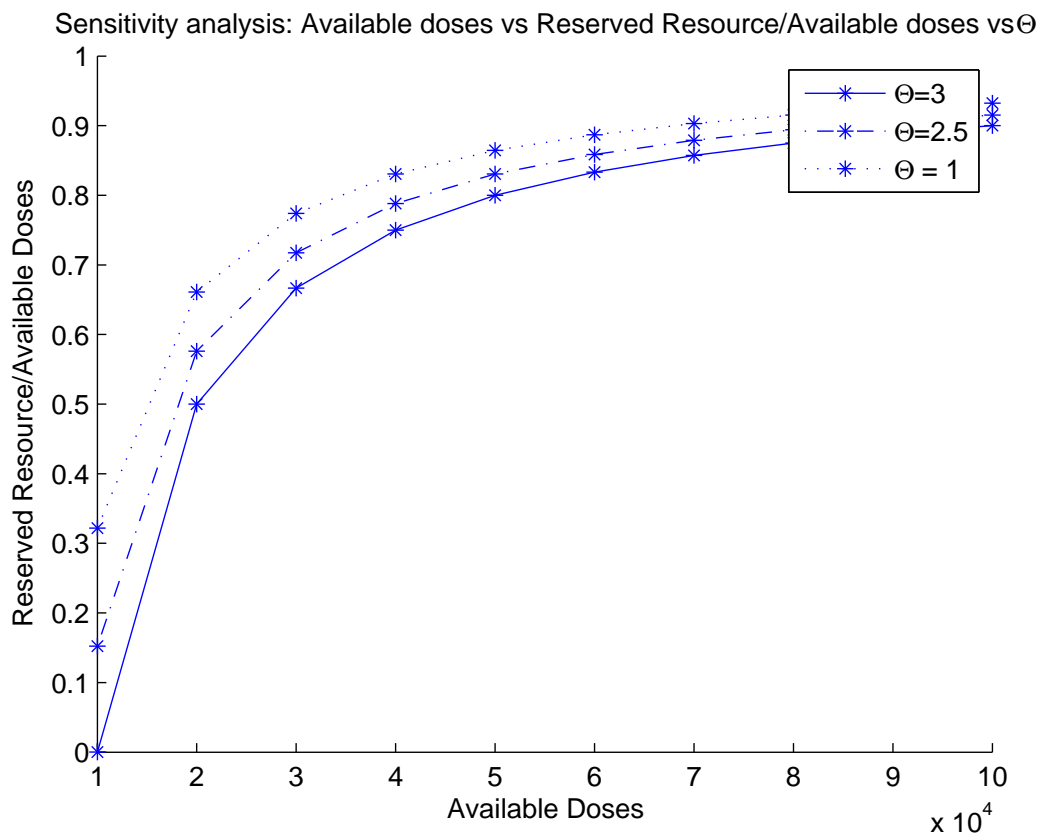
Suppose y^* is the optimal solution generated by the RO model (3.29) with certain θ and x^* is the optimal solution generated by the SP model (3.8). y^* is a reliable solution even when the uncertain demand ζ is incomplete or wrong, but the overall performance might be poor. x^* is the solution that can maximize or minimize the expected objective with known distributional information. Given the fact that the estimation on ζ could be wrong or incomplete in reality, especially during some unexpected pandemic diseases, allocating resources by x^* could also be problematic.

In order to obtain a balanced resource allocation plan, we present our metrology named " Resource Reservation". The idea of reserving a well determined amount of resources will be an effective way to deal with uncontrollable uncertainty, perturbation, and poor estimations.

The resource reservation method for vaccines allocation with multiple deliveries can be describe as follows:

Step 1: Collect necessary information from decision makers, such as the amount of available resources c, selecting resource requesting counties d_i, predicted infection rates, and estimated error which are used to calculated the upper bound $d_i^* + \Delta d_i$ and lower bound of demands $d_i^* - \Delta d_i$ combined with the population of each selected county.

Step 2: Solve the RO model (3.29) for the amount to be reserved and released.

Step 3: Solve the SP model (3.8) with the released resource. The solution from the SP model will be the allocation plan for this period.

Step 4: Depending on the number of deliveries that decision makers input and if the number of distribution is more than one, Step 2 and Step 3 will repeat multiple times. In the last time period, only the SP model will be used since we are trying to

release as much of the remaining resources as we can. If the number of distribution is just one, only the RO model will be ran.

3.2. Solution Technique and Computational Complexity

A solution technique for a class of optimization problems is an algorithm that computes a solution of the problem (to some given accuracy), given a particular problem form the class, i.e., an instance of the problem. A large effort has gone into developing algorithms for solving various classes of optimization problems since the late 1940s (Boyd and Vandenberghe (2004), Nemirovski (2006), Nesterov and Nesterov (2004)). The effectiveness of these algorithm varies considerably, and depends on factors such as the particular forms of the objectives and constraint functions, how many variables and constraints there are, and special structure, such as "sparsity". (A problem is "sparse" if each constraint function depends on only a small number of the variables).

It was proved that generic convex problems, under mild computability and boundedness assumptions, are polynomially solvable. Consider a family of convex problems (2.2) of a given analytical structure, like the family of LP problems, or Linearly constrained Quadratic problems, or Quadratically constrained Quadratic ones, etc. Assume p is a problem instance of a generic mathematical programming problem $\mathcal{P}$ with finite-dimensional data vector $\mathcal{D}(p)$. $\mathcal{D}(p)$ can be understood as the collection of the numeric coefficients in analytical expressions for the objective and the constraints. The dimension of the data vector is called the size $\mathcal{L}(p)$ of the problem instance. A solution technique for solving problems from the family is a routine which, given on input the data vector, generates a sequence of approximate solutions to the problem in such a way that every of these solutions is obtained in finitely many operations of precise real arithmetic, like the four arithmetic operations, taking square roots, exponents, logarithms and other elementary functions; We

call a solution technique "convergent", if, for any positive ε and for any problem instance p

from the family, the approximate solutions x_i generated by the method, starting with certain

$i = i^*(\varepsilon, p)$, are ε-solution to the problem, i.e., they belong to G and satisfy the relations

$$f(x_i) - f^* \leq \varepsilon, g_i(x_i) \leq \varepsilon, j = 1, \ldots, m, \tag{3.33}$$

where f^* is the optimal value in the problem. We call a solution technique "polynomial",

if it is convergent and the arithmetic cost $C(\varepsilon, p)$ of ε-solution, i.e.., the total number of

arithmetic operations at the first $i^*(\varepsilon, p)$ steps of the technique as applied to p, admits an

upper bound as follows:

$$C(\varepsilon, p) \leq \pi(\mathcal{L}(p)) \ln\left(\frac{\upsilon(p)}{\varepsilon}\right), \tag{3.34}$$

where π is certain polynomial independent on the data and $\upsilon(p)$ is certain data-dependent

scale factor. The ratio $\frac{\upsilon(p)}{\varepsilon}$ can be interpreted as the relative accuracy which corresponds

to the absolute accuracy ε, and the quantity $\ln\left(\frac{\upsilon(p)}{\varepsilon}\right)$ can be understood as the number of

accuracy digits in ε. With this interpretation, the polynomially of a method means that for

this method the arithmetic cost of an accuracy digit is bounded from above by a polyno-

mial of the problem size, and this polynomial can be thought of as the characteristic of the

complexity of the problem.

Model (3.29) is a convex optimization because the objective is a composition of a norm on

an affine function and the constraints are linear and conic in nature. Thus, the convexity is

perfectly preserved, and model (3.29)'s ε-optimal solution can be obtained within polyno-

mial operational counts.

As a result, the robust optimization method outperforms other uncertainty methods due to its extremely efficient solution techniques. In practice, model (3.29) can be solved by the modern interior point algorithm in milliseconds.

