Design of Tactical and Operational Decisions for Biomass Feedstock Logistics Chain

Rahul Ramachandran

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial and Systems Engineering

Subhash C. Sarin, Chair
Douglas R. Bish
John S. Cundiff

June 15, 2016
Blacksburg, Virginia

Keywords: Biomass Logistics, Optimization, Simulation, Vehicle Routing Problem, Mixed-Integer Programming

Design of Tactical and Operational Decisions for Biomass Feedstock Logistics Chain

Rahul Ramachandran

(ABSTRACT)

The global energy requirement is increasing at a rapid pace and fossil fuels have been one of the major players in meeting this growing energy demand. However, the resources for fossil fuels are finite. Therefore, it is essential to develop renewable energy sources like biofuels to help address growing energy needs. A key aspect in the production of biofuel is the biomass logistics chain that constitutes a complex collection of activities, which must be judiciously executed for a cost-effective operation.

In this thesis, we introduce a two-phase optimization-simulation approach to determine tactical biomass logistics-related decisions cost effectively in view of the uncertainties encountered in real-life. These decisions include number of trucks to haul biomass from storage locations to a bio-refinery, the number of unloading equipment sets required at storage locations, and the number of satellite storage locations required to serve as collection points for the biomass secured from the fields. Later, an operational-level decision support tool is introduced to aid the "feedstock manager" at the bio-refinery by recommending which satellite storage facilities to unload, how much biomass to ship, how to allocate existing resources (trucks and unloading equipment sets) during each time period, and how to route unloading equipment sets between storage facilities. Another problem studied is the "Bale Collection Problem" associated with the farmgate operation. It is essentially a capacitated vehicle routing problem with unit demand (CVRP-UD), and its solution defines a cost-effective sequence for collecting bales from the field after harvest.

Design of Tactical and Operational Decisions for Biomass Feedstock Logistics Chain

Rahul Ramachandran

(GENERAL AUDIENCE ABSTRACT)

The global energy requirement is increasing at a rapid pace and fossil fuels have been one of the major players in meeting this growing energy demand. However, the resources for fossil fuels are finite. Therefore, it is essential to develop renewable energy sources to help address growing energy needs. Biofuels are one such renewable sources of energy that are gaining in importance. Over the past decade, there has been a surge in research on "Green energy" in the United States to reduce dependence on the foreign supply of fossil fuels and also to reduce greenhouse gas emissions. Several states in the USA now require a compulsory 10% ethanol blend with gasoline. A key aspect in the production of biofuel is the biomass logistics chain that constitutes a complex collection of activities, which must be judiciously executed for a cost-effective operation.

In this thesis, we introduce an approach to determine cost-effective biomass logistics-related decisions in view of the uncertainties encountered in real-life. This approach enables determination of tactical decisions after accounting for variability and dynamics of the system. The tactical decisions include number of trucks to haul biomass from storage locations to a bio-refinery, the number of unloading equipment sets required at storage locations, and the number of satellite storage locations required to serve as collection points for the biomass secured from the fields. The model considers both capital and biomass logistics operating costs to arrive at cost-effective decisions. Later, an operational-level decision support tool is introduced to aid the "feedstock manager" at the bio-refinery by recommending which satellite storage facilities to unload, how much biomass to ship from those storage facilities, and how to allocate and use existing resources i.e. trucks and unloading equipment sets during each time period.

We also introduce the "Bale Collection Problem" that refers to an operational-level decision encountered in the farmgate operation. The solution to this problem defines a cost-effective sequence for collecting bales from fields after harvest. Though some argue that an experienced operator does a good job in judging the sequence in which to pick up bales at minimum operating cost, the problem becomes complicated when multiple vehicles operate in a single field or when the visibility is poor, for example during night-time operation.

Dedication

I dedicate this thesis to my mom and dad, who always encouraged me, stood by me and prayed for me.

Acknowledgments

“ You have to dream before your dreams can come true.” - A.P.J. Abdul Kalam

First and foremost, I offer my sincerest gratitude to my advisor, Dr. Sarin who has supported me throughout my thesis research with his patience and knowledge while allowing me the room to work in my own way. I would like to record my gratitude for his supervision, advice and guidance from the very early stage of this work and patiently correcting my work. I have never met someone who is more mindful and conscientious as him.

I would like to thank my committee members Dr. Cundiff and Dr. Bish for their invaluable guidance, and support. A warm felt thank you goes out to Dr. Cundiff for using his precious time to teach me basics of biomass feedstock logistics and in reading the thesis, and providing valuable comments about it.

I would like to thank Antares Group for providing me with images of different biomass logistics operation for use in this thesis. I would also like to thank my lab-mate, Maichel for helping me learn to use CPLEX concert technology. He was always accessible and willing to help with his wide knowledge.

Words fail me to express my appreciation to my family, without whose encouragement and understanding the completion of my thesis would have been impossible.

I have to thank my beloved, Lakshmi, who have always been there for me, cheering me and standing by my side through the good and bad times.

I also wish to thank my friends, and roommates, who were always willing to help and support during the course of my graduate study. You all gave me a home away from home, here at Blacksburg.

Above all I praise the God Almighty who has been an unfailing source of strength, comfort and inspiration to me throughout my work.

Contents

1 Motivation and Background **1**

1.1 Motivation 1

1.2 Background 3

1.3 Problem description and research objectives 4

1.4 Outline of chapters 7

1.4.1 Chapter 2: A detailed description of biomass feedstock logistics chain 7

1.4.2 Chapter 3: A taxonomic review - Biofuel supply chain problems . . . 8

1.4.3 Chapter 4: A two-phase optimization-simulation approach for tactical decisions in a biomass feedstock logistics chain 8

1.4.4 Chapter 5: An operational-level decision support tool for biomass feedstock logistics 9

1.4.5 Chapter 6: Bale collection problem 10

2 A Detailed Description of Biomass Feedstock Logistics Chain **11**

2.1 Introduction 11

2.2 Overview of biofuel supply chain 12

2.3 Biomass feedstock logistics chain for the Southeast United States 13
2.3.1 Farmgate operation 14
2.3.2 Highway hauling operation 19
2.3.3 Receiving-facility operation 20
2.4 Biomass feedstock logistics chain for the Midwest United States 22

3 A Taxonomic Review - Biofuel Supply Chain Problems 24

3.1 Introduction 24
3.2 Classification framework 25
3.2.1 Inherent optimization problems in the biofuel supply chain design . . 27
3.2.2 Modeling methodology : Mathematical programming 30
3.2.3 Modeling methodology: Simulation 32
3.2.4 Modeling methodology: Simulation-optimization 33
3.3 Results and Discussion 34
3.3.1 Inherent optimization problems in the biofuel supply chain design . . 34
3.3.2 Modeling methodology 38
3.4 Conclusion 48

4 A Two-Phase Optimization-Simulation Approach for Tactical Decisions in a Biomass Feedstock Logistics Chain 50

4.1 Introduction 50
4.2 System description 52
4.3 Problem statement and literature review 54

4.4 Proposed modeling methodology 58
4.4.1 Phase 1: Optimization module 68
4.4.2 Phase 2: Simulation module 71
4.5 Results and Discussion 79
4.5.1 Results 79
4.5.2 Properties of the solution obtained from the proposed methodology . . 86
4.5.3 Discussion 90
4.6 Conclusion 91

5 An Operational-Level Decision Support Tool for Biomass Feedstock Logistics 92

5.1 Introduction 92
5.2 Need for an operational-level decision support tool 93
5.3 Problem statement and model formulation 96
5.4 Implementation and results 100
5.5 Conclusion 106

6 Bale Collection Problem 107

6.1 Introduction 107
6.2 System description 108
6.3 Problem Statement and literature review 110
6.4 Model formulation 113
6.5 Methodolodgies explored 118

6.5.1 Column generation 118

6.5.2 Cutting plane method 120

6.6 Results and discussion 121

6.7 Conclusion 126

7 Concluding Remarks and Directions for Future Research **127**

Bibliography **129**

Appendix A Input modeling **144**

List of Figures

2.1 Biofuel Supply Chain. 12

2.2 Biomass feedstock logistics chain for the Southeast United States. 13

2.3 Mowing Operation. (*Source: Antares Group*) 14

2.4 Baling Operation. (*Source: Antares Group*) 14

2.5 In-field hauling of large rectangular bales. (*Source: Antares Group*) 17

2.6 Single-layer ambient storage of round bales. (*Source: Picture by author*) . . 18

2.7 Side-loading rack option. 19

2.8 Process of loading bales on the rack. 20

2.9 Recieving-facility operations. 21

2.10 Midwest biomass feedstock logistics chain. 22

3.1 Classification framework for the taxonomic review. 26

3.2 Distribution of inherent optimization problem in the publications reviewed. . 36

3.3 Distribution of publications based on different solution approaches used. . . 42

3.4 Distribution of publications applying different modeling approaches. 45

4.1 Framework of two phase optimization-simulation approach. 66

4.2 A flowchart of the proposed two-phase optimization-simulation approach. . . 67

4.3 Process logic of S-BLCP-P2. 73

4.4 Process logic of SSL operation in the simulation model. 74

4.5 Process logic of highway hauling and receiving facility operation in the simulation model. 75

4.6 Replication analysis. 78

4.7 Two-dimensional picture displaying SSL(s) and bio-refinery at Gretna, VA. (Source: Liu et al. [79], InTech, Open Access, 2013) 79

4.8 Probability of meeting demand. 82

4.9 Robustness of solution with associated capital investment. 83

4.10 Performance of biomass logistics chain over the entire planning horizon. . . . 84

4.11 Weekly performance of the biomass logistics chain across different iterations. 84

4.12 Average productivity factor for load-out equipment sets. 85

4.13 Observed capacity of tractor-trailers. 85

4.14 SSLs with equal inventory equidistant from bio-refinery. 87

4.15 SSLs with unequal distance from the bio-refinery with unequal inventory. . . 88

5.1 Components of a decision support tool. 93

5.2 Flowchart of the biomass logistics desion support tool. 95

5.3 Features of the BL-DST in a time-expanded network. 101

5.4 Biomass deilvered to bio-refinery for orginial and BL-DST schedules. 103

5.5 Biomass deilvered to bio-refinery for orginial and BL-DST schedules. 104

5.6 A sample output report for BL-DST. 105

6.1 Bale collection. (*Source: Antares Group*) 109

6.2 An ariel depiction of bale distribution and field-side storage in a production field. 110

6.3 Features of the bale collection problem. 111

6.4 Standard benchmark data set with 48 nodes. 122

6.5 BCP data set with 48 nodes. 122

A.1 Parameter estimation. 145

A.2 Goodness of fit test. 146

A.3 Q-Q plot. 146

List of Tables

3.1 Publications classified based on an inherent optimization problem sub-structure. 35

3.2 Publications classified based on the mathematical programming approach used. 39

3.3 Publications classified based on the solution approaches used. 43

3.4 Publications classified based on modeling approaches. 46

4.1 Parameters evaluated from simulation model. 61

4.2 A comparison between direct solution by CPLEX and B&P approach [1]. . . 66

4.3 Standard library of objects in Simio. 72

4.4 Biomass inflow as a percentage of total annual harvest. 80

4.5 Parameters used for simulation of 199 SSL dataset. 81

4.6 Fleet size across different iterations. 83

4.7 Comparison of fleet size for start of operation and a matured operation. . . . 86

4.8 Number of load-out equipment sets and tractor-trailers required for Case 1. . 87

4.9 Number of load-out equipment sets and tractor-trailers required for Case 2. . 89

5.1 Other parameters and their values used in BL-DST. 102

6.1 BCP problem instances. 123

6.2 Computational results for direct application of CPLEX on BCP-C instances. 124

6.3 Computational results for cutting plane-based methodology on BCP-C instances. 125

Chapter 1

Motivation and Background

1.1 Motivation

The global energy demand has been constantly increasing at the rate of 5.3% each year [98], and fossil fuels have been one of the major players in meeting this growing energy demand. However, the resources for fossil fuels are finite. By considering the available fossil fuel reserves, it is estimated that fossil fuel products like oil and coal will last only for about 35 and 107 years, respectively [105]. Therefore, it is essential to develop renewable energy sources to help address growing energy needs. Biofuels are one such renewable source of energy that are gaining in importance. They can be blended with fossil fuels to produce alternative fuels like ethanol-blended-gasoline. Several states in the USA now require a compulsory 10% ethanol to be blended with gasoline. Besides, biofuels contribute less to greenhouse gas emissions. Biofuels are essentially made from plants or crops, which consume carbon dioxide for photosynthesis, thereby reducing net greenhouse gas emissions.