Speaking of interior-point methods, they are used for solving convex optimization problems that include inequality constraints,

$$\min \quad f_0(x) \tag{3.35}$$

$$\text{subject to} f_i(x) \leq 0, i = 1, \ldots, m$$

$$Ax = b,$$

where $f_0, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ are convex and twice continually differentiable, and $A \in \mathbf{R}^{p \times n}$ with $\mathbf{rank} A = p \leq n$. We assume that the problem is solvable, i.e., an optimal x^* exists. We denote the optimal value $f_0(x^*)$ as p^*. We also assume that the problem is strictly feasible, i.e., there exists $x \in \mathcal{D}$ that satisfies $Ax = b$ and $f_i(x) \leq 0$ for $i = 1, \ldots, m$. This means that Slater's constraint qualification holds, so there exist dual optimal $\lambda \in \mathbf{R}^m$, $\upsilon \in \mathbf{R}^p$, which together with x^* satisfy the KKT conditions

$$Ax^* = b, f_i(x^*) \leq 0, i = 1, \ldots, m \tag{3.36}$$

$$\lambda^* \succeq 0$$

$$\nabla f_0^*(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + A^T \upsilon^* = 0$$

$$\lambda_i^* f_i(x^*) = 0, i = 1, \ldots, m.$$

Interior-point methods solve the problem (3.35) or (3.36) by applying Newton's method to a sequence of equality constrained problems, or to a sequence of modified versions of the KKT conditions. The barrier method can be applied to find the ε-suboptimal solution. The convergence analysis and complexity analysis have been done for the barrier method in Boyd and Vandenberghe (2004).

In terms of the stochastic programming models, typical linear programming models in commercial scales always lead to a sparse Newton system and as a result, modern supercomputers are able to solve commercial linear programs with up to 10^6 variables (see Nemirovski (2000)). Typical nonlinear convex programs, however, usually lead to dense Newton systems and the supercomputer "Tianhe-2" is only capable of solving a nonlinear convex programming with up to 10^8 variables. Of course, thanks to the long-term trend of computing hardware, as embodied in Moore's law, the increasing trend on the maximum number of variables in a tractable dense Newton system is not expected to stop for another decade at least (see Kanellos (2009)).

CHAPTER 4

CASE STUDY: KENTUCKY SNS VACCINE ALLOCATION

In this chapter, we will show the value of our model by applying our proposed methodology to the Kentucky SNS Vaccines Allocation problem. The models presented in Chapter 3 are the general form of an SP and RO model. They can be used for any type optimal resource allocation under perturbation problem. In Chapter 3, we give the formation of the RR models for our problem in particular. Given the fact of lacking real data, we conduct numerical experiments on testing data. Based on the test problem, the formulation of the models is changed accordingly. All the numerical experiments are performed on Matlab 7.12.0 (R2011a), Windows 7 64 operation system with 4GB RAM. In order to solve the RR models, we utilize "Optimization Toolbox" within Matlab and a downloadable "CVX Toolbox", which is a Matlab-based modeling system for convex optimization developed by Stanford University. CVX turns Matlab into a modeling language, allowing constraints and objectives to be specified using standard Matlab expression syntax.

Numerical results from the case study will be presented to show how the model can help decision makers such as the Kentucky SNS coordinators, to develop Kentucky vaccines allocation strategies and plans. Moreover, sensitivity analysis is applied to the mathematical models, which is used to demonstrate the effects of different of risk factor settings on the output. At last, a user-friendly web application is developed by the RTDSS team and will be present at the end of this chapter. The interface of the RR heuristic for the Kentucky SNS is co-designed by Dr. Holman from the department of Industrial Engineering and web pages are developed by Anala Panit from the Department of Computer Engineering and

Computer Science at the University of Louisville.

4.1. Numerical Results for One Time Period Resource Allocation

This experiment is conducted for the scenario that only one delivery is required during the whole planning horizon. When there is only one time period, the RR heuristic only needs to run the RO model. That is why we want to have a separate session for the experimental results of one time period resource allocation.

In this experiment, we are investigating the effect of the changes of "Available doses" and "θ" on the allocation solutions. The population of ten selected counties is available in the table 4.1, as well as the PIR and "Estimated Error" as user inputs. Column "d^*" and column "Δd_i" are the estimated affected populations and its corresponding range, which will be calculated automatically in our mathematical programming models.

By default, the PIR is calculated by the number of cases to date, and the number of new cases increased within a certain number of days, which are input by the users. The way to calculate the PIR can be expressed in this formula:

$$PIR = \frac{\text{Number of new Cases/Periods of time by days}}{\text{Number of cases to date}} \qquad (4.1)$$

Using the PIR times each selected county's population, we can get the average estimated affected population and using the estimated error times, the average estimated affected population, we can get the upper and lower bound of each estimation. Once again, the administer is able to change the default PIR, which is automatically calculated by the formula (4.1) in the web application. We keep the PIR and the "Estimated Error" fixed at 10% and 5%, and change the value of "Available doses" and "θ" gradually. We do not consider the

effect of the risk level that the user is required to manually select in the web application, since the essence of the manually selected risk level is to change the actual amount of " Available doses" input to the RR models.

Table 4.1. Population size of selected counties

County Name	Population	PIR	d^*	Estimated Error	Δd_i
Adair	18656	10%	1866	5%	93
Allen	19956	10%	1996	5%	100
Anderson	21421	10%	2142	5%	107
Ballard	8249	10%	825	5%	41
Barren	42173	10%	4217	5%	211
Bath	11591	10%	1159	5%	58
Bell	28691	10%	2869	5%	143
Boone	118811	10%	11881	5%	594
Bourbon	19985	10%	1999	5%	100
Boyd	49542	10%	4954	5%	248

The feasible range of θ is $[0, \sqrt{n}]$ and n is the number of the decision variables or the number selected counties. In this case study, we select ten counties, so the feasible range of θ in this case is $[0, \sqrt{10}]$ ($\sqrt{10} \approx 3.16$). As we discussed in Chapter 3, we have known that the value of θ can be understood as a trade-off between less risk averse and more risk averse. When the volume of the ellipsoid is larger, it means that the feasible region of solution is larger. The feasible region contains more worse scenarios and thus the solution generated by the RO model is more conservative.

Suppose only one delivery is needed, depending on the amount of available resource and estimated demand, two scenarios that could happen in reality: first, when the available resource is more than the estimated demand, the allocation plan will reserve a certain amount

of resource as a safe cushion. In addition, adjusting the value of θ will be able to change the conservativeness of the results that generated by the RO models (see the results in 4.2, 4.3, 4.4). Table4.2 shows the allocation plans when available doses of a certain type of vaccine is equal to 40000 and also when the θ is 3, 2.5, and 1 respectively. We repeat the same experiment with different available doses i.e.: 30000, 20000, and 10000. From the results, we can see that the allocation plans are affected by the value of θ and "available doses". Increasing the amount of available doses of vaccines is not necessarily changing the allocation plan, especially when the available doses are above the absolutely necessary amount.

From Table 4.2, 4.3, and 4.4, we can see even though the amount of available doses is decreased from 40000 to 30000, and then 20000, the vaccines allocation plan does not necessarily change because the amount of available resources is still higher than the absolutely necessary amount. However, when the amount of available doses is decreased to 10000, the vaccines allocation plan changes when θ = 1, because the available doses is no longer enough to satisfy the absolutely necessary amount.

In general, sensitivity analysis helps to determine how sensitive the optimal solution is to the changes in data values, which includes analyzing changes in an objective function coefficient and a right hand side value of a constraint. For the vaccines allocation problem, we can utilize the sensitivity analysis to determine how different amounts of available resources and θ would change the rate of the reserved doses by the available doses.