Biofuels are generally divided into four generations. The first generation of biofuels are essentially made from feedstock which are traditionally used for food and feed. The production and use of first generation biofuels has become controversial due to the use of food crops for fuel production [88]. The second generation of biofuels has been developed to overcome shortcomings of first generation biofuels. They are produced from non-food crops like wood, switchgrass, or agricultural residues like corn stover and wheat straw. Cellulosic ethanol is one of the major second generation biofuels, which is produced from cellulose, a major component of cell walls in plant cells. Second generation biofuel feedstock cost is higher due to additional processing requirements like shredding and transportation requirements for the relatively low bulk density materials. Considerable research has been carried out to minimize feedstock cost and thereby make the cost of biofuels more competitive with petroleum fuels. Third generation biofuels, similar to the second generation biofuels, are produced from non-food feedstock like algae. The feedstock algae can be genetically modified to produce a wide range of hydrocarbons like ethanol, methane, butanol, jet fuel, among others. Third generation of biofuels are also known as advanced biofuels, or green hydrocarbons. Fourth generation biofuels are designed to be carbon negative, i.e., they produce sustainable energy by capturing and storing carbon dioxide from the environment and converting it into fuels.

The production of ethanol in the United States has rapidly increased in the last decade making it the world's largest producer of ethanol. The United States Department of Energy is targeting a capacity of 60 billion gallons of ethanol by 2030 (De La Torre Ugarte et al. [34]). This is mainly to reduce the reliance of imported crude oil, reduce greenhouse gas emissions (and the subsequent impact of global climate change), and to improve rural economies. It is estimated that 35-60% of the cost of biofuel is contributed by biomass logistics (Fales et al. [46]) leaving very little margin for farmers and bio-refinery operators. To make biofuels an economically viable alternative, biomass feedstock logistics must be designed and operated cost effectively.

1.2 Background

Many high technology product industries, like personal computers, electronics and the automotive industry, take advantage of many operations research tools for effectively designing and managing their supply chain. Over the past decade, the biofuel industry has also seen considerable increase in the application of operations research tools to improve its supply chain, but not to an extent used in other industries. This can be mainly attributed to reasons like a different organization of the agricultural industry as compared to high technology product industries. Also, the uncertainty due to weather impacts in agriculture is greater than the rest, requiring stochastic models which are difficult to solve. Despite these facts, there is a huge effort to improve supply chain management in biofuel industry, especially in the area of biomass feedstock logistics.

Many farm owners in the Piedmont region in Virginia grew tobacco for the cigarette industry before decline in the tobacco market. These farms can now grow biomass feedstock, like switchgrass, to become economically stable and also to help ramp up the United States' biofuel production. Specially designed biomass logistics systems are necessary for the Southeast region as they have many small-sized farms (less than 80 hectares) unlike the Midwest region where farms of 250 hectares area are common.

Switchgrass is harvested and converted into round or rectangular bales. An effective operation in this regard requires addressing several challenges including contracting with farmers, determining the logistics capability required, and managing the logistics equipment effectively, among others. Recently, Judd et al. [62] developed a mathematical model to determine optimal locations for temporary storages called Satellite Storage Locations (SSL) considering locations of potential production fields and a potential bio-refinery. The goal was to achieve the most cost effective operation possible. The authors consider a novel logistics design involving temporary storage locations, sharing of loading equipment sets among these storage sites and use of side-loading racks to transport bales from temporary storage locations to the bio-refinery. This model is a mixed integer linear program and does not consider seasonality in the availability of biomass from production fields and uncertainties in logistics equipment operations.

In this thesis, we expand on this novel biomass logistics methodology by including uncertainty arising in the availability of biomass from production fields and in the availability of resources because of variation in speed, loading (unloading)times, and equipment breakdowns. To address this uncertainty, we employ a simulation-optimization based methodology. We designate the underlying problem as the ***Stochastic - Biomass Logistics Chain Problem*** **(S-BLCP)**. The methodology developed does not require mathematical modeling of stochastic problems, and it provides an implementable solution with a reasonable computation time. We then introduce an operational-level decision support tool ***Biomass Logistics Decision Support Tool*** **(BL-DST)** to help a feedstock manager at the bio-refinery to make effective real-time decisions. Finally, we introduce and formulate mathematical models for the ***Bale Collection Problem*** **(BCP)** in order to provide a sequence for collecting switchgrass bales from a production field that minimizes cost.

1.3 Problem description and research objectives

The biomass logistics problem consists of a myriad of complex sub-problems pertaining to the various functionalities involved. In this section, we provide a brief overview of the problems that we address in this research, research objectives and the contributions that we make.

Problem Description:

S-BLCP: Given a set F of pre-located SSLs, expected inflow of biomass during each time period t at each SSL i, A_i^t, planning horizon of length T, and variability in the operational parameters of logistics equipment, determine minimum fleet size of logistics equipment, amount of biomass shipped from SSLs to a bio-refinery, routing of mobile equipment during each time period, so as to minimize the total cost incurred (consisting of capital investment and operational costs) and maximize the amount of biomass delivered considering the uncertainty and dynamics of the system.

BL-DST: Given a set F of pre-located SSLs, promised inflow of biomass during each time period t at each SSL i, A_i^t, planning horizon of length T, and operational parameters of logistics equipment, current inventory levels y_i^0 and location of logistics equipment, determine: (1) routes of mobile equipment during each time period to load-out biomass at the SSLs, and (2) amount of biomass to be shipped from SSLs to the bio-refinery, so as to (1) minimize the mobilization cost for the load-out equipment sets, and (2) total hauling cost for feedstock delivered at the bio-refinery. Also, a third goal was to minimize the risk of not meeting the desired demand of the bio-refinery.

BCP: Given a set, N, of bale locations in a production field and a set, D, of farm-side storage depots (refereed to as SSLs), a non-negative cost $C_{i,j}$ associated with traveling from locations i to j, capacity c of a in-field hauling vehicle, determine an optimal sequence to collect bales from the production field in order to minimize the total cost incurred. We assume $C_{i,j} = C_{j,i}$ making this problem a symmetric capacitated vehicle routing problem with unit demand (SCVRP-UD). A vehicle routing problem (VRP) is concerned with determining optimal routes to visit a set of customers by a fleet of vehicles that are based at one or more depots. The capacitated vehicle routing problem with unit demand (CVRP-UD) is a special case of the basic VRP in which there exists a limit on the maximum number of customers that can be visited in each trip. The CVRP has been extensively studied in the literature and many exact algorithms and heuristics approaches have been devised for its solution.

We address the above problems in the following chapters. But, first we present, in Chapter 2, a detailed description of the biomass logistics system, its components and important decisions associated with it. Then, in Chapter 3, we review the work reported in the biomass logistics area and classify this work based on the modeling and solution approaches used, for a variety of biomass logistics problems, and highlight opportunities for further work in this area. In

Chapter 4, we discuss the S-BLCP and present a new method for its solution. In Chapter 5, we present the methodology behind the BL-DST. Chapter 6 then introduces the BCP along with its extensions, and presents a capacitated vehicle routing problem-based formulation for its solution.

Research objectives: The main objective of this research is to understand, model, and analyze some important problems encountered in the biomass feedstock logistics chain, and create an effective decision support system. Some specific objectives are as follows:

1. Understand biomass feedstock logistics chain, its components and major challenges.
2. Highlight potential areas in biomass feedstock logistics that can be studied using advanced operations research techniques.
3. Determine tactical decisions pertaining to operation of the biomass feedstock logistics chain in the face of uncertainty.
4. Develop a methodology to help make operational-level decisions.
5. Mathematically model and analyze the bale collection problem, and investigate a method for its solution.

Contributions of thesis research:

The major contributions of this thesis research are as follows:

1. Classification of biomass logistics problems based on the inherent operations research problem, to aid future research.
2. Development of a novel methodology to model uncertainty and dynamics of a biomass feedstock logistics system using a two-phase optimization-simulation approach.
3. Application of optimization-simulation approach to the real-life database of potential production fields within a 48-km radius of Gretna, Virginia, a potential location for a bio-refinery.

4. Development of an effective decision support tool for operational control of a biomass feedstock logistics system pertinent to the Piedmont physiographic region - a region across five southern states (VA, NC, SC, GA, AL).

5. Mathematical modeling of a capacitated vehicle routing problem and its solution methodology.

1.4 Outline of chapters

1.4.1 Chapter 2: A detailed description of biomass feedstock logistics chain

The biomass logistics chain problem (BLCP) is associated with transporting baled biomass from temporary storage locations, called satellite storage locations (SSL), to a bio-refinery for conversion into biofuel, which can be blended with fossil fuels in the downstream supply chain. In Chapter 2, we present a detailed description of the biomass feedstock logistics system under consideration, and also, present its advantages over the existing system. The capital investment and storage cost associated with round bales are two folds lower than that for the rectangular bales. It is also worth noting that the operational aspects of a round-bale logistics system are more challenging as compared with that for a rectangular-bale system [61]. Significant cost savings for the farmers and contractors associated with the biomass feedstock production and logistics system can be achieved by improving operational efficiencies of the round-bale biomass logistics system.

The logistics system under consideration is specifically designed for handling round bales of switchgrass biomass produced in the Southeast region of the United States. The system can be also deployed for fields in the the Midwest for the corn stover harvest, if appropriate design modifications are applied. The principle contribution of this chapter is two-fold. First, a detailed description of biomass feedstock logistics is presented to help understand the underlying problems and challenges associated with biomass feedstock logistics operation. Second, this chapter also forms a background for subsequent chapters where we discuss opportunities for applying operations research tools in biomass logistics, and for chapters where we have developed optimization and simulation models for a specific biomass logistics problem.

1.4.2 Chapter 3: A taxonomic review - Biofuel supply chain problems

In Chapter 3, we review different problems pertaining to the biomass feedstock logistics chain and identify the fundamental operations research problems involved. The selected publications are classified based on: (1) fundamental operations research (OR) problem, (2) modeling methodology, and (3) solution approach. The review of literature aims at identifying fertile avenues for future research that focuses on relevant applications of operations research tools to efficiently design and operate a biomass logistics system. We believe that this chapter will not only motivate practitioners to apply relevant models to assure economic viability of biofuel industries, but also, will motivate development of better solution approaches for the problems involved in the biomass feedstock logistics system.

1.4.3 Chapter 4: A two-phase optimization-simulation approach for tactical decisions in a biomass feedstock logistics chain

In Chapter 4, we propose a novel modeling approach to determine tactical decisions like fleet sizes of equipment and trucks by integrating simulation and optimization methodologies. The optimization model is a multi-period mixed integer linear programming model that minimizes the total capital cost of equipment and operational costs incurred for this equipment. Initially, the logistics equipment input parameters are fixed at their expected values for the optimization model. The simulation model accounts for uncertainty in operational parameters, unexpected delays and breakdowns of logistics equipment. The simulation model iteratively updates the operational parameters in the optimization model. This approach develops robust and cost-effective solutions in the face of uncertainties. It terminates at the desired solution in a finite number of iterations.

The proposed optimization-simulation modeling approach saves the user from developing and solving a large-scale stochastic mathematical model, and it provides an implementable, robust solution in the face of uncertainties. We demonstrate applicability of this methodology by applying it to a real-life production region in the Southeast United States (Gretna, VA).

1.4.4 Chapter 5: An operational-level decision support tool for biomass feedstock logistics

In Chapter 5, we first highlight the need for a decision support tool for operating a biomass feedstock logistics chain. The fleet sizes of trucks and load-out equipment sets for the planning horizon are determined by using the models presented in Chapter 4 before starting operation in a particular region. Several uncertain events are encountered during the harvesting of biomass. These pertain to weather, equipment availability, delay in processing, among others. Hence, the biomass inflow to the storage facilities cannot be determined with certainty ahead of time. The decision support tool developed considers current status of the system, which includes inventory level, current location of load-out equipment sets, inflow of biomass from fields to storage locations, and determines the storage locations to unload biomass from at any given time, amount of biomass to be shipped from each of these locations, allocation and routing of trucks and unloading equipment sets in order to minimize mobilization cost for the load-out equipment sets and hauling cost for biomass delivered at the bio-refinery. The planning horizon considered is two weeks, as the feedstock manager would have details of biomass inflow for this time interval with high confidence. The problem is mathematically modeled as a mixed integer linear program, implemented in Microsoft Visual Studio 12.0, and is solved using CPLEX 12.6 solver.

1.4.5 Chapter 6: Bale collection problem

In Chapter 6, we introduce "Bale collection problem" - an operational-level decision problem encountered in farmgate operations. The solution to this problem gives a cost-effective sequence for collecting bales from a farm after harvest. The problem can be modeled as a capacitated vehicle routing problem with unit demand (CVRP-UD), a special case of the vehicle routing problem (VRP). The VRP is NP-hard because it includes the Traveling Salesman Problem (TSP) as a special case (Laporte [72]). The coordinates of the bales in a field is obtained from a Global Positioning System (GPS) fitted on the baler. The location of each bale is recorded as it is ejected from the baler. Since distance is calculated from these coordinates, a symmetric cost $C_{i,j}$ is considered. The CVRP, in general, has been extensively studied and many exact algorithms and heuristics approaches have been devised for its solution. In this chapter, we first mathematically model the BCP as mixed integer programming model, then present several extensions of the BCP relevant to biomass feedstock logistics, and finally, explore some solution methodologies to obtain minimum cost solutions.

Chapter 2

A Detailed Description of Biomass Feedstock Logistics Chain

2.1 Introduction

The biomass feedstock logistics chain is an important part of the biofuel supply chain. The design and operation of the biomass feedstock logistics chain plays a major role in determining the price of biofuel. By some accounts, it is estimated that 35-60% of cost of biofuel is contributed by biomass logistics [46]. Biofuel industries use different feedstocks like corn, sugarcane, switchgrass, and corn stover, among others, based on geographic location and climatic conditions. Switchgrass and woody biomass are commonly used in the Southeast and corn stover in the Midwest. The type of biomass used defines the features of the biomass feedstock logistics chain. The chapter aims at providing a detailed understanding of various features of biomass feedstock logistics which serve as a base for the mathematical and simulation models discussed in later chapters.

2.2 Overview of biofuel supply chain

A supply chain comprises a complex network of activities used to deliver product and services to end customers through engineered flow of goods, information and finance. The biofuel supply chain consists of activities pertaining to farming, harvesting of biomass, conversion biomass to biofuel and delivery of biofuel to end customers. A biofuel supply chain can be compared with a petroleum-based fuel supply chain. The petroleum-based fuel supply chain is generally categorized into upstream, midstream and downstream supply chain activities. Upstream activities comprise petroleum exploration, extraction and transportation to crude-oil refineries. Midstream activities consist of petroleum refinement, and downstream activities comprise of processes after refining i.e., storage and distribution to end customers.

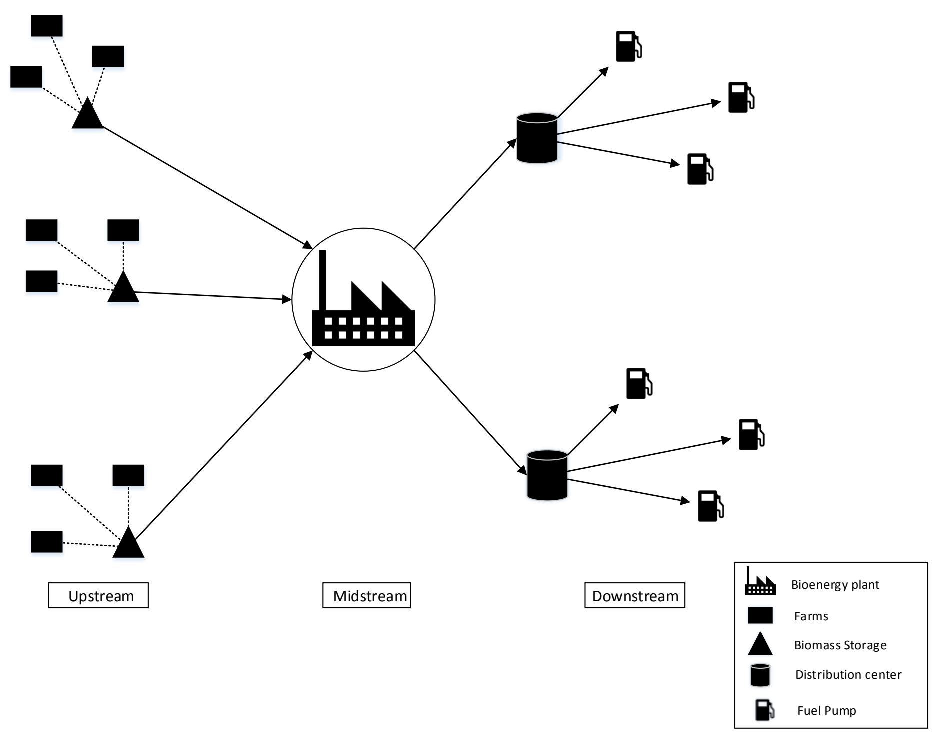


Figure 2.1: Biofuel Supply Chain.

The biofuel supply chain can also be similarly categorized to upstream, midstream and downstream activities. Upstream activities in the biofuel supply chain include farming, harvesting, storing, and transporting biomass to the bio-refinery. Midstream activities pertain to conversion of biomass to biofuel. Downstream activities include storage and distribution similar to the petroleum fuel supply chain. Upstream and midstream activities are the two popular areas of research in the biofuel supply chain. The biomass feedstock logistics chain is the term used here for the upstream activities of a biofuel supply chain. It includes activities pertaining to farming, harvesting, storing and transporting biomass feedstock to a bio-refinery.

2.3 Biomass feedstock logistics chain for the Southeast United States

The biomass feedstock logistics chain includes a collection of activities ranging from farming to transporting biomass to a bio-refinery. These activities have to be planned and executed judiciously for cost-effective production of biofuel. The key features of a biomass feedstock logistics chain operation can be categorized into: (1) Farmgate operation, (2) Highway hauling operation, and (3) Receiving facility operation (Liu et al. [79]).

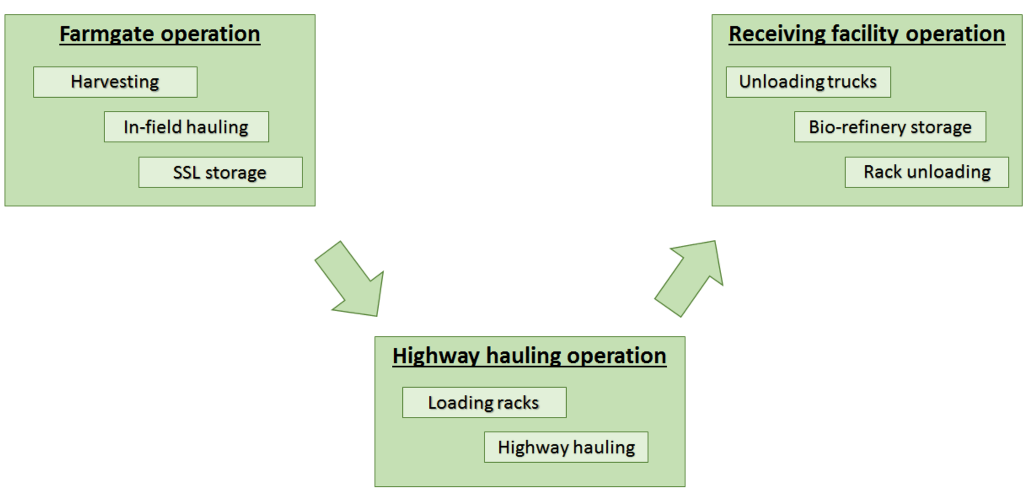


Figure 2.2: Biomass feedstock logistics chain for the Southeast United States.

2.3.1 Farmgate operation

The farmgate operation refers to activities such as production of biomass, harvesting, baling, and delivery of biomass to a satellite storage. Different biomass feedstock such as switchgrass or woody material can be used in the production of biofuels. This study considers the use of switchgrass for the production of biofuel in the Southeast United States.

2.3.1.1 Harvesting and Baling

The harvesting operation is the cutting and baling of switchgrass on the production fields. The switchgrass is cut with a mower-conditioner and later baled with a tractor pulling a baler. Switchgrass can be harvested in both short and extended harvesting time windows [53]. The short harvesting time window lasts for about three months where the aim is to delay the harvest till the switchgrass field enters into a dormant state. This harvesting strategy facilitates transfer of nutrients from the above-ground plant parts to below-ground plant parts thereby maximizing biomass yield and reducing the need for fertilizing production fields. Since the harvesting is completed in a short period of time, more harvesting equipment and large storage facilities are required.

Figure 2.3: Mowing Operation.
(*Source: Antares Group*)

Figure 2.4: Baling Operation.
(*Source: Antares Group*)

Another strategy is to extend harvesting over eight to nine months enabling effective use of harvesting equipment and also reducing need for large capacity storage sites. The delayed harvest of certain fields allows the switchgrass to dry standing in the field. It then can be cut and baled on the same day. The disadvantage of delayed harvest is some loss of yield due to leaf loss. The switchgrass, after being cut, is dried and later baled for densification. When the switchgrass has a moisture content greater than 15% when cut, it is dried laying in the field and baled later. Generally, biomass feedstock is baled into either a large round bale or a large rectangular bale depending on climate and the biomass logistics design. Both types of bales have their own advantages and disadvantages. It is more advantageous to use round bales in the Southeast United States due to the high annual rainfall. The round bales, made by wrapping continuous layers of switchgrass, shed moisture effectively when placed in single layer ambient storage. The round baler is generally operated with a smaller tractor, and the capital cost associated is much lesser for a round baler ($23,000) in comparison to a rectangular baler ($87,000). The round bale, which is almost 5 feet in diameter and 4 feet long is typically not transported over large distances due to difficulty in handling. An efficient "rack system concept" was proposed in Cundiff [29] to handle and transport round bales effectively. The rectangular bales, on the other hand, are more suitable for Midwestern and Western United States as these regions receive lower amounts of rainfall. The rectangular bales are made by pressing layers upon layers of biomass in a bale chamber and then tying the resulting bale with twine. When left in the field, the rectangular bale act like a sponge and absorb more moisture during a rainfall event. There is no practical way to remove this moisture from the bale. Rectangular bales are stored as stacks in a covered storage facility with special floor surface to prevent the bottom bales from absorbing moisture from the ground. A 12 ton-per-hour capacity rectangular baler is expected to cost around $87,000 against the cost of $23,000 for a 5 ton-per-hour capacity round baler [61]. The capital investment per ton-per-hour capacity clearly indicates that producing round bales is a more cost-effective densification option. But, the question remains, how to cost effectively haul the round bales.

The bales are protected from moisture to prevent microbial activity and subsequent dry matter loss. Several storage options are: (1) storing in a covered shed, (2) tarping bales in a pyramid stack, and (3) net wrapping bales and storing them in single layer ambient storage. Storing bales in a shed is the most effective option to limit dry matter loss, but it is expensive. Tarping bales is cheaper than the use of sheds, but it is labor intensive and is often unsafe as it involves workers climbing on stacks of bales. The most cost-effective option is to wrap bales in a polymer net and place them in single-layer ambient storage, if the bales do not require long storage before processing. The dry matter loss for net wrapping is more than that for shedding or tarping, but it is considered to be effective if the bales are not expected to be stored for too long (less than 6 months). Net wrapped bales can be left in the field for several days before being hauled to an SSL. This "uncoupling" of baling and in-field hauling gives an equipment scheduling advantage.

2.3.1.2 In-field hauling

The process of collecting and moving switchgrass bales from different locations on the field to an SSL is called in-field hauling. The bales are collected using an implement capable of lifting and placing bales on a flat-bed vehicle. The equipment has a pickup arm for self loading and a pusher arm to locate the collected bales on the vehicle. The capacity of this in-field hauling equipment can vary from six to twelve bales depending on the size. The in-field hauling equipment moves the bales from field to SSL where they are placed in ambient storage as a single layer. The in-field hauling equipment is designed to be operated in fields as well as paved roads. This equipment has limited road speed making it less efficient than a tractor-trailer truck for highway hauling. Hence, only a short-distance hauling to a storage site is carried out using this equipment. A transfer point is established in the logistics chain to uncouple in-field hauling and highway hauling. Bales from nearby production fields are brought to this transfer point and later transported to a bio-refinery in racks on tractor-trailers trucks. The farmgate contract signed by the feedstock producer includes a storage fee for the contract holders to establish and maintain an SSL. The contract holder informs the bio-refinery when feedstock is placed in the SSL, and the bio-refinery informs the contract holder when the feedstock is hauled out of the SSL.

Figure 2.5: In-field hauling of large rectangular bales. (*Source: Antares Group*)

2.3.1.3 Satellite storage location

A satellite storage location (SSL) is a pre-designated location on a production field for storing the biomass provided by fields within a defined radius of the location. It also acts like a transfer point between in-field hauling and highway hauling. These storage locations store bales from a single large field or from multiple smaller fields before it is hauled to a bio-refinery storage yard. The risks of fire and destruction of inventory are mitigated by distributing the entire bio-refinery inventory across different SSLs and by reducing inventory at the bio-refinery storage yard. An SSL incurs a one-time setup cost and an annual operational cost for the duration it is under contract. The one-time setup cost is for the construction of an SSL. The area designated for an SSL by the farmgate contract holder must be graded to a minimum slope specification, must be near a state-maintained road, and must have a compacted gravel base. The gravel base prevents bale degradation due to moisture absorption at the bottom of the bale, and also, it provides suitable surface for equipment operation. The locations of SSLs influence both the cost of in-field hauling and cost of transporting biomass to the bio-refinery. The number of SSLs to be opened and their locations are determined by balancing the setup cost required to create SSLs, the cost of transporting biomass from the field to the SSL, the cost of loading out the SSL, and the cost of hauling the biomass to a bio-refinery. It is probable that "custom contractors" will be employed by certain feedstock producers to haul their bales and place them in the SSL. These contractors can afford to invest in higher capacity industrial-grade equipment needed for high annual use operation. In this case, each farmgate contract holder will not have to own in-field hauling equipment.