Figure 3.1 summaries the ratio of reserved resources to the available capacity of resources against the available capacity of resources. The available capacity of resource are changed from 10000 to 100000. The experiment is tested with different θ values: 3, 2.5 and 1. The graph once again demonstrates one of our conclusions that adjusting the value of θ will

Table 4.2. One Time Period Vaccines Allocation Plan Result #1

Available doses	40000		
θ	3.0	2.5	1.0
Adair	373	466	746
Allen	399	499	798
Anderson	428	536	857
Ballard	165	206	330
Barren	843	1054	1687
Bath	232	290	464
Bell	574	717	1148
Boone	2376	2970	4752
Bourbon	400	500	799
Boyd	991	1239	1982
Released doses	6781	8477	13563
Released %	16.95%	21.19%	33.91%
Reserved%	83.05%	78.81%	66.09%

change the conservativeness of the results. As the value of θ grows, the feasibility of the uncertainty set expands as well. The more "worst-case" scenarios will be incorporated by the model, the more conservative the solutions are likely to become. As we can see in the graphic 3.1, when the amount of the available doses is fixed, the model tends to reserve more resource when the value of θ grows bigger. In practice, we prefer $θ \in [0.5, 1.5]$ to remove the concern of risk sensitiveness. In the following case study, we will adopt $θ = 1$ in our model and we will no longer consider the effect of θ on the results any more. Another reason that we select $θ = 1$ is because when the user changes the number of selected counties, we do not need to worry if the value of θ is valid or not. For example, if the user only selects four counties, and if we set the value of $θ = 2.5$, it may not give the user a valid solution since the default value of θ is out of the valid range for this instance. Therefore, if we set the default value of θ as one, it will always be valid no matter how many counties

Table 4.3. One Time Period Vaccines Allocation Plan Result #2

Available doses	30000		
θ	3.0	2.5	1.0
Adair	373	466	746
Allen	399	499	798
Anderson	428	536	857
Ballard	165	206	330
Barren	843	1054	1687
Bath	232	290	464
Bell	574	717	1148
Boone	2376	2970	4752
Bourbon	400	500	799
Boyd	991	1239	1982
Released doses	6781	8477	13563
Released %	22.60%	28.26%	45.21%
Reserved%	77.40%	71.74%	54.79%

that the end user selects. If for some reason, the user wants to change the default value of θ, the user can always log in as the administer to change the default value, if the administer role is authorized.

All in all, this session gives the readers an idea of how our proposed methodology works for a one time period resource allocation. More importantly, it shows the effect of θ on the conservativeness of the RO model and suggest a default value of θ for practice use. In the next session, we will continue to explore the performance of our proposed methodology on multiple time periods resource allocation.

Table 4.4. One Time Period Vaccines Allocation Plan Result #3

Available doses	20000		
θ	3.0	2.5	1.0
Adair	373	466	746
Allen	399	499	798
Anderson	428	536	857
Ballard	165	206	330
Barren	843	1054	1687
Bath	232	290	464
Bell	574	717	1148
Boone	2376	2970	4752
Bourbon	400	500	799
Boyd	991	1239	1982
Released doses	6781	8477	13563
Released %	33.91%	42.39%	67.82%
Reserved%	66.10%	57.62%	32.19%

4.2. Numerical Results for Multiple Time Periods Resource Allocation

In this session, we investigate the performance of our proposed methodology on "Multiple Time Periods Resource Allocation" due to the difference in models that will be ran. We will start from two time periods resource allocation and then increase to three time periods, due to the similarity between multiple time periods resource allocation. We think numerical results from the two time periods resource allocation and three time periods should be enough to give the reader an overall perspective of how to utilize our methodology and what to expect from the results.

Two Time Periods Resource Allocation

Table 4.5. One Time Period Vaccines Allocation Plan Result #4

Available doses	10000		
θ	3.0	2.5	1.0
Adair	373	466	447
Allen	399	499	484
Anderson	428	536	527
Ballard	165	206	176
Barren	843	1054	1203
Bath	232	290	258
Bell	574	717	750
Boone	2376	2970	4202
Bourbon	400	500	485
Boyd	991	1239	1466
Released doses	6781	8477	9998
Released %	67.81%	84.77%	99.98%
Reserved%	32.19%	15.23%	0.02%

If the number of distribution is more than one, for each time period except the last one, the RR model will need to run both the RO and SP models for solutions. The RO model will determine the absolutely necessary amount to be distributed for each time period and the SP model will further determine how to allocate the absolutely necessary amount of resources to each selected county in order to achieve better overall performance. In the last time period, only the SP model will be used to determine how to allocate all the remaining resources. Most of the input parameters that the RR model takes in are listed in Table 4.6. The value of θ that we use in the RO model is 1. Since we already discussed the effect of adjusting the value of θ in last session, we only need to demonstrate how the resource allocation plans change as the available doses change.

Table 4.6. Inputs to RR models for two time period resource allocation

	Population	Time Period 1				Time Period 2			
		PIR	d_i^*	Est.Error	Δd_i	PIR	d_i^*	Est.Error	Δd_i
Adair	18656	10%	1866	5%	93	10%	1866	5%	93
Allen	19956	10%	1996	5%	100	10%	1996	5%	100
Anderson	21421	10%	2142	5%	107	10%	2142	5%	107
Ballard	8249	10%	825	5%	41	10%	825	5%	41
Barren	42173	10%	4217	5%	211	10%	4217	5%	211
Bath	11591	10%	1159	5%	58	10%	1159	5%	58
Bell	28691	10%	2869	5%	143	10%	2869	5%	143
Boone	118811	10%	11881	5%	594	10%	11881	5%	594
Bourbon	19985	10%	1999	5%	100	10%	1999	5%	100
Boyd	49542	10%	4954	5%	248	10%	4954	5%	248

Table 4.7 shows the results for two time periods vaccines allocation when the available doses are 10000 and 20000. When the amount of available doses are equal to 10000 and θ is equal to 1, the available resources barely meet the absolutely necessary amount for the first time period, so there are not many resources to be reserved for the second time period. When we increase the number of 20000, we can see that there are some vaccines for the second time period allocation. As we discussed earlier, the absolutely necessary amount for each time period is generated by the RO model and that the factors that influence the results, including the population size of selected counties, PIR, Estimated Errors, and θ. After the absolutely necessary amount is determined, the SP model will take in this amount as the available dose, and then further determine how to allocate the available does to each selected county in order to achieve better overall performance. Since there is no reason for us to reserve any resources in the last time period, we will just apply the SP model to determine how to allocate all the remaining resources.

Table 4.8 is the results of RR models when the available doses increase to 30000, and 40000. We can get similar allocation plans for the time period #1, but since we have more available resources to allocate during the second time period, the SP model will generate different allocation plans based on the amount of available resources.

When we increase the available doses to 50000, and 60000, we notice that there are still some resources remaining even after the allocation for the second time period is done. This result may suggest that the available doses at hand are enough to cover the demands even under uncertain scenarios.

Table 4.7. Two Time Period Vaccines Allocation Plan Result #1

Available doses	10000		20000	
	Time Period #1	Time Period #2	Time Period #1	Time Period #2
Adair	48	0	67	21
Allen	53	0	76	22
Anderson	60	0	94	24
Ballard	0	0	13	10
Barren	868	0	1795	35
Bath	28	0	35	12
Bell	104	0	537	34
Boone	7266	0	8381	6195
Bourbon	53	0	76	22
Boyd	1518	0	2488	62
Released doses	9998	0	13562	6437
Reserved doses	2	2	6438	1
Released %	99.98%	0.00%	67.81%	99.98%
Reserved%	0.02%	100.00%	32.19%	0.02%

Table 4.8. Two Time Period Vaccines Allocation Plan Result #2

Available doses	30000		40000	
	Time Period #1	Time Period #2	Time Period #1	Time Period #2
Adair	67	419	68	1178
Allen	75	520	78	1325
Anderson	93	626	99	1488
Ballard	13	50	13	133
Barren	1793	2231	1801	3507
Bath	35	122	35	410
Bell	535	1152	551	2236
Boone	8388	7929	8350	10626
Bourbon	76	522	78	1328
Boyd	2488	2865	2489	4207
Released doses	13563	16436	13562	26438
Reserved doses	16437	1	26438	0
Released %	45.21%	99.99%	33.91%	100.00%
Reserved%	54.79%	0.01%	66.10%	0.00%

Three Time Periods Resource Allocation

Suppose three deliveries are planned in the beginning of allocation, and the "PIR" and the "Estimated Error" are the same as the previous two time periods. Depending on the available doses, the RR models will generate solutions accordingly. For example, when the amount of available doses is equal to 10000, there are no vaccines allocated to the second and third time period (see Table 4.10). This is because the available resources barely meet the absolutely necessary amount for the first time period. The results indicate it might not be a good idea to distribute the available resources in three different times under the current prediction of demands. If the prediction of demands still holds, the users may either seek

Table 4.9. Two Time Period Vaccines Allocation Plan Result #3

Available doses	50000		60000	
	Time Period #1	Time Period #2	Time Period #1	Time Period #2
Adair	67	2063	67	2618
Allen	76	2195	75	2754
Anderson	94	2345	92	2907
Ballard	13	920	13	1106
Barren	1796	4473	1791	5023
Bath	35	1342	35	1791
Bell	539	3087	530	3653
Boone	8380	12465	8397	12814
Bourbon	76	2198	75	2757
Boyd	2489	5234	2487	5770
Released doses	13565	36322	13562	41193
Reserved doses	36435	113	46438	5245
Released %	27.13%	99.69%	22.60%	88.71%
Reserved%	72.87%	0.31%	77.40%	11.29%

ways to increase the inventory or reduce the number of deliveries.

When we increase the available resources to 50000 or 60000, we can see there are enough resources for three deliveries in Table 4.11. The resource allocation plan for each selected county can be clearly seen from Figure 4.1.

It is also very interesting to compare the results generated by the RR heuristics with one delivery and two deliveries if the rest inputs are the same. Figure 4.2 presents the solutions from the RO model and the RR models side by side when the available dose is 10000 and $\theta = 1$. From the Figure 4.2, we can clearly see the results generated by the RO model is more conservative, because the solution spreads the risk to counties, which have less

Table 4.10. Three Time Periods Vaccines Allocation Plan Result #1

Available doses	10000			20000		
	Time Period #1	Time Period #2	Time Period #3	Time Period #1	Time Period #2	Time Period #3
Adair	46	0	0	68	20	0
Allen	51	0	0	77	22	0
Anderson	58	0	0	94	23	0
Ballard	0	0	0	13	11	0
Barren	869	0	0	1789	39	0
Bath	27	0	0	36	12	0
Bell	105	0	0	531	32	0
Boone	7271	0	0	8395	6168	0
Bourbon	51	0	0	77	22	0
Boyd	1519	0	0	2484	86	0
Released doses	9997	0	0	13564	6435	0
Reserved doses	2	2	2	6436	1	1
Released %	99.97%	0.00%	0.00%	67.82%	99.98%	0.00%
Reserved%	0.03%	100.00%	100.00%	32.18%	0.02%	100.00%

Table 4.11. Three Time Period Vaccines Allocation Plan Result #2

Available doses	50000			60000		
	Time Period #1	Time Period #2	Time Period #3	Time Period #1	Time Period #2	Time Period #3
Adair	65	67	738	67	66	1809
Allen	74	76	862	76	75	1935
Anderson	91	94	1007	94	92	2077
Ballard	13	13	115	13	13	800
Barren	1791	1796	3095	1794	1791	4089
Bath	34	35	120	35	35	1124
Bell	528	540	1746	537	531	2782
Boone	8405	8376	10511	8382	8397	11521
Bourbon	74	76	865	76	75	1938
Boyd	2489	2490	3814	2488	2487	4804
Released doses	13564	13563	22873	13562	13562	32879
Reserved doses	36436	22873	0	46438	32876	-3
Released %	27.13%	37.22%	100.00%	22.60%	29.20%	100.01%
Reserved%	72.87%	62.78%	0.00%	77.40%	70.80%	-0.01%

estimated demands. While, the SP model tends to allocate more resources to the counties, which have highest estimated demands in order to pursue better performance.

As we continue to increase the inventory to 70000 or 80000, there are unused resources appearing at the end of the third time period. If this happens, the user could either save the
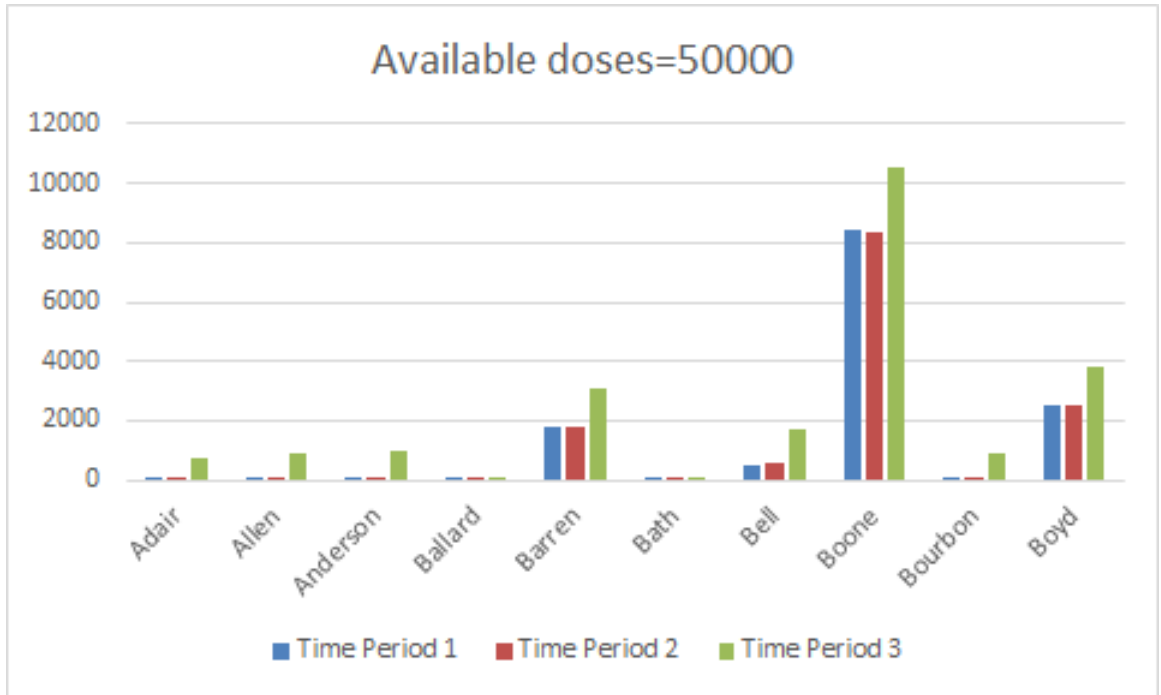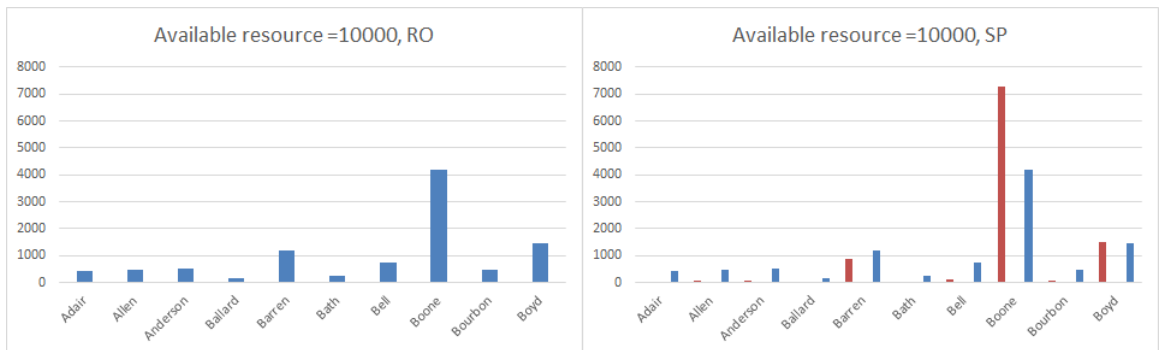
Figure 4.1. Vaccines allocation plan when available doses = 50000

(a) RO result when available doses = 10000, θ = 1(b) RR result when available doses = 10000, θ = 1

Figure 4.2. Compare RO and SP results with the same available dose and θ

inventory for future use or increase the value of the "PIR" of any desired time periods.

There are some advantages of the RR heuristics that we want to point out:

(1) The most obvious advantage of the RR heuristic is to make resource allocation plans with the consideration of uncertainty. The RO model helps to determine the "absolutely necessary amount" of resources to be allocated for each time period based on the " PIR", the "Estimated Error", and θ, which are the most important factors for constructing the uncertainty set. On the other hand, the SP model further determines how to allocate the "absolutely necessary amount" of resources to achieve better performance, due to the conservativeness of the RO models. The SP model also takes the uncertainty into consideration but requires the distributional information to be known.

(2) The RR heuristic can also help decision makers to determine a better resource allocation plan. It can be used for one time period resource allocation planning or multiple time periods resource allocation planning. Depending on the outputs, the users are able to tell if the available amount of resources are good enough for the pre-determined number of deliveries. If there are no resources to be distributed for some time periods, then the decision makers can utilize RR models to work out a more realistic plan to better use the available resources.