Figure 2.6: Single-layer ambient storage of round bales. (*Source: Picture by author*)

Specialized equipment is used for loading-out biomass from the SSLs. This equipment can be either stationed permanently at an SSL, or it may be mobile and shared among a number of SSLs so that fewer sets of equipment are needed [62]. Biomass logistics chain design with mobile equipment reduces the cost for loading-out an SSL as capital cost of equipment is shared among multiple SSLs. Such a design may have fewer equipment sets for many small-sized SSLs compared to having dedicated equipment at each large-sized SSL. A multibale handling unit is proposed for highway hauling of round bales as individual handling of bales is not cost-effective. A telehandler with a spear attachment is used to pick and load bales in the rack. The bales are loaded into a rack either from the rear of the rack (rear-loading rack option) or from the side of the rack (side-loading rack option). We will consider the "side-loading rack option" for this study. The telehandler picks two bales at a time and loads them into a rack while it is still attached to the trailer. The trailer is a unique low-profile design developed specifically to haul racks. Since racks and trailers are handled as one unit in the"Side-loading rack option", it eliminates the need for a forklift at an SSL to off-load an empty rack and load a full rack. The telehandler will cycle back and forth to load a rack with 20 bales, 10 bales from the right side and 10 bales from the left side. The trailers with loaded racks are later towed to the bio-refinery using truck tractors. The design calls for two trailers behind a truck tractor, thus a truck load is 40 bales.

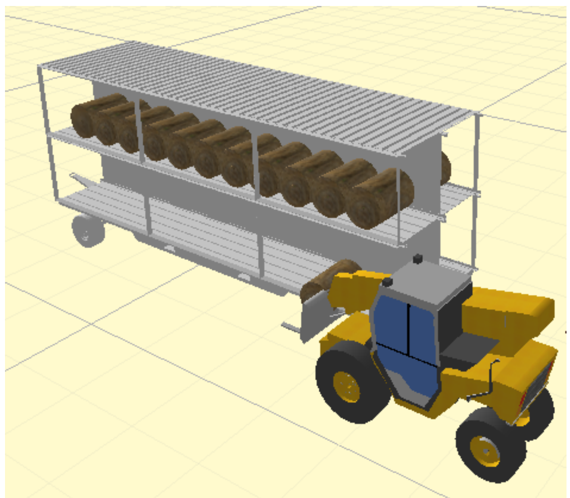


Figure 2.7: Side-loading rack option.

2.3.2 Highway hauling operation

The highway hauling operation includes hauling of biomass from a satellite storage location to a bio-refinery storage yard. In-field hauling equipment is generally not preferred for long distance hauling due to its limited road speed. Hauling contractors use truck tractors to haul biomass in racks attached to the trailers. The loading of tractor-trailers in the SSL is a challenging task to maintain cost-effectiveness of the biomass feedstock logistics chain as these operations have poor labor productivity (Mg handled per hour per worker). To improve labor productivity, multibale handling is proposed using a rack system (Cundiff [29]). The proposed rack is approximately the size of a standard shipping container with capacity of 20 bales, 10 on each side. Each truck tractor pulls two trailers each carrying one rack with 20 bales. A full truck load has 40 bales (16.3 Mg) of biomass and each rack has a bar code that is read when the truck crosses the scale at the bio-refinery. These data identify the source of biomass to track quality and for bookkeeping purposes to pay the farmgate contract holder.

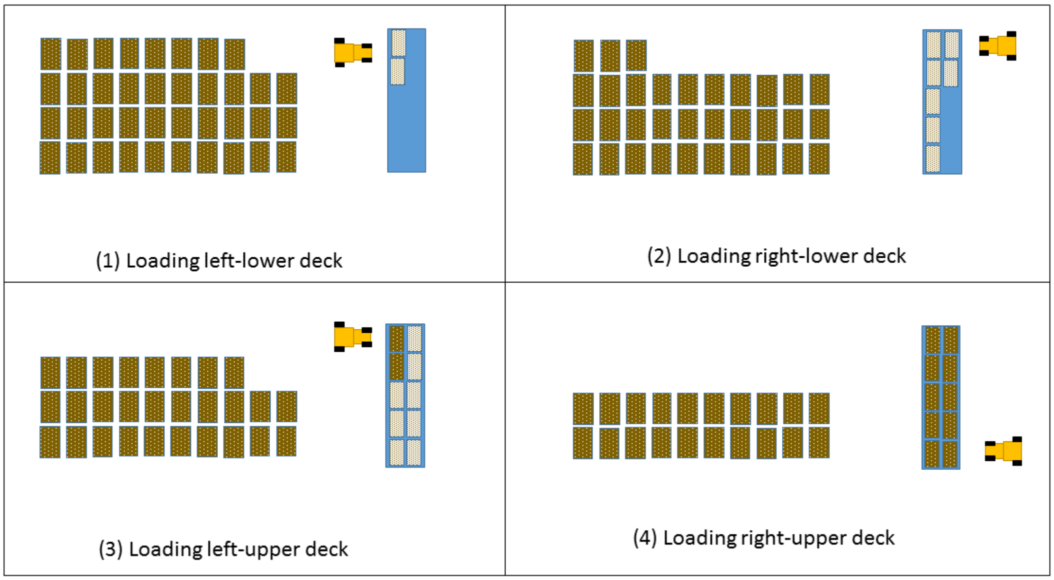


Figure 2.8: Process of loading bales on the rack.

Two significant factors determining highway hauling costs are:

1. Weight of each truck-load
2. Time required to load, haul, unload, and return to an SSL for next load

Some decisions that the hauling contractors make include number of load-out equipment sets, number of racks and trailers required, number of truck tractors required, and some operational level decisions like which SSL is to be unloaded at what time to meet the bio-refinery demand for year-round operation.

2.3.3 Receiving-facility operation

The receiving-facility operations essentially include all operations performed at the bio-refinery storage yard including weighing, unloading biomass and finally feeding bales into the bio-refinery as needed. The bio-refinery would ideally want to operate at “just-in-time” (JIT) inventory levels. The dynamics and variations of the real-world scenario make it challenging to achieve JIT delivery of raw materials. Hence, a minimum inventory is stored at the bio-refinery. The inventory at bio-refinery is stored in racks until processing to avoid individual bale handling, as this is a key part of the multibale handling concept. Reduction in cost of receiving-facility operation at the bio-refinery is a key advantage of the rack system concept.

The truck tractor brings two loaded racks on trailers and they first pass over a scale where the biomass load is weighed. Each load identification label is scanned to identify the source for bookkeeping purposes. The truck then moves to the unloading area where the rack is unlocked from the trailers. A fork lift (10-ton capacity) unloads full racks from trucks and moves them to the bio-refinery storage yard and loads the trailer with two empty racks. The receiving facility may have one or two forklifts for handling the racks depending on the biomass demand at the bio-refinery. The truck with two empty racks then passes across the scale to record weight and then returns to an SSL for picking up the next load. A full rack, either directly from the truck or from the bio-refinery storage yard, is loaded into a rack unloader. The rack unloader pushes bales out of the rack onto a conveyor which delivers a single line of bales into the bio-refinery for size reduction and subsequent processing.

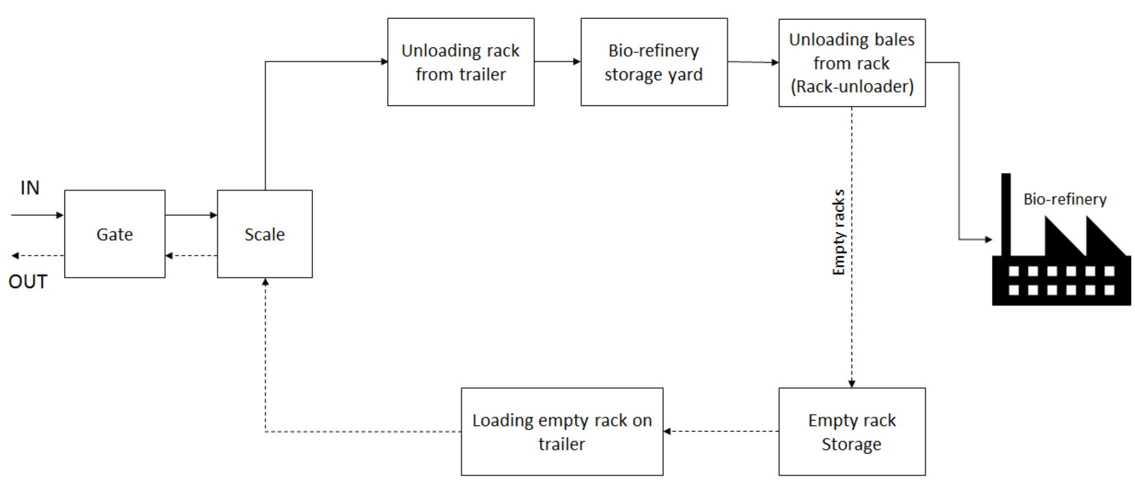


Figure 2.9: Recieving-facility operations.

2.4 Biomass feedstock logistics chain for the Midwest United States

Biomass feedstock produced in the Midwestern and Western United States differs from the feedstock produced in the Southeastern United States due to soil composition, climate, rainfall pattern and crops grown. Corn stover is the major cellulosic biomass feedstock used in the Midwest region for production of ethanol. Corn stover is the residue after corn grain is harvested. It is inherently contaminated with dirt and foreign materials, making its harvesting and processing a complex process. A listing of the Midwest unit operation is given in Figure 2.10.

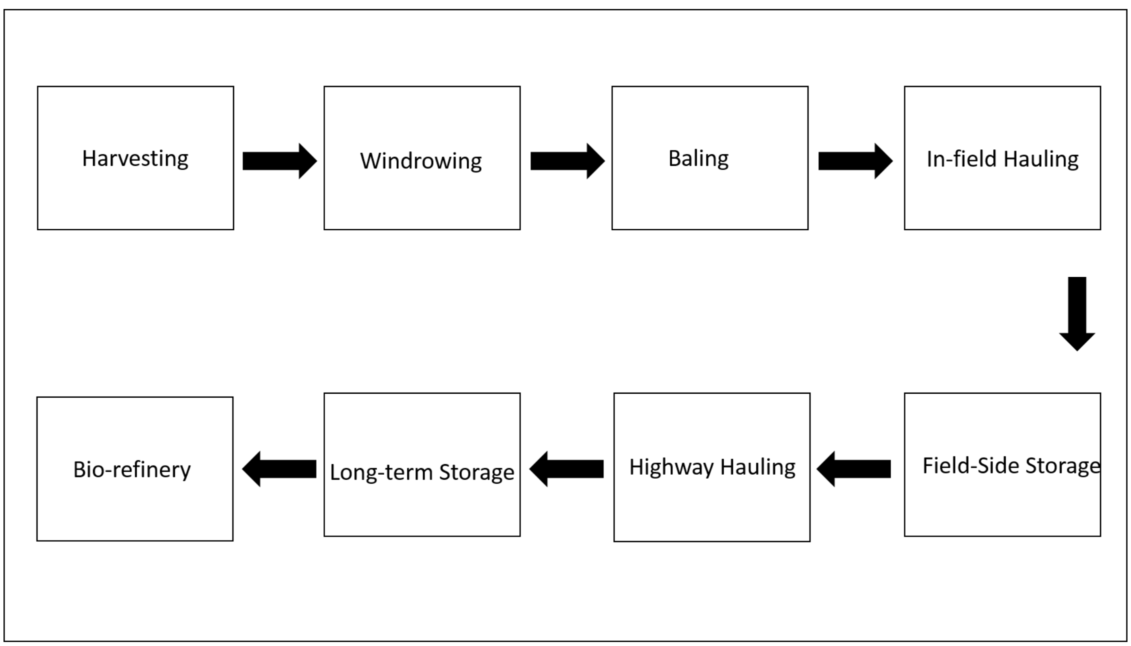


Figure 2.10: Midwest biomass feedstock logistics chain.

The average size of corn fields in the Midwestern United States is larger than the smaller switchgrass fields in the Southeast. Corn stover, unlike switchgrass, is harvested in a short time period, requiring more equipment and larger storage facilities to ensure a year-round supply of biomass to the bio-refinery. Two baling technologies, both rectangular and round bales are used in the Midwest. The large rectangular-bales are typically stacked six high and covered with a tarp. Some important decision-support problems in this system include determining which farms to contract with, fleet size of equipment, scheduling of harvest and in-field hauling, and inventory management, among others.

Chapter 3

A Taxonomic Review - Biofuel Supply Chain Problems

3.1 Introduction

A growing global energy requirement, and perceived effects of global warming caused by the release of fossilized carbon into atmosphere have encouraged researchers to explore sustainable alternatives to fossil fuels. Renewable sources of energy like hydro power, solar, wind, geothermal, and biomass are widely studied and considered for development. Biofuels produced from biomass is an attractive renewable energy product as it can substitute for motor fuels like gasoline and diesel. A study of biomass feedstock logistics, which constitutes a major portion of biofuel cost at the pump, is important to make it a cost-effective alternative to fossil fuels. Thus, the farmgate-to-bio-refinery supply chain has been a prominent research area over the past decade, supporting the development of a sustainable and cost effective bio-energy supply chain.

Most review papers on the biofuel supply chain generally classify publications based on supply chain level: upstream, midstream, and downstream (An et al. [8], Sharma et al.