(3) The RR heuristic in combination of the web application allows users to select any affected counties. It is also very easy to apply the RR models to other states.

4.3. Model Interface on the Web Application

The web application is an effective way to allow the users in health-care field to comfortably utilize the models that are developed by industrial engineering researchers. It will help the health-care decision makers improve the decision effectiveness and make better decisions. Moreover, it will substantially reduce the decision cycle time. Some functions in the web application will give the decision makers more control on the outputs. For example,

Table 4.12. Three Time Period Vaccines Allocation Plan Result #3

Available doses	70000			80000		
	Time Period #1	Time Period #2	Time Period #3	Time Period #1	Time Period #2	Time Period #3
Adair	67	67	2475	68	66	2801
Allen	76	76	2610	77	75	2938
Anderson	95	94	2761	96	93	3090
Ballard	13	13	945	12	13	1175
Barren	1795	1795	4866	1795	1795	5198
Bath	35	35	1682	35	35	1922
Bell	539	538	3503	541	537	3835
Boone	8377	8382	12646	8373	8383	12937
Bourbon	77	76	2613	78	75	2941
Boyd	2488	2489	5610	2487	2490	5941
Released doses	13562	13565	39711	13562	13562	42778
Reserved doses	56438	42873	3162	66438	52876	10098
Released %	19.37%	24.04%	92.62%	16.95%	20.41%	80.90%
Reserved%	80.63%	75.96%	7.38%	83.05%	79.59%	19.10%

the decision makers can manually reserve a certain percentage of resources by selecting different risk levels. The snapshots of the web application are presented below and we will provide step-to-step instructions on how to utilize the application to generate the solution. Figure 4.3 provides an overview of this web application and from there the users can assess different models. If the user clicks "Planning and Analysis Tools", it will show all the tools

Figure 4.3. Home page of RTDSS web application

available under this category (see Figure 4.4).

Planning and Analysis Tools

Figure 4.4. Available Planning and Analysis Tools

Users will enter different tools by clicking different buttons. The button "SNS resource reservation" is the one we are going to demonstrate. If the user clicks this button, then s/he will be brought to this page (see Figure 4.5). This is the place where authorized users or administers log in, in order to access the model or change default parameters. Suppose the
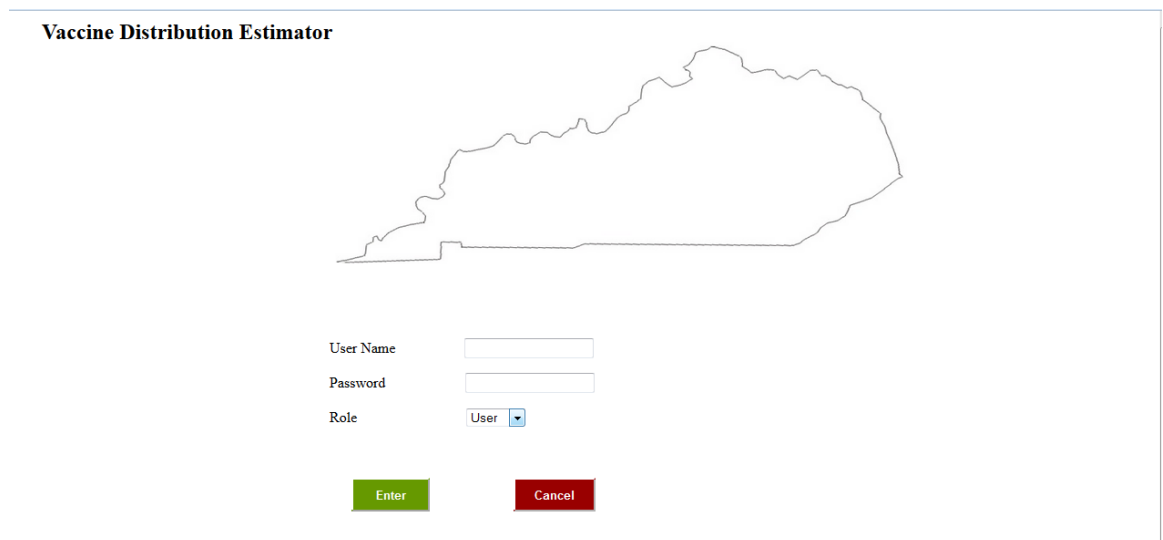
Figure 4.5. Vaccine Distribution Estimator - Page 1:User login

user logs in as a user, then s/he will see this page (see Figure 4.6). "Show Saved Allocation" is used when users want to revisit the solution generated previously. We are going to click "Create New Allocation" in order to demonstrate how to generate solutions from the beginning. After clicking "Create New Allocation" button, users should be able to see a
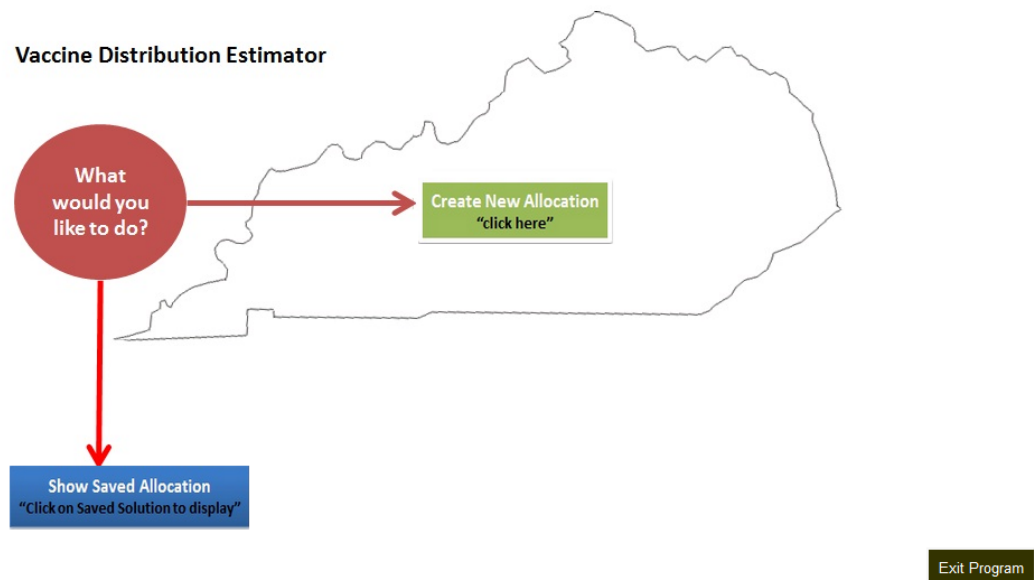
Figure 4.6. Vaccine Distribution Estimator - Page 2

page with blank input fields and instructions (see Figure 4.7). In the web application, there is a function that allows the end users to manually reserve resources from the available resources based on their knowledge of the risk level of the disease or disaster. By default, if the user selects "high risk", then 10% of the available doses of vaccines will be reserved for the future or the next time period. The RR heuristic will further determine how to allocate 90% of the available does of vaccines. Depending on the uncertain demands, it might reserve some resources from 90% along with 10% manually reserved. "Medium-high" means 80% of the available doses will be provided as the input to the RR models and 20% of the available doses will be reserved for the future or next distribution. "Medium-low" indicates 70% of the available doses will be sent to the RR models for determination and 30% will be reserved. "Low" indicates 60% of the available doses will the sent to the RR

Figure 4.7. Vaccine Distribution Estimator - Page 3: User Inputs

models for further determination and 40% will be reserved manually. If the end user logs into the web application as an administer, the administer can change the default setting. With that said, the administer is able to change the percentage of resources to be manually reserved for each risk level based on the type of the epidemics or diseases. This function gives the users more power and control on the allocation plans and hopefully the allocation plans will work better in reality.

All of the date entries need to follow mm/dd/year format. After all the fields have been filled as it shows in Figure 4.8, click "Set Demand". The application will bring the user to another page shown in Figure 4.9, and 4.10. The user can select the counties that require vaccines on this page. By default, all the counties are selected. When there are only a few counties that request vaccine supply, the user can uncheck the check mark on the left side of "County_name", so all the check marks by the counties will disappear. It will be easier for the user to select the counties that request vaccines. "Priority" is an option for the users to expedite the shipping date. If any selected county requests a fast shipping delivery

date, then the user should check this priority box and then the estimated shipping date will be expedited. There is another column called "Estimated Error", which measures the prediction accuracy and also are useful to generate the upper and lower bounds of predicted infectious populations. The default value of "Estimated Error" is 5% for all the counties and it can be changed by administers. After the regional demand is properly set, the user clicks "Save and Close" button and then the results will pop up.