[106]), decision level: strategic,tactical, operational, and integrated (An et al. [8], De Meyer et al. [35], Sharma et al. [106], Yue et al. [119]), optimization methodology: linear and non-linear programming, pure and mixed integer programming, deterministic and stochastic models (Baos et al. [16], De Meyer et al. [35], Shabani et al. [104], Sharma et al. [106]), or type of objective function: economic, social, and ecological based (De Meyer et al. [35]). Mafakheri and Nasiri [81] organize publications based on the biomass supply chain operations like harvesting, storage, transportation and the conversion process. Yue et al. [119] segregate publications based on the type of biomass feedstock, end product and conversion process. However, to the best of our knowledge, none of the review papers classify the literature based on the inherent optimization problems encountered in the farmgate-to-bio-refinery supply chain. We believe that this is essential to develop efficient solution approaches that exploit the structure of underlying problems.

In this chapter, we present a taxonomic review of selected publications in the field of biofuel supply chain based on the inherent optimization problems encountered like lot sizing, traveling salesman, and location-allocation problems, among others. This classification allows readers to identify these underlying optimization problems, and hence, take advantage of the state-of-the-art modeling and solution approaches from the extensive research on these well-studied optimization problems. It will also serve as a motivation for developing solution methodologies for new problems originating within the biofuel supply chain. We also classify the publications based on relevant modeling methodologies and solution approaches to present a comprehensive review of work in this area.

3.2 Classification framework

Thirty three scientific publications were selected and reviewed with a focus on the biofuel supply chain problem, and the methodologies used to model it. A framework of classification is presented in Figure 3.1. First, the publications are classified based on the inherent optimization problems encountered in the biofuel supply chain. The group of problems identified include single and multiple item lot sizing, location-allocation problem, traveling salesman problem and its extensions like multiple traveling salesman problem (m-TSP) and high multiplicity multiple traveling salesman problem (HMTSP). Second, the publications are classified based on the modeling methodologies like mathematical programming, simulation, and simulation-optimization. The publications under mathematical programming modeling approach are further classified into linear and non-linear programming, pure and mixed integer programming, and deterministic and stochastic programming. The publications in the mathematical programming domain are also sub-classified based on the solution approaches to evaluate popularity in use of customized algorithms to solve biofuel supply chain problems.

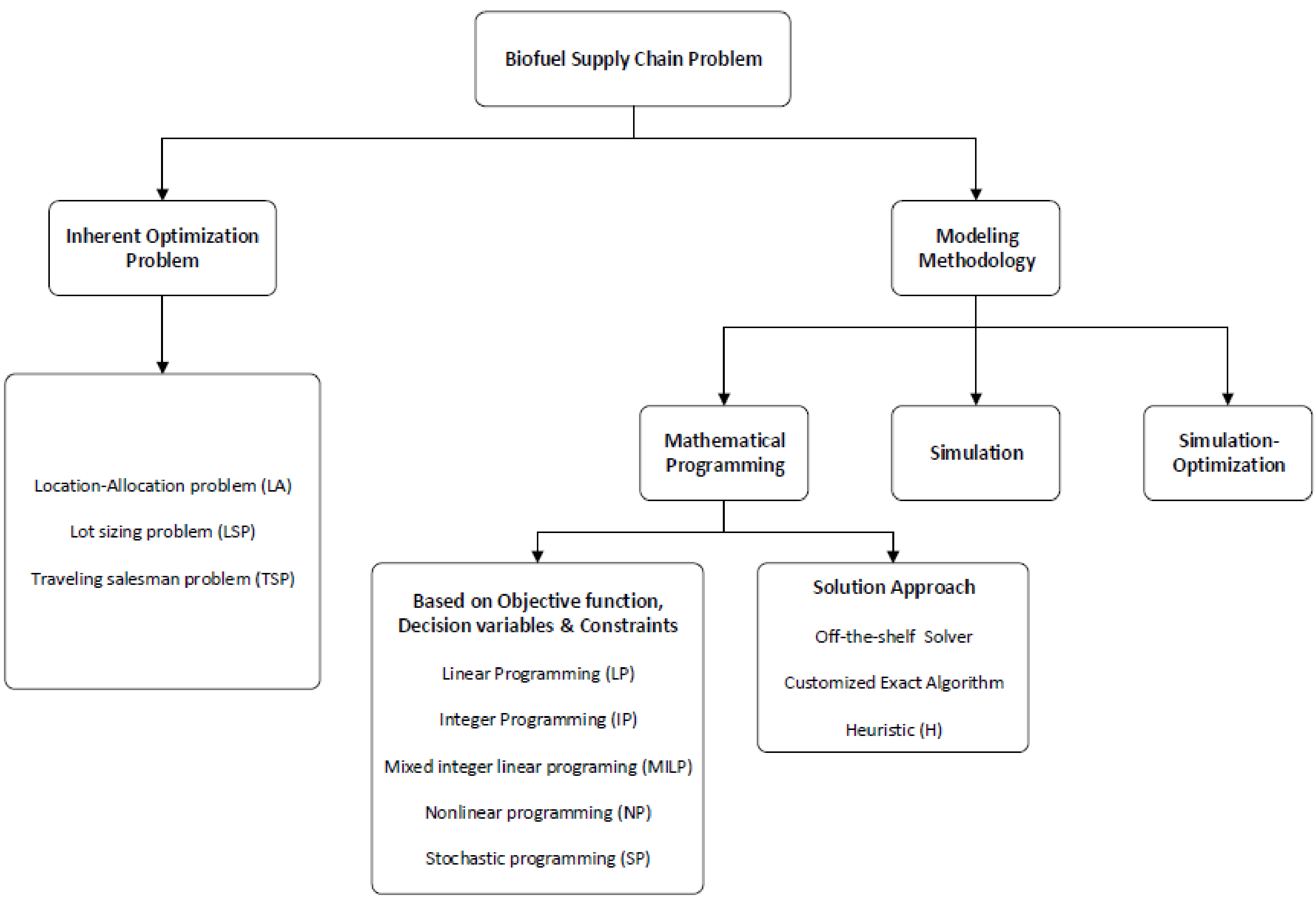


Figure 3.1: Classification framework for the taxonomic review.

3.2.1 Inherent optimization problems in the biofuel supply chain design

3.2.1.1 Location-allocation

The location-allocation problem (LA), in general, is to locate an optimal number of facilities to satisfy demand for a set of customers such that the overall cost of setting up facilities and transporting goods between customers and facilities is minimized [12]. Some exact and approximate algorithms proposed for this problem in the literature include branch-and-bound-based algorithm (Kuenne and Soland [68]), simulated annealing (Murray and Church [87]) and tabu search (Ohlemller [90]). Gong et al. [50] have proposed a hybrid algorithm involving a Lagrange relaxation method and a genetic algorithm. Brimberg et al. [21] presented heuristic approach using variable neighborhood search for the problems involving a large number of facilities.

The location of storage sites (facilities) and the allocation of production fields (customers) to those storage sites in the farmgate-to-bio-refinery supply chain is similar to a special case of the location-allocation problem, namely, single source capacitated facility location problem (SS-CFLP). The objective of SS-CFLP is to locate optimal number of facilities (biomass storage sites and bio-refineries) that serve a set of costumers (production fields and blending stations) at a minimum cost, where each costumer is served by a unique capacitated facility. Mathematical programming-based methodologies proposed for the SS-CFLP include Bender's decomposition (Geoffrion and Graves [49]), Lagrangian relaxation (Barceló and Casanovas [18], Klincewicz and Luss [67], Beasley [19], Sridharan [110], and Holmberg et al. [59]); and column generation (Díaz and Fernández [38]). Heuristic-based approaches to solve large-scale instances of SS-CFLP include kernel search heuristic (Guastaroba and Speranza [54]), hybrid algorithm combining Lagrangian heuristic and ant colony system (Chen and Ting [24]) and a multi-exchange heuristic (Ahuja et al. [2]).

3.2.1.2 Lot-sizing problem

The lot-sizing decisions identify when and how much to produce without violating any capacity constraint in order to minimize the cost associated with setup, production and inventory holding [64]. The lot-sizing problem can be classified as capacitated or un-capacitated based on the presence of capacity constraints in the problem. The capacitated lot-sizing problem (CLSP) involves discrete periods, finite horizon, resource with a finite capacity, and multiple products with known demands. The items or products in the problem correspond to different types of biomass (e.g. switchgrass and corn stover), and lot sizing corresponds to the determination of the amount of biomass to be produced in fields or delivered from storage locations to bio-energy plants during each time period. The capacity constraints include the availability of biomass at production fields or storage sites, and the capacities of the load-out equipment sets or tractor-trailers. Some relevant mathematical programming-based methodologies proposed for the CLSP include reformulating the problem using valid inequalities (Barany et al. [17], Leung et al. [76]), use of variable re-definition to provide tighter linear programming relaxation (Eppen and Martin [45]), branch-and-bound-based heuristic method(Hindi [58], Armentano et al. [10]), Lagrangian relaxation and its heuristic implementation (Thizy and Van Wassenhove [114], Diaby et al. [37]), and set partitioning and column generation approach (Manne [83], Cattrysse et al. [23], Chen and Thizy [25], Degraeve and Jans [36]). Karimi et al. [64] gives a comprehensive review of modeling methodologies for the capacitated lot-sizing problems.

Apart from classification of the lot-sizing problem as capacitated and un-capacitated, it can be also classified as single item or multi-item lot-sizing based on the number of products involved and as single facility or parallel facility based on the number of facilities involved. In the context of the farmgate-to-bio-refinery supply chain, parallel facility lot sizing problems are common due to the presence of multiple fields, storage sites or bio-refineries. Lot-sizing problems with parallel facility have been studied by Zangwill [120], Sung [111], and Sambasivan and Yahya [99].

3.2.1.3 The traveling salesman problem and its extensions

The routing of harvesting equipment or load-out equipment to unload storage sites can be modeled as a traveling salesman problem (TSP), where the salesmen and cities are represented by the equipment sets and production fields, respectively. The traveling salesman problem (TSP) can be defined as follows: Given a complete graph with vertex set N in which city 1 denotes the base city, distance matrix $[c_{ij}]$, $i, j \in N$, determine an optimal tour that starts and ends at a base city after having visited city i exactly once, $\forall i \in N$, while minimizing the total distance traveled. Let x_{ij} be a binary variable, which equals 1 if city i directly precedes city j in the TSP solution, and zero otherwise, $\forall i, j \in N,\ i \neq j$. If $c_{ij} = c_{ji}$, the problem is denoted as the symmetric traveling salesmen problem (STSP) and if distance matrix $[c_{ij}]$, $i, j \in N$ is asymmetric, then the problem is denoted as the asymmetric traveling salesmen problem (ATSP). Comparison of polynomial-length formulations for the ATSP has been presented in Öncan et al. [91], Roberti and Toth [97], and review of exact methods for the ATSP have been presented in Laporte [71], Fischetti et al. [47], D'Ambrosio et al. [32] and Roberti and Toth [97].

Another extension of ATSP is mATSP, where m salesmen (instead of one) are located at the base city. Review of polynomial-length formulation for the mATSP have been presented by Kara and Bektas [63] and Sarin et al. [102], where the latter present a comprehensive review of 32 formulations.

The high-multiplicity asymmetric traveling salesman problem (HMATSP) is another extension of the traveling salesman problem, where a salesman departs and then returns to the base city after having visited each city multiple times. It can be concisely defined as follows: Given a complete graph with vertex set N, an asymmetric distance matrix $[c_{ij}]$, and positive integers n_i, $i, j \in N$, determine an optimal tour that starts and ends at a base city after having visited city i exactly n_i times, $\forall i \in N$, while minimizing the total distance traveled. Grigoriev and van de Klundert [52] and Sarin et al. [101] propose formulations for the HMATSP of exponential and polynomial lengths, respectively.

3.2.2 Modeling methodology : Mathematical programming

Mathematical programming is an analytical decision tool in which a real-world system is represented in mathematical terms. It supports the user to make better decisions or to understand system characteristics better [117]. A mathematical program in general involves an objective function (goal) which we seek to minimize or maximize, decision variables that we seek to determine, and constraints that portray conditions under which the decisions have to be made. Next, we present general forms of mathematical programs.

3.2.2.1 Classification of mathematical programming

Based on the type of objective function, decision variables and constraints, mathematical programs can be classified as follows:

Linear programming (LP) : A mathematical programming model for which the objective function and all its constraints are linear [117].

Nonlinear programming (NLP) : A mathematical programming model for which some of the constraints or the objective function are non-linear.

Pure and mixed integer linear programming : A mixed integer linear program (MILP) consisting of both integer and continuous decision variables, whereas a pure integer program (IP) has all integer decision variables [27]. The general representation is as follows :

Mixed integer program:

$$\text{Minimize} \;\; c^T x + d^T y$$

$$\text{subject to} \;\; Ax + Gy \leq b$$

$$x \in \mathbb{Z}_+^p$$

$$y \in \mathbb{R}_+^n$$

Pure integer program:

$$\begin{aligned}&\text{Minimize } \ c^T x\\ &\text{subject to } \ A(x) \le b\\ &x \in \mathbb{Z}^n_+\end{aligned}$$

where A is (m x n) matrix, G is (m x p) matrix, b is a vector in $\mathbb{R}^n$

Deterministic and Stochastic programming : Mathematical programming models can also be classified as deterministic or stochastic programming models. A deterministic model does not include any uncertainty, and the output of the model is fully determined by the parameter values and initial conditions. Deterministic linear and mixed integer linear programming models are the most extensively used modeling methodologies. Stochastic programming models consider uncertainty, and hence, the same set of parameters and initial conditions will yield different outputs based on the uncertainty considered in the model. A deterministic equivalent of stochastic programming model is built by considering different possible scenarios with their probability and optimizing the expected value of the objective function. Expected value and risk-based objective functions are commonly used in stochastic programming models. Modeling of variability is essential for the problems encountered in the farmgate-to-bio-refinery supply chain, as there are a number of uncertainties involved in the production, inventory, and distribution of biomass feedstock, which directly affect the decision making process.