Figure 4.8. Vaccine Distribution Estimator - Page 3: Inputs Are Filled

Figure 4.9. Vaccine Distribution Estimator - Page 4: User Inputs

Regional Demand

r_name	Vaccine_Doses_Requested	Select Priority	Expected_Error	Regional_id
air	18656	☐	10	112
en	19956	☐	10	113
rson	21421	☐	10	114
lard	8249	☐	10	115
ren	42173	☐	10	116
th	11591	☐	10	117
ell	28691	☐	10	118
one	118811	☐	10	119
rbon	19985	☐	10	120
yd	49542	☐	10	121
yle	28432	☐	10	122
ken	8488	☐	10	123
thitt	13878	☐	10	124
nridge	20059	☐	10	125
litt	74319	☐	10	126
ler	12690	☐	10	127

Figure 4.10. Vaccine Distribution Estimator - Page 4: User Inputs continued

CHAPTER 5

CONCLUSION

Public-health emergencies, such as bioterrorist attacks or pandemics, demand fast, efficient, large-scale dispensing of critical medical countermeasures (i.e., vaccines,drugs, and therapeutics). This large-scale dispensing is complex and requires careful planning and coordination from multiple federal, state, and local agencies. The multifaceted nature of dispensing and the complexity of the public-health emergencies make the process highly unpredictable. The impact of such events might be tremendous to the public health. Thus, it is necessary for emergency managers and public-health administrators to make quick and effective response as an emergency occurs.

In strategic planning, resource allocation is a plan for using available resources to achieve the goad or generate better performance. The optimization model is a powerful way to determine resource allocation in an efficient and effective fashion. Uncertainty cannot be ignored, especially in dealing with the pandemic and emergency disasters. We have shown that even a small uncertainty in the data can make the optimal solution to the problem completely meaningless in practice.

In order to improve KY SNS to better prepare for public-health emergencies and better utilize the scarce resources, we propose to include the data uncertainty in the demands as a factor to make strategic resource allocation plan. There are two popular methods that are capable of dealing with uncertain data in optimization, which are the stochastic programming and robust optimization. SP models are fundamentally based on the assumption

that the random variables' distribution information is known. However, the data in reality are more likely to be perturbed. When the data are incomplete, erroneous, there will be consistency and stability issues reported in the literature of stochastic programming. In contrast, RO models do not need to assume known distribution information. The negative impacts of unwanted perturbations are minimized for the RO models because the uncertain data is modeled as unknown-but-bounded variables without unrealistic assumptions. RO model is aimed to yield stable, feasible solution for the practical implementation. However the RO solution usually cannot achieve better expected objective than the corresponding SP solution. In our opinion, when the government agency is preventing the loss of life and optimizing the benefits from the vaccines allocation under uncertainty, neither RO nor SP models will be solely qualified. Therefore, we propose the RR heuristic to solve the vaccines allocation problem during the public-health emergencies.

Traditional resource allocation practices adopted by the State SNS during public health emergencies are based on the projected area's demands. If the resources are not enough to satisfy the demands, the resources will be allocated by the projected area's population size. Since the demands of some public health emergencies are highly unpredictable, allocating the scarce resources by the current policy may cause serious losses on the public health.

We propose a "Resource Reservation" heuristic to improve the accuracy of traditional decision making strategy by taking the consideration of uncertainty of demands. The basic concept of the RR heuristic is fairly straightforward: when the critical factor such as the amount of demands that affect the allocation plan is quite uncertain and changing all the time, people should be more conservative in deciding how to allocate the resources. In order to avoid that scenario that allocates all the resources and regrets it later, the RR heuristic suggests that we only allocate an absolutely necessary amount of resources. The resources

not being allocated can be treated as safety stock in logistics in case there will be uncertain demands in the future.

The resource reservation heuristic is a combination of robust optimization and stochastic programming with multiple iterations. The absolutely necessary amount of resources will be determined by the RO model with the unknown-but-bounded random variables and risk attitude as input information. Since the RO model can minimize loss under the "worst scenario", the corresponding decision from the RO model tends to be conservative. Therefore, the RO model is suited to make a rough allocation plan prior to possessing accurate information. The absolutely necessary amount will feed the SP model, which basically includes the total amount of resources to be allocated. By taking the amount of resources from the RO model, the SP model provides a plan on to optimally assign those resources to each request, with best expected objective.

By adjusting θ, we can either increase or decrease the perturbation and the corresponding results will be adapted to the adjustment. To be more specific, When $\theta = 0$, the RO model becomes a linear programming with nominal parameters. When $\theta = 1$, the worst-case scenarios are defined on the hull of the uncertainty set, i.e., the ellipsoid. The uncertainty set becomes the largest volume ellipsoid contained in $\mathcal{B} := \{d \mid |d_i - d_i^*| \leq \Delta d_i^*\}$. When $\theta = \sqrt{n}$, the uncertainty set enlarges to cover all the possible scenarios and the uncertainty becomes the minimum volume ellipsoid contained in $\mathcal{B}$. Thus, the value of θ can be understood as a trade-off between less risk averse and more risk averse. When the volume of the ellipsoid is larger, it means the feasible region of the solution is larger. Figure 3.1 further prove the theoretical conclusion that adjusting the value of θ will change the conservativeness of the results. As the value of θ grows, the feasibility of the uncertainty set expands as well. The more "worst-case" scenarios will be incorporated by the model, the more conservative the

solutions are likely to become. As we can see in Figure 3.1, when the amount of the available doses is fixed, the model tends to reserve more resource when the value of θ grows bigger.

In this dissertation, we also demonstrated how to incorporate uncertainty into stochastic programming models and robust optimization models by two examples: a farmer's problem and a portfolio selection problem. We made a thorough comparison of the SP and RO methods, and then brought the concerns on computational complex, stability issue of stochastic programming, and the conservativeness of robust optimization. All the concerns can be addressed by our proposed RR heuristic with solid theoretical proof.

After we demonstrated the RR heuristic, we applied this method to the KY SNS vaccines allocation problem. In order to help the public-health coordinators or administers easily access and utilize the models, our team developed a real-time decision support web tool. Authorized users can log in to the website, fill in the required information, run the models and the results can be shown on the web tool. The data collected from the users will be processed in certain way in order to satisfy the requirement by the model. For example, PIR is used to calculate the nominal demands of decision variables and it can be obtained by this equation (4.1). In addition, the users need to determine how many time periods that the available resources will be distributed. Depending on the number of time periods that the public-health coordinators or administers determine, the RR heuristic will run the RO model and/or SP model by the number of time periods. For example, if only one time period is determined, the RR heuristic will only call the robust optimization model once. If only the RO model is run, the solution tends to spread the risk among the decision variables, and therefore the result of the allocation plan will be rather conservative. If two or more time periods are determined, the RR heuristic will call the RO model first, in order

to calculate the amount to be reserved for each time period, and then call the SP model to further determine how to allocate the released resources in order to achieve better performances for each time period.

In order to evaluate the performance, we conduct numerical experiments on the RR models with testing data. We tested the RR models on one time period allocation, two time periods allocation, and three time periods allocation. With each number of time periods, we also gradually changed the number of available resources and observe the changes of the results. Additionally, for the one time period allocation, we also tested the effect of θ on the conservativeness of the results. The results can suggest how many resources need to be reserved under the "worst-case" scenario. Depending on the different combinations of the amount of available resources and estimated demands, the vaccine allocation plan varies. For example, sometimes there are a lot of resources to be reserved, and sometimes there will not be any resource to be reserved.

In our opinion, our research distinguishes itself as follows:

- Our proposed resource reservation heuristic has a wide application, especially for the complicated optimization problems involving uncertain parameters. In this research work, we apply our methodology to solve the Kentucky SNS vaccine allocation problem, which helps the public health emergency managers or coordinators make better decision on how to allocate scarce resources during the pandemic, or other types of public health emergencies.

- Only a few analytical models in the literature include the discussion of the effect of parameter of uncertainty on the vaccination policies. Some researchers apply the stochastic programming method with chance constraints to determine the vaccination policies, which requires the complete information of the epidemiological

characteristics of the virus. However, in reality, the epidemiological information is not necessarily available before the allocation decision has to be made. Most of the time, only the rough estimation of the demands will be known when the public health emergency occurs. Our proposed method does not require the complete epidemiological information or any distributional information to be exactly known. It is designed for the situation that limited information of disease is known and the estimation of the demands is under severe uncertainty. Since the models are designed for the severe uncertainty situation, the resource reservation heuristic will first determine the "absolutely necessary amount" of resources to release, and if there are resources left, then it will be reserved for the future use.

- The web tool developed by the RTDSS team is convenient for users to utilize the models without any mathematical background and industrial engineering knowledge. The user only needs to provide a few data points to make the model work, such as the amount of available resources, number of new cases, number of cased identified, and period of times, number of time periods during the whole planning horizon, and other general information associated with the outbreaks. The user can also select the projected counties, and then the allocation decision will only be made among those projected counties. The user can also manually reserve the available doses of resources by selecting "risk level" in the user input interface. In addition, all the default value that we use in the model is changeable by the administers.

- The model can also help the public health emergency managers or administers determine the number of distribution that the current available resources are good for. For example, in the later part of the dissertation, you will see there are no resource to be allocated for some multi-time periods allocation. In this case, the results suggest the number of time periods that the users pre-determined is too

high. Based on the results they get, the user can find out how many time periods that the available resources can be best utilized and distributed.

- In the dissertation, we also investigate the effect of an uncertainty parameter "θ" on the conservativeness of the resource allocation plan with an example. Changing this parameter allows the users to adjust the conservativeness of the allocation plan. But there is no certain linear relationship between the conservativeness and the value of "θ" and the feasible range of "θ" varies as the number of selected counties changes. Therefore, we determine to use 1 as the default value of "θ" since it works no matter how many counties that the users select.

- Our method perfectly preserves convexity and overcomes the stability issue of the stochastic programming, therefore it can be efficiently solved by modern interior point solvers on average desktop computers in milliseconds.

REFERENCES

S. Bansal, B. Pourbohloul, and L. A. Meyers. A comparative analysis of influenza vaccination programs. *PLoS Medicine*, 3:e387, 2006.

A. Ben-Tal. The entropic penalty approach to stochastic programming. *Math. Oper. Res.*, 10(2):263–279, 1985. ISSN 0364-765X.

A. Ben-Tal and A. Nemirovski. Robust solutions to uncertain programs. *Operations Research Letters*, 25:1–13, 1999.

A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88:411–424, 2000.

A. Ben-Tal, A.and Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:769–805, 1998.

J. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer-Verlag, Berlin, 1997a.

J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer-Verlag, New York, 1997b. ISBN 0-387-98217-5.

S. Boyd and L. Vandenberghe. *Convex Programming*. Cambridge University Press, 2004.

D. Clancy and N. Green. Optimal intervention for an epidemic model under parameter uncertainty. *Mathematical biosciences*, 205:297–314, 2007.

W. L. Cooper. Asymptotic behavior of an allocation policy for revenue management. *Operations Research*, 50(4):720–727, 2002.

G. Dantzig. Linear programming under uncertainty. *Management Science*, 1:197–206, 1955.

G. B. Dantzig and G. Infanger. Multi-stage stochastic linear programs for portfolio optimization. *Ann. Oper. Res.*, 45(1-4):59–76, 1993. ISSN 0254-5330.

J. Dushoff, J. B. Plotkin, C. Viboud, L. Simonsen, M. Miller, M. Loeb, and J. David. Vaccinating to protect a vulnerable subpopulation. *PLoS medicine*, 4:e174, 2007.

M. B. Harrington. Forecasting areawide demand for health care services: A critical review of major techniques and their application. *Inquiry*, 14:254–268, 1977.

P. Kall and S. Wallace. *Stochastic Programming*. Wiley, Chichester etc., 1994.

R. D. Kamenetzky, L. J. Shuman, and H. Wolfe. Estimating need and demand for prehospital care. *Operations research*, 30:1148–1167, 1982.

M. Kanellos. New life for moore's law. *CNET*, 2009.

E. P. Kao and G. G. Tung. Forecasting demands for inpatient services in a large public health care delivery system. *Socio-Economic Planning Sciences*, 14:97–106, 1980.

I. M. Longini Jr, E. Ackerman, and L. R. Elveback. An optimization model for influenza a epidemics. *Mathematical Biosciences*, 38:141–157, 1978.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

K. Moidu and O. Wigertz. Computer based information systems in primary health carewhy? *Journal of medical systems*, 13:59–65, 1989.

J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios. Robust optimization for large-scale systems. *Operations Research*, 43:264–281, 1995.

A. Nemirovski. Beyond the scope of interior point polynomial time methods: Simple methods for extreme large-scale convex programs. Technical report, Israel Institute of Technology, 2000.

A. Nemirovski. Advances in convex optimization: conic programming. *Proceedings oh the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, pages 413–444, 2006.

A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM J. Optimization*, 19:1574–1609, 2009.

Y. Nesterov and I. E. Nesterov. *Introductory lectures on convex optimization: A basic course*. Springer, 2004.

F. Pezzella, R. Bonanno, and B. Nicoletti. A system approach to the optimal health-care districting. *European Journal of Operational Research*, 8:139–146, 1981.

G. C. Pflug, A. Ruszczyński, and R. Schultz. On the Glivenko-Cantelli problem in stochastic programming: linear recourse and extensions. *Math. Oper. Res.*, 23(1):204–220, 1998. ISSN 0364-765X.

W. P. Pierskalla and D. J. Brailer. Applications of operations research in health care delivery. *Handbooks in OR & MS*, 6:4–7, 1994.

S. T. Rachev and W. Römisch. Quantitative stability in stochastic programming: the method of probability metrics. *Math. Oper. Res.*, 27(4):792–818, 2002. ISSN 0364-765X.

E. Rizakou, J. Rosenhead, and K. Reddington. Aidsplan: a decision support model for planning the provision of hiv/aids-related services. *Interfaces*, 21(3):117–129, 1991.

R. Rockafellar and R. J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.*, 16(1):119–147, 1991. ISSN 0364-765X.

W. Römisch. Stability analysis of stochastic programs. *Investigación Oper.*, 14(2-3):175–194, 1993. ISSN 0257-4306. Workshop on Stochastic Optimization: the State of the Art (Havana, 1992).

W. Römisch. Stability of stochastic programming problems. *Handbooks in Operations Research and Management Science*, 10:483–554, 2003.

W. Römisch and R. Schultz. Lipschitz stability for stochastic programs with complete recourse. *SIAM J. Optim.*, 6(2):531–547, 1996. ISSN 1052-6234.

W. Römisch and R. J.-B. Wets. Stability of ε-approximate solutions to convex stochastic programs. Stochastic Programming E-Print Series, http://www.speps.org , 2006.

A. Ruszczynski and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. Elsevier, 2003.

A. Shapiro. Asymptotic properties of statistical estimators in stochastic programming. *Ann. Statist.*, 17(2):841–858, 1989. ISSN 0090-5364.

A. Shapiro. On differential stability in stochastic programming. *Math. Programming*, 47(1 (Ser. A)):107–116, 1990. ISSN 0025-5610.

A. Shapiro. Asymptotics of minimax stochastic programs. Optimization Online, `http://www.optimization-online.org`, 2006.

M. W. Tanner, L. Sattenspiel, and L. Ntaimo. Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. *Mathematical biosciences*, 215: 144–151, 2008.

S. Vogel. Stochastic stability concepts. In *Operations research proceedings 1990, Pap. 19th Annu. Meet. DGOR, Vienna/Austria 1990, 57-64*, 1992.

R.-B. Wets. Challenges in stochastic programming. *Mathematical Programming*, 75(2): 115–135, 1996.