3.2.2.2 Solution approaches of mathematical programming

Off-the-shelf solver : Commercial solvers like Cplex, Xpress, Gurobi, and so on, are directly used to solve linear and mixed integer linear programming models. The system to be modeled is studied, formulated and coded in softwares like OPL, C++, Matlab etc. The solver engines are then used to solve the optimization model. This solution approach is relatively user friendly and extensively used. Many real-life problems often modeled as integer or mixed integer programs are not efficiently solved to optimality when directly solved using these commercial solvers.

Customized exact algorithm : A modeling and solution approach that is relatively efficient, like branch-and-price, branch-and-cut, and relaxation and decomposition techniques, among others, is developed specific to a problem by exploiting its structure. Identifying the inherent optimization problem contained in the farmgate-to-bio-refinery supply chain is imperative to develop a customized algorithm for a given problem.

Heuristics : Integer and mixed integer programming optimization problems belong to the class of NP-hard problems. They are difficult to solve and require excessive computational efforts to find optimal, or sometimes even a feasible solution, for a large-sized problem. Heuristic approaches like simulated annealing, tabu search, and genetic algorithms search for a good feasible solution, but not necessarily global optimal solutions, to reduce computational effort. Heuristic methods could be used to find a suitable initial solution or to tighten the feasible domain of solutions and to direct the search for an optimal solution.

3.2.3 Modeling methodology: Simulation

Simulation is another modeling methodology widely used to design and analyze supply chain systems apart from mathematical programming. It is a computerized imitation of the random behavior of a system for the purpose of estimating its measures of performance. The choice of an optimization method, or simulation, is based on the system being modeled and the business needs. Optimization is preferred when the objective is to determine the best possible decision while simulation is preferred either when the system is too complex to be modeled by mathematical programming models, or when the degree of uncertainty in the model is large enough that ignoring it would lead to biased results. Simulation models are extensively used to design and analyze a supply chain because of its ability to capture the inherent variability and dynamics of the system, thereby yielding a better decision making capability.

Simulation models can be broadly classified as continuous and discrete [112]. Continuous models are used when the behavior of the system changes continuously with time. These models usually use difference-differential equations to describe interactions among different elements of the system. Discrete models, or discrete event simulations, are generally used to study the measures related to queuing in the system that change only when an entity enters or leaves the system. Verification and validation of simulation is an important step while modeling a complex supply chain. Verification is the process of verifying the process logic implemented in the simulation model and confirming that the implementation of the model is correct. Validation evaluates how accurately the model represents the real system being modeled by using a known dataset as input and evaluating output from the simulation model.

3.2.4 Modeling methodology: Simulation-optimization

Optimization and simulation are both widely used to model supply chain problems. The performance of supply chain systems are susceptible to uncertainties and dynamics in the real world, making stochastic optimization models a more realistic approach. These optimization models in general are computationally difficult to solve. Simulation models on the other hand do not determine the best solution, but they measure system performance against uncertainties. Traditionally, optimization has been used to model supply chain problems with relevant assumptions, followed by the use of simulation models to evaluate the performance of the solution obtained in the face of uncertainties. Another modeling approach which optimizes the objective function via a simulation model is called simulation-optimization. The simulation-optimization approaches are comprehensively reviewed by Amaran et al. [7], Carson and Maria [22], Tekin and Sabuncuoglu [113].

Almeder et al. [6] propose a novel methodology for supply chain design problems using a combination of optimization and discrete event simulation. The simulation model includes stochastic elements, whereas the optimization model represents a simplified version of the problem without considering the uncertainties. The solution of the optimization model provides a set of decision rules for the discrete event simulation model. The optimization and simulation modules are run iteratively until the termination criteria is reached. This approach enables modeling of real-life problem instances by effectively incorporating dynamism and uncertainties. Bilgen and Çelebi [20] and de Keizer et al. [33] use this modeling methodology for a production scheduling and distribution planning problem in a dairy supply chain and for designing a logistics network for distributing perishable products, respectively. Marques et al. [84] develops a decision support tool for short-term planning of forest operations by combining optimization and simulation methodologies. The approach is used to study a raw material reception problem in a Portuguese pulp mill where the performance of deterministic schedules for the raw material trucks developed by optimization module is analyzed considering the uncertainty through the simulation module.

3.3 Results and Discussion

3.3.1 Inherent optimization problems in the biofuel supply chain design

The strategic and tactical-level farmgate-to-bio-refinery supply chain models, which seek optimal location of a bio-refinery, or storage facilities, have location-allocation problems embedded in them. A multi-commodity network flow is proposed in Zhu and Yao [122] to model the logistics system for multiple biomass feedstock. The problem is formulated as a mixed integer linear program to determine locations of facilities (like warehouses and bio-refinery), the size of the harvesting team, types and amount of biomass to be harvested, stored and processed for every time period in order to maximize overall profit. The model presented contains two location-allocation problems, one of which being a multi-period location-allocation problem. The warehouse is optimally located for every time period from a list of potential locations, while the bio-refinery is optimally located for the entire planning horizon using

Table 3.1: Publications classified based on an inherent optimization problem sub-structure.

Publications	**Capacitated lot-sizing with parallel facility (Single Item)**	**Capacitated lot-sizing with parallel facility (Multiple Item)**	**LA**$^{+}$	**TSP***
Lin et al. [78]	X			
Dyken et al. [40]		X		
Zhu and Yao [122]		X	X	
Xie et al. [118]	X	X		
Huang et al. [60]	X		X	
Sokhansanj et al. [109]			X	
Sharma et al. [107]	X			
Akgul et al. [3]			X	
Gunnarsson et al. [55]	X			
Judd et al. [62]			X	X
Zhang et al. [121]	X		X	
Ekiolu et al. [44]	X	X	X	
Ravula et al. [94]				X
Aguayo et al. [1]	X			X
Ekiolu et al. [43]	X		X	
Cundiff et al. [30]	X			
Kim et al. [66]			X	
Alex Marvin et al. [4]			X	
Lin et al. [77]			X	
Kim et al. [65]			X	
Ebadian et al. [42]		X	X	
Dal-Mas et al. [31]	X		X	
Mol et al. [86]			X	

LA^{+} Location allocation, TSP* Traveling salesman problem

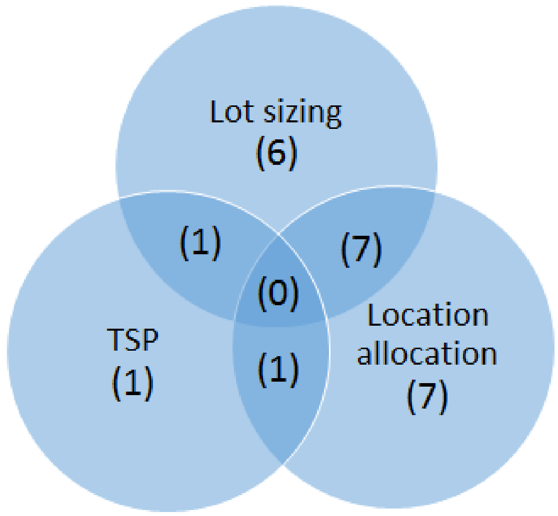


Figure 3.2: Distribution of inherent optimization problem in the publications reviewed.

binary variables in the optimization model. Huang et al. [60] also present a multi-stage mixed integer optimization model having a location-allocation problem within it. The objective of the model is to minimize total system cost over the entire planning horizon. The location of bio-refinery is optimally selected from the list of potential locations for every time period in the planning horizon. If a bio-refinery is opened, it cannot be closed or relocated; only additional bio-refineries can be opened during subsequent time periods. The optimal locations of bio-refineries are determined in Alex Marvin et al. [4] and Ekiolu et al. [43] to maximize net present value and minimize overall cost, respectively, for the entire supply chain. Judd et al. [62] address a special case of location-allocation problem called single source capacitated facility location problem (SS-CFLP), which involves determining optimal number and locations of storage facilities that serve a set of production fields at a minimum cost, where each production field is allocated to only one storage facility. Table 3.1 lists all the publications addressing the location-allocation feature.

The lot sizing problem (LSP) is another common optimization problem ingrained in farmgate-to-bio-refinery supply chain models. Lin et al. [78] introduce a mixed integer programming model to optimize operations ranging from harvesting, in-field transportation, storage, preprocessing, highway hauling to ethanol production and distribution. The key decisions include number, location and capacities of facilities, biomass and bio-ethanol distribution patterns. The model contains three capacitated single item lot sizing problems with parallel facilities (multiple farms)at production field, storage facilities for biomass feedstock and at the biorefinery. Xie et al. [118] present a multi-stage mixed integer programming model with two capacitated lot sizing problem with parallel facilities. The first lot sizing problem determines the amount and type of biomass feedstock transported from each production field to the storage facilities (capacitated multiple-item lot sizing problem with parallel facilities), while the second lot sizing problem determines the amount of biofuel allocated from the bio-refinery to demand points for each time period (capacitated single-item lot sizing problem with parallel facilities). Another mixed integer programming model developed by Ekiolu et al. [44] to design and manage the biomass supply chain determines the amount of biomass to be shipped between different facilities, in addition to the optimal location of facilities.The model presents two capacitated lot sizing problems with multiple and single items associated with the biomass feedstock and bio-ethanol lot sizes, respectively. Table 3.1 lists all the publications containing the lot sizing feature.

Tactical and operational level supply chain models, which consider routing of equipment or trucks, will have the traveling salesman problem (TSP) or its extensions associated with it. Judd et al. [62] propose a pure integer programming model to determine optimal number and location of storage facilities. The authors consider sharing of load-out equipment among storage facilities in order to improve resource utilization. The routing of these load-out equipment sets amount to the mATSP optimization problem. The model is effectively solved by developing a decomposition technique-based solution approach. The routing of load-out equipment sets in Aguayo et al. [1] can be viewed as a time-dependent, selective, high-multiplicity, multiple traveling salesman problem (TDS-HMmTSP). The proposed model seeks to determine optimal number of load-out equipment sets and trucks needed for the planning horizon by minimizing the total cost incurred by mobilization of equipment and hauling of biomass using trucks. A branch-and-price-based approach is developed to effectively solve the large-size instances of the problem to obtain near-optimal solutions. Ravula et al. [94] uses a TSP optimization routine within the simulation for routing load-out equipment among storage facilities.

3.3.2 Modeling methodology

3.3.2.1 Mathematical programming

Types of mathematical programming models

Table 3.2 identifies publications that employ a linear programming-based approach. Alfonso et al. [5] present a methodology to optimize locations of facilities from a logistics point of view, and provide necessary data to analyze different biomass energy options from technical, economic, and environmental viewpoints. The central module for optimization and computation interacts with five other modules, namely, biomass resource module, demand module, logistic module, technology characterization module, and environmental module. The optimization methodology is employed in two steps: the first step provides a list of optimal locations based on minimum transportation costs and distribution of biomass resources, while the second step optimizes economic suitability and greenhouse gas emissions with the help of the results provided by various supporting modules. Frombo et al. [48] use a linear programming-based approach for a wood-based forest biomass supply chain problem. The objective is to minimize different costs associated with felling and processing woody biomass, collection and transportation, plant installation and maintenance. The model also considers the benefits resulting from energy sales at the plant. The LP model expresses the costs as a function of the annual harvested biomass quantity and plant capacity, constrained by the forest biomass collection and the continuity equation at the energy plant. Another linear programming based model with objective of developing the most sustainable bio-ethanol supply chain is presented in Ren et al. [95] by minimizing the total ecological footprint, along with satisfying the stakeholder's requirements.

Table 3.2: Publications classified based on the mathematical programming approach used.

Publications	**LP**	**IP**	**MILP**	**NLP**	**Deterministic**	**Stochastic**
Lin et al. [78]					X	
Dyken et al. [40]			X		X	
Zhu and Yao [122]			X		X	
Velazquez-Marti and Fernandez-Gonzalez [115]					X	
Xie et al. [118]			X		X	
Huang et al. [60]			X		X	
Sharma et al. [107]			X			X
Akgul et al. [3]			X		X	
Gunnarsson et al. [55]			X		X	
Judd et al. [62]		X			X	
Zhang et al. [121]			X		X	
Ekiolu et al. [44]			X		X	
Aguayo et al. [1]					X	
Ekiolu et al. [43]			X		X	
Cundiff et al. [30]	X					X
Kim et al. [66]			X		X	
Alex Marvin et al. [4]			X		X	
Lin et al. [77]			X		X	
Kim et al. [65]			X			X
Dal-Mas et al. [31]			X			X
Ren et al. [95]	X				X	
Shabani and Sowlati [103]				X	X	
Mol et al. [86]			X		X	
Alfonso et al. [5]	X				X	
Frombo et al. [48]	X				X	

Integer programming models have all its decision variables as integer variables. In the publications reviewed, only Judd et al. [62] propose a pure integer programming model. The objective is to minimize costs related to storage and transportation for different equipment options at the storage site by optimizing location of storage facilities and the allocation of biomass from production fields to these storage facilities. Two integer (binary) decision variables define whether a farm is selected as a storage location (location) and whether a production field uses a given storage facility (allocation).