MATLAB CODE

```matlab
function []= RR_main(cap,num,t,distrate,PIR,esterror,MyArray,Kd)

%function []=  main (cap ,num ,t,PIR ,SF ,TNC1 ,NNC ,PT ,Stob ,Stov ,Etov )

%function []=  main (cap ,num ,t,Var ,PIR ,SF ,Kd )

% MUST  INPUT  PARAMETERS   TO RUN  THE  PROGRAM

% CAP =capacity ;The  number  of vaccines   to be  distributed

% N=NUM : NUMBER  OF TIME  PERIODS  for  distribution

% t =theta

% PIR =Pred_inf : Predicted   Infection   Rate

%size =1* n triangular   distribution   or input  by user

% Esterror =scalefactor  ; size =1* V

capacity = cap;  % the  quantity  of available  doses  of vaccincine

n = num; % number  of time  periods  in the  entire  planning   horizon

theta = t;  % risk  parameter   from  0.5 to 2

distr=distrate/100;    % percentage   of resource  to be  distributed

%in  the  first  time  period

%High :  90%  release ,medium -high  80%  release ,

%medium -low  70%  release  ,low  60%  release

PopulationArray=MyArray;

% MyArray =population   of vaccines -requested   counties ,

% size =  1* V, V=Var =number   of counties

Var=length(PopulationArray);

V=Var;
```

```matlab
Pred_inf=PIR;
% size : 1*n predicted  infection  rates  are generated  by user  input
m=ones(n,1);
scalefactor =(m*esterror)    '; %size  of  scalefactor   = V*n
population = PopulationArray    '; % size = V*1
rev_r = distr*capacity;
% resource  to be released  for  the  first  time  period.  size =1*1
K=Kd;
% the  number  of  realizations   of  random  variable ;
u_nom = population*Pred_inf;
% predicted  demands  of  resource  in  each  time  period  and  for  each  county ;
% size :V*n
u_range = u_nom .* scalefactor;
% range  of  demands  in  each  time  period  and  for  each  county ; size :V*n
if  n==1
    TP=1;
    % When  Time  Period  is  equal  to  1,
    % all  the  resource  must  be  allocated  to each  county.
    x=RO(V,theta,u_nom,u_range,rev_r,TP);
    x_rd= round(x);
    rev_r = capacity - min(rev_r, sum(x_rd));
    % Check  if there  is any  reserved  resource.
    dlmwrite( 'singlerelease.txt  ', x_rd, 'newline ', 'pc','-append ');
    % Released  resource  for  single  time  period  planing
    dlmwrite( 'singlereserve.txt  ', rev_r, '-append ');
    % Reserved  resource  after  planning
else
     for  TP = 1:n-1
            %u_n  = [u_nom ;zeros (1, n)];
```

```matlab
        %u_r = [u_range ;zeros (1, n)];
         %V_n =V+1;



%x = ROnew (V_n ,theta ,u_n ,u_r ,rev_r ,TP );

%rel_r = rev_r - round (x(V_n ,1));
    x=RO(V,theta,u_nom,u_range,rev_r,TP);
    rel_r = sum(round(x));
    % use RO model to determine the absolutely necessory demand
    %rev_r = rev_r - rel_r ;
    upperbd = u_nom(:,TP) + u_range(:,TP);
    lowerbd = u_nom(:,TP) - u_range(:,TP);



        y = SP(V,K,rel_r,u_nom,u_range,TP,upperbd,lowerbd);
        % use SP model to further determine how to
        % allocate the absolutely necessary demand
        y = round(y(1:V));
        rev_r= rev_r-sum(y)+ (1-distr)*capacity;
        %calculate the available resource for next round allocation
        distr=1;
        string1=strcat( 'multirelease ', num2str(TP));
        filename1 = strcat(string1,  '.txt ');
        string2= strcat( 'multireserve ', num2str(TP));
        filename2 = strcat(string2,  '.txt ');
        dlmwrite(filename1,y,  'newline ', 'pc');
        % Released resource from 1 to n-1 time period  Size :120*1
        dlmwrite(filename2,rev_r);
        % Reserved resource corresponding to each time period  Size :1*1
    end
```

```matlab
    for  TP=n


        % x=RO(V,theta ,u_nom ,u_range ,rev_r ,TP );

        % x_rd =  round (x);

        % rev_r  =  rev_r  - min (rev_r ,  sum (x_rd ));

        % rev_r  =  rev_r  - sum (x_rd );

         upperbd = u_nom(:,TP) + u_range(:,TP);

         lowerbd = u_nom(:,TP) - u_range(:,TP);

         y = SP(V,K,rev_r,u_nom,u_range,TP,upperbd,lowerbd);

         y = round(y(1:V));

        %rev_r  = rev_r  - min (rev_r ,  sum (y));

        % Check  if  there  is  any  reserved  resource.

        rev_r=rev_r-sum(y);

        string1=strcat(  'multirelease ', num2str(TP));

        filename1 = strcat(string1,    '.txt ');

        string2= strcat(  'multireserve ', num2str(TP));

        filename2 = strcat(string2,    '.txt ');

        %dlmwrite (filename1 ,x_rd ,'newline ', 'pc ');

        % Released  resource  for  nth  time  period  Size :120*1

        dlmwrite(filename1,y,   'newline ', 'pc ');

        dlmwrite(filename2,rev_r);

        % Reserved  resource  for  nth  time  period  Size : 1*1  (Should  be  0).

    end

    fprintf(1,   'Done ! \n');

end



function  x=RO(V,theta,u_nom,u_range,rev_r,TP)
```

100

```matlab
cvx_begin

variable x(V,1)

minimize(theta * norm(diag(u_range(:,TP))) -sum(x- u_nom(:,TP)));

subject to

sum(x)<= rev_r;

2*theta * u_range(:,TP)+x <= u_nom(:,TP)/2;

-x <= zeros(V,1);

cvx_end

end




function  x = ROnew(V_n,theta,u_n,u_r,rev_r,TP)

cvx_begin

variable x(V_n,1)

minimize(theta * norm(diag(u_r(:,TP)))-sum(x-u_n(:,TP)));

subject to

sum(x)<= rev_r;

x+2* theta * u_r(:,TP) <= u_n(:,TP)/2 ;

-x<=zeros(V_n,1);

cvx_end

end




function  y =SP(V,K,rel_r,u_nom,u_range,TP,upperbd,lowerbd)

A = zeros(V*K+1,V*(K+1));

A (1,1:V)=ones(1,V);

d = zeros(V*K+1,1);

d(1,1)= rel_r;   % release  the  released  resource  previous  allocated  by  RO  model
```

```matlab
for  k=0:(K-1);

    A((V*k+2):(V*(k+1)+1),1:V) = - eye(V);

    d((V*k+2):(V*(k+1)+1),1)= - round(lowerbd + (upperbd - lowerbd)* rand(1,1));

end

A(2:(V*K+1), (V+1):V*(K+1))= - eye(V*K);

dlmwrite( 'demands.txt ', d, 'newline ', 'pc','-append ', 'roffset ', 1)

f = zeros(V*(K+1),1);

f((V+1):V*(K+1), 1)= (1/(2*K))*ones(V*K,1);

lb=zeros((V+1)*K,1);

%lb=zeros (V);

y= linprog(f ',A,d,[],[],lb);

%y= linprog (f',A,d);

end
```

CURRICULUM VITAE

NAME:	Dongxue Bridgeman
ADDRESS:	Department of Industrial Engineering
	Speed School of Engineering, University of Louisville
	Louisville, KY 40292
DOB:	Shenyang, Liaoning Province, China - Jan.12, 1986

Education

- **B.E., Logistics Engineering**

  Beijing University of Posts and Telecommunications

  2004 - 2008

- **M.S., Industrial Engineering**

  University of Louisville

  2008 - 2009

- **Ph.D., Industrial Engineering**

  University of Louisville

  2008 - 2013

Awards & Honors

- **Graduate Research fellowship**

  University of Louisville

  August 2008 - March 2012; April 2013 - July 2013

Professional Membership

- **Student Chapter of INFORMS at University of Louisville**

- **Institute of Industrial Engineers**

- **APICS The Association for Operations Management**

Professional Experiences

- **Research Fellow**

  Rtdss Project, University of Louisville,

  August 2008 - March 2012; April 2013 - July 2013

Publications

- **"Rehabilitation Planning by Highway Contractors: A Robust Optimization Approach "**

  Lijian Chen, Qingbin Cui, Yujie Lu, and Dongxue Ma,

  Structure and Infrastructure Engineering