Mixed integer linear programming (MILP) models have both continuous and integer decision variables. Most of the approaches in farmgate-to-bio-refinery supply chain analysis use a mixed integer linear program-based model. We discuss few papers with MILP models and readers are directed to Table 3.2 for a complete list of all the publications belonging to this classification. Akgul et al. [3] develop a mixed integer linear programming-based model to design a bio-ethanol supply chain by optimizing total supply chain cost. The model optimizes the location and scale of the plant, flow of biomass and bio-ethanol between regions, and the number of transport units required for the transfer of biomass and bio-ethanol by using the neighborhood flow approach to avoid long computational time. Kim et al. [66] propose a MILP model that enables the selection of fuel conversion technologies, capacities, biomass locations, and the logistics design to maximize the overall profit. The decisions determined by the model include optimal number and location of the processing plants, the amount of biomass, intermediate products, and final products to transport among different locations. Zhang et al. [121] present a multi-period mixed integer linear programming model to optimize the total annualized switchgrass-based bio-ethanol supply chain costs by integrating all supply chain and logistics decisions. The integer decision variables are associated with determining the location of preprocessing and bio-refinery facilities, while the continuous decision variables are associated with amount of biomass feedstock and bio-ethanol transported.

A mathematical programming model is called non-linear if the objective or some of the constraints contain non-linear functions. A mixed integer non-linear programming (NLP) model is introduced in Shabani and Sowlati [103] for tactical value chain optimization for a woody biomass-based power plant. The model considers biomass procurement, storage, energy production and ash management. The non-linear constraints relating to energy production and some integer decision variables make it a mixed integer non-linear programming model. The model is decomposed into a non-linear programming sub-model and mixed integer linear master problem. An outer approximation algorithm [39] iteratively solves the sub-problem and the relaxed version of master problem with the assumption of convexity of the non-linear sub-model.

Most of the reviewed papers use deterministic models. Modeling uncertainties is imperative as decision making in farmgate-to bio-refinery supply chain is more susceptible to the dynamics and uncertainties encountered. Cundiff et al. [30] consider the uncertainty in biomass production levels based on weather while optimizing the design of storage facilities and logistics of biomass delivery system. The authors consider four weather scenarios by combining good and poor weather conditions for crop growth along with good and poor weather conditions during the harvest period. A multi-stage stochastic linear programming model is used to determine the shipment of biomass between facilities to optimize cost. The paper also evaluates storage expansion schedule for each producer by considering monthly harvest for each of the four weather scenarios and transportation of biomass from on-farm storage locations to a centrally located plant. Sharma et al. [107] develop a scenario-based optimization model to address weather-based uncertainties to minimize cost associated with harvest, transportation, and storage. The weather scenarios used in the optimization model are built from daily weather data collected from Oklahoma Mesonet. The model determines the required number of harvesting and transporting equipment sets. In Dal-Mas et al. [31], a multi-echelon mixed integer linear programming modeling framework is presented, which serves as a strategic decision support tool to assess economic performance and risk on investment considering uncertainty in the price. The two different objective functions considered are the expected net present value and conditional value-at-risk (CVaR). The proposed model optimizes the economic performances and minimizes the financial risk on investment by determining the best biomass cultivation site locations, ethanol production plant capacity and location, and transport logistics. A mixed integer linear program that enables selection of fuel conversion technologies, capacities, biomass location and logistics of biomass and end products under uncertainty, is considered in Kim et al. [65]. This strategic model considered maximizes the expected value of overall profit.

Types of solution approaches

A distribution of the publications based on the different solution approaches used is given in Figure 3.3.

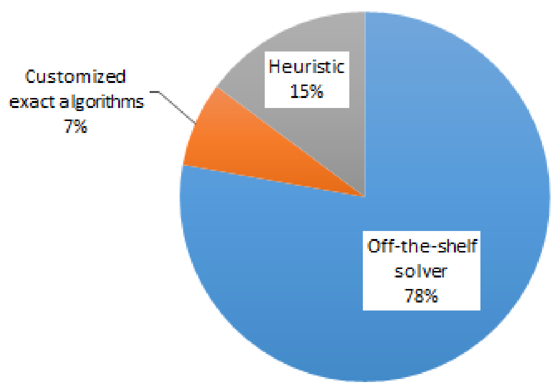


Figure 3.3: Distribution of publications based on different solution approaches used.

It is evident from Figure 3.3 that using an off-the-shelf solver directly is the most common solution approach used for solving optimization models in the farmgate-to-bio-refinery supply chain. The mixed integer linear programming models that are employed are generally large-scale MILP models and are not easily solvable to optimality. Often, the problem resolution is adjusted to reduce the number of integer variables associated with a problem, and sometimes customized exact or approximate algorithms are developed to solve these large-scale models.

Aguayo et al. [1] developed an algorithm for the biomass logistics supply chain design problem (BL-SCDP) which models operations performed at storage facilities and transportation of biomass from these storage locations to a bio-refinery. The BL-SCDP is modeled as a

Table 3.3: Publications classified based on the solution approaches used.

Publications	**Off-the-shelf solver**	**Customized exact algorithms**	**Heuristic**
Lin et al. [78]	X		
Dyken et al. [40]	X		
Zhu and Yao [122]	X		
Velazquez-Marti and Fernandez-Gonzalez [115]			X
Xie et al. [118]	X		
Huang et al. [60]	X		
Sokhansanj et al. [109]	X		
Sharma et al. [107]	X		
Akgul et al. [3]	X		
Gunnarsson et al. [55]			X
Judd et al. [62]	X		X
Zhang et al. [121]	X		
Ekiolu et al. [44]	X		
Aguayo et al. [1]		X	
Ekiolu et al. [43]	X		
Cundiff et al. [30]	X		
Kim et al. [66]	X		
Alex Marvin et al. [4]			
Lin et al. [77]	X		
Kim et al. [65]	X		
Ebadian et al. [42]	X		
Dal-Mas et al. [31]	X		
Ren et al. [95]	X		
Shabani and Sowlati [103]		X	
Mol et al. [86]			X
Alfonso et al. [5]	X		
Frombo et al. [48]	X		

multi-stage mixed integer linear program, and it comprises single-item capacitated lot sizing problem with parallel facilities and time-dependent selective high-multiplicity multiple traveling salesmen problem. The model determines fleet size of trucks and load-out equipment sets, amount of biomass to be shipped from storage locations to biorefinery, and storage capacity. It is pointed out by the authors that direct application of best in class commercial solvers could not produce acceptable (near optimal) solutions, and hence, they developed a customized algorithm (based on branch-and-price approach). Computational investigation reveals efficacy of the proposed approach over direct application of a commercial solver by showing its ability to produce near-optimal solutions for large-sized problem instances.

Heuristic approaches are also used to obtain satisfactory but not necessarily optimal solutions for complex problems with lesser computational effort. Judd et al. [62] presents the biomass logistics problem involving location-allocation of storage sites with feature of sharing load-out equipments among storage facility (BLP-M). The problem features the mATSP as its substructure, and direct application of a solver is unable to solve large instances of the problem. The problem is decomposed by first solving for the locations of the storage facilities, and then, solving for equipment routes between storage sites. The solution methodology does not guarantee attainment of optimal solutions; however, it finds near-optimal solutions within 1% of optimality gap. Gunnarsson et al. [55] proposes a heuristic solution approach for a forest fuel-based supply chain for a heating plant to obtain solutions within a reasonable computational time. The supply chain system is modeled as multi-period, large-scale mixed integer linear programming model. The proposed heuristic approach can be interpreted as a depth-first search in a restricted branch-and-bound scheme. Mol et al. [86] formulates a multi-feedstock biomass supply chain problem by using a mixed integer linear programming model. The overall problem is too complex to be directly solved using a commercial software, and hence, the original problem is decomposed into three sub-problems based on the biomass type. The sub-problems are individually solved, and then, a knapsack model is used to integrate them back together. The approach gives good solutions requiring reasonable computational effort.

3.3.2.2 Simulation models

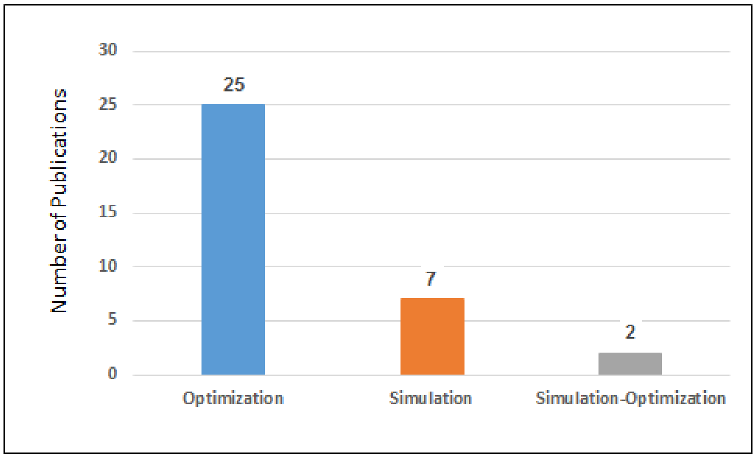


Figure 3.4: Distribution of publications applying different modeling approaches.

Simulation models are extensively used in farmgate-to-bio-refinery supply chain due to the presence of inherent uncertainties. Sokhansanj et al. [108] present a framework of a dynamic "integrated biomass supply analysis and logistics" (IBSAL) model to simulate the operations of collection, storage and transportation of biomass feedstock to a bio-refinery. The model considers uncertainties like weather, moisture content and dry matter loss of feedstock. Kumar and Sokhansanj [69] study cost, energy input, and carbon emissions for delivering switchgrass to a bio-refinery for three transport options (i.e. by baling, grinding, and chopping of biomass) by using the IBSAL model developed in [108]. The IBSAL-MC model (Ebadian et al. [41]) is based on the framework of IBSAL described in [108] with a hybrid push-pull logistics system. The field operations to harvest and collect biomass within the harvest window works under push logistics system, whereas the downstream operations required to meet daily demand of ethanol from bio-refinery, operate under a pull system. Some other features added to the basic IBSAL model includes ability to model multiple biomass feedstock, equipment breakdowns and realistic storage capacity against the unlimited storage capacity assumed in the basic IBSAL model.

Table 3.4: Publications classified based on modeling approaches.

Publications	Optimization	Simulation	Simulation-Optimization
Lin et al. [78]	X		
Dyken et al. [40]	X		
Zhu and Yao [122]	X		
Velazquez-Marti and Fernandez-Gonzalez [115]	X		
Xie et al. [118]	X		
Huang et al. [60]	X		
Sokhansanj et al. [109]			X
Sharma et al. [107]	X		
Akgul et al. [3]	X		
Kumar and Sokhansanj [69]		X	
Gunnarsson et al. [55]	X		
Judd et al. [62]	X		
Zhang et al. [121]	X		
Ekiolu et al. [44]	X		
Ravula et al. [94]		X	
Aguayo et al. [1]	X		
Ekiolu et al. [43]	X		
Cundiff et al. [30]	X		
Sokhansanj et al. [108]		X	
Kim et al. [66]	X		
Alex Marvin et al. [4]	X		
Lin et al. [77]	X		
Kim et al. [65]	X		
Ebadian et al. [42]			X
Dal-Mas et al. [31]	X		
Ren et al. [95]	X		
Shabani and Sowlati [103]	X		
Mobini et al. [85]		X	
Windisch et al. [116]		X	
Ebadian et al. [41]		X	
Mol et al. [86]	X	X	
Alfonso et al. [5]	X		
Frombo et al. [48]	X		

Mobini et al. [85] develop a discrete event simulation approach based on the IBSAL framework to investigate forest biomass logistics for a power plant. The simulation evaluates the cost of transporting biomass, equilibrium moisture content, and the carbon footprint from logistics operation over the service life of the power plant. The model incorporates different harvesting systems like conventional harvesting, satellite yard harvesting and full-tree chipping. Windisch et al. [116] develop a discrete event simulation model to improve the efficiency of a wood-based supply chain by including uncertainties in the logistics system. Simulation models are also used to model specific parts of biomass logistics instead of modeling the entire farmgate-to-bio-refinery supply chain. For instance, a discrete event simulation model is used to evaluate truck scheduling strategies and to assess the cost required to load, haul and deliver a weekly supply of biomass from storage locations to a bioenergy plant in Ravula et al. [94].

3.3.2.3 Simulation-optimization models

The use of simulation and optimization together to model supply chain problems is gaining importance due to its capability of generating acceptable solutions for real-life problems with less computational effort. Mol et al. [86] evaluate a logistics system for biomass collection with both optimization and simulation model independently. The paper points out advantages of both modeling methodologies, motivating the reader to think about an integrated approach comprising both optimization and simulation modules.

Sokhansanj et al. [109] presents a simulation/optimization model for the cellulosic biomass supply chain to evaluate different biomass feedstock storage systems. The optimization model is a mixed integer linear programming model that prescribes the optimum number of farms to contract and their locations, optimal number of storage sites to open and their locations, and optimal assignment of farms to storage sites thereby minimizing harvesting, collection and hauling costs. The output from the optimization model serves as an input to the simulation model. Day-to-day variations like changing moisture content or dry matter loss are not modeled in the optimization module to avoid a complex and large-sized optimization model; they are modeled as a part of the simulation model. The simulation model which is an extension of IBSAL-MC [41] evaluates the total logistics costs, capacity of storages, number of equipments required, resource utilization, dry matter losses, and carbon foot print for different storage systems.

Ebadian et al. [42] propose to integrate the tactical and operational planning levels in the biomass supply chain using an optimization and a simulation model. The overall objective of the integrated model is to assure that the daily biomass demand of the bio-refinery is satisfied year-round at a minimum delivery cost. The optimization module produces a supply chain design considering a five-year planning horizon to minimize the logistics cost. Given the supply chain design, the simulation module simulates the biomass feedstock flow in the supply chain to meet demand at the bio-refinery. If the bio-refinery demand is not satisfied, outputs from the simulation model is used to update optimization model parameters. The optimization and simulation modules are iteratively run until the termination criteria like convergence to a desired supply chain design or a limit on computational time is reached.

3.4 Conclusion

Bio-energy is an attractive and emerging renewable source of energy with ability to substitute for the use of fossil fuel. Governments of different countries are focusing on ramping up biofuel production to meet major portion of energy demand. However, a variety of barriers and uncertainties inhibit the development of a strong, sustainable and cost-effective supply of bio-energy. Overcoming these hurdles is imperative for making biofuel a viable and attractive option.

The review of publications shows that the farmgate-to-bio-refinery supply chain system is mostly modeled with a mixed integer linear programming approach containing binary and continuous variables in order to determine location of facilities and flow of material. Modeling a real-life system would need a very large scale MILP model and huge computational effort for its solution. The farmgate-to-bio-refinery supply chain is often susceptible to large uncertainties, hence modeling stochasticity of the supply chain system is necessary. The number of publications reporting on stochastic programming models is comparatively low. Some researchers have used simulation models and simulation-optimization approaches successfully to model the stochasticity present in the system in order to determine solutions that are robust in the face of uncertainties. The common solution approach followed in most of the publications is to directly use a commercial solver, which would require limiting the problem resolution to avoid huge computational times. Customized algorithms comprising both exact and approximate approaches have been developed in a few publications. There exists a great potential to develop better solution approaches to these large-scale problems by understanding and exploiting the inherent optimization problem sub-structure.

Chapter 4

A Two-Phase Optimization-Simulation Approach for Tactical Decisions in a Biomass Feedstock Logistics Chain

4.1 Introduction

Biofuels constitute a potential solution for the challenges arising because of growing energy needs, depleting fossil fuel, and greenhouse gas (GHG) emissions. They are becoming increasingly popular as they can substitute for fossil fuel equivalents used in the transportation sector. Cellulosic ethanol, a biofuel manufactured from processing cellulosic biomass feedstock, is blended with gasoline to form flex-fuels like E10 (10% ethanol is blended with 90% gasoline). Biofuels are essentially made from plants, which consume carbon dioxide for photosynthesis, thereby reducing net greenhouse gas emissions. A key aspect in the production of biofuel is the biomass logistics chain that constitutes a complex collection of activities, which must be judiciously executed for a cost-effective operation. The cost associated with biomass feedstock logistics is said to be 35-60% of the total cost of the cellulosic ethanol paid at retail pump Fales et al. [46].

Decisions in the farmgate-to-bio-refinery supply chain can be classified into: (1) strategic level, (2) tactical level, and (3) operational level decisions. The strategic-level decisions refer to capital intensive decisions like technology selection, location determination and size of bio-energy plant, central storage facility, capacity enhancement of existing facilities, among others. These decisions are revised once every few years (around five). For example, Judd et al. [62] propose a strategic-level biomass logistic chain problem where locations of storage facilities and fleet size of load-out equipment set is determined. The tactical-level decisions are revised approximately in six months to one year and pertains to fleet sizes of load-out equipment and trucks, among others. The operational-level decisions are short-term decisions (hourly, daily, or weekly) like inventory control, scheduling and routing decisions.

In this chapter, we propose a two-phase optimization-simulation approach to help make tactical decisions like fleet sizes of load-out equipment and tractor-trailers in view of the uncertainties encountered. The approach minimizes the total cost considering both the capital and biomass logistics operating costs while ensuring the probability of meeting bio-refinery demand with required reliability. We designate this problem as the *Stochastic Biomass Logistics Chain Problem* (S-BLCP).

The chapter is organized as follows: In Section 4.2, we present a brief description of the biomass logistics chain under study. In Section 4.3, we define the problem and present a brief review of literature related to our problem. Section 4.4 presents the proposed modeling methodology consisting of optimization and simulation modules. The approach is applied to a real-life data set collected for the farms within a 48-km radius of Gretna, VA, and results are presented in Section 4.5. Finally, concluding remarks are made in Section 4.6.

4.2 System description

The biomass feedstock considered in this study is switchgrass, a promising bio-energy crop suitable for the Southeastern United States, that grows well in marginal soil and in climatic conditions that are not ideal for traditional crops. Switchgrass is cut, then dried, baled and transported to satellite storage locations (SSL) using in-field hauling equipment owned by the feedstock producer. The round bale is commonly used in the Southeastern United States due to its ability to shed water when stored in single-layer ambient conditions. The inflow of biomass to storage location depends on the harvesting scenario adopted by the feedstock producers, and these scenarios are studied in Grisso et al. [53].

The SSL is a temporary storage facility that is used to store biomass, either from a single large farm or from several smaller farms, before hauling it to the bio-refinery storage yard. These storage locations serve as a transfer point between in-field hauling and highway hauling. A telehandler is used in an SSL to load biomass feedstock bales into the trucks. Resop et al. [96] and Judd et al. [62] have studied strategic decisions like determining locations of these storage facilities for a given set of production fields. All the activities ranging from production of biomass, harvesting, baling, and delivery of biomass to a storage facility are collectively known as "farmgate operations". The in-field hauling equipment has a limited road speed making it less efficient for long hauling; hence, it is only used to haul biomass from fields to an SSL. A truck tractor coupled with a specially-designed low-deck trailer is used for the highway hauling operation. A rack system is used for highway hauling of biomass in order to improve efficiency of the hauling operation. The telehandler loads the bales into racks during a 10-hour workday. A rack, having dimensions that emulate a 20 feet ISO shipping container, can accommodate two levels of 10 bales, and these bales, once loaded in a rack, become a single handling unit for all subsequent operations. Each truck hauls two filled-racks on separate trailers coupled in tandem to the truck tractor giving a total load of 40 bales. This unit will hereafter be referred to as a "tractor-trailer".

During each time period, a specific number of tractor-trailers is allocated to the SSLs being unloaded to perform out-and-back trips to the bio-refinery. The number of out-and-back trips that a tractor-trailer can perform during each time period depends on the distance between the SSL to which it is assigned and the bio-refinery. The tractor-trailers arrive at SSLs with empty racks, and the trailers are unhooked. These racks are then loaded with bales using a telehandler. Meanwhile, the trailers with loaded racks are towed to the bio-refinery. The tractor-trailers do not have to wait to be loaded, as the rack loading operation is decoupled from the hauling operation, improving overall efficiency of the logistics chain.

The utilization of load-out equipment sets is low, when placed permanently at an SSL. To improve resource utilization, Judd et al. [62] recommend sharing these equipment sets among a number of SSLs. At the end of each time period, a decision is made whether to keep the load-out equipment at the same SSL or move it to a different one. The load-out equipment is either driven on the road for short distances, or transported on a truck for longer distances, to move it from one SSL to the other. This transfer of load-out equipment generally takes place during the non-working day of the current week, to avoid delays and loss of productivity in the subsequent week.

The operations performed at the bio-refinery storage yard, unloading of biomass and feeding a uniform flow of material to the bio-refinery as needed, are termed the receiving-facility operations. The tractor-trailer passes over a scale, where the load is weighed and the source of biomass is noted for book-keeping purposes, before moving to the storage yard. A forklift, 10-ton capacity, is permanently stationed in a bio-refinery storage yard to unload filled racks from the trailer. The unloaded racks are either stocked in the bio-refinery storage yard or are directly placed on a rack unloader to unload bales onto a conveyor for delivery into the bio-refinery. The tractor-trailer is ready to return to the SSL for picking up the next load, after the forklift places two empty racks on the trailers. Section 2.3 provides a detailed description of the biomass feedstock logistics chain designed for the Southeastern United States.

4.3 Problem statement and literature review

The S-BLCP can be concisely defined as follows: Given a set F of pre-located SSLs, expected inflow of biomass during each time period t at each SSL i, A_i^t, planning horizon of length T, and variability in operational parameters of logistics equipment and tractor-trailers (Load-out rate ($U(\xi)$) and number of out-and-back trips ($N_i(\xi)$); where ξ, is a random variable), determine minimum fleet sizes of load-out equipment sets and tractor-trailers to meet the bio-refinery demand with required reliability level (β), so as to minimize the total cost which includes the capital investment in the equipment and the operational costs of that equipment.

The S-BLCP is a new and challenging problem that inherently contains single-item capacitated lot-sizing with parallel facilities (CLSP-PF), high-multiplicity multiple traveling salesman problem (HMmTSP), and stochasticity associated with the availability of equipment. The single-item corresponds to bales, lot sizing corresponds to the determination of the amount of biomass to be shipped from each SSL, and multiple facilities are represented by SSLs. Capacities or resources are: (1) the rate at which an equipment set can load-out biomass, (2) the rate at which a tractor-trailer can haul the biomass to the plant, and (3) the availability of biomass at each SSL. The capacity parameters for both load-out equipment sets and tractor-trailer are uncertain. Additionally, the availability of bales from period to period can vary, as can the set-up cost, which in this study is the mobilization cost of the equipment sets.

Judd et al. [62] and Aguayo et al. [1] are the only studies that consider routing of load-out equipment among SSLs for improving resource utilization. Judd et al. [62] determine location of storage facilities, number of equipment sets and their routing to minimize the overall cost, whereas Aguayo et al. [1] solve a multi-period problem to determine the following parameters: number of load-out equipment sets, their routing, number of trucks for highway hauling, which defines the amount of biomass delivered from each SSL during every period. the overall goal is to meet the bio-refinery demand while minimizing both capital and operational costs associated with the logistics chain. Both models do not consider the stochasticity involved in the system. Mixed integer linear programming models are widely used to model farmgate-to-bio-refinery supply chain problems and determine optimal decisions pertaining to location, technology selection, capital and investment, production planning, and inventory management (Akgul et al. [3], Kim et al. [66], and Zhang et al. [121]).

Modeling uncertainties is imperative in farmgate-to-bio-refinery supply chain problems, but computational complexities limit the decision making capabilities of such models. Ignoring uncertainties in the system will yield a Type-III error - "the error associated with solving the wrong problem precisely" (Hester and Adams [57]). The uncertainties present in the system are generally modeled with a stochastic programming model, simulation, or simulation-optimization approaches. Cundiff et al. [30] consider the uncertainty caused by weather in determining biomass production levels while optimizing the design of storage facilities and logistics of biomass delivery system. The authors use a multi-stage stochastic linear programming model to determine the shipment of biomass between facilities to optimize cost and also to evaluate storage expansion schedule for each feedstock producer. Sharma et al. [107] develop a scenario-based optimization model to address weather-based uncertainties to minimize cost associated with harvest, transportation, and storage. Some publications on stochastic programming-based biomass logistics models include Dal-Mas et al. [31] and Kim et al. [65].

Using a simulation modeling approach, Sokhansanj et al. [108] present a framework of a dynamic integrated biomass supply analysis and logistics model (IBSAL) to simulate operation of collection, storage and transportation of biomass feedstock to a bio-refinery. The model considers uncertainties like weather, moisture content variation, dry matter loss of feedstock, among others. Kumar and Sokhansanj [69] study cost, energy input, and carbon emissions for delivery of switchgrass to a bio-refinery for three transport options (i.e. baling, grinding, and chopping) using the IBSAL model developed in [108]. IBSAL-MC model (Ebadian et al. [41]) is based on the framework of IBSAL model [108] with a hybrid push-pull logistics system. The field operations to harvest and collect biomass within the harvest window works as a "push" logistics system, whereas the downstream operations to meet daily demand of ethanol from biorefinery operates as a “pull” system. Some other features added to the basic IBSAL model includes ability to model multiple biomass feedstocks, equipment breakdowns and realistic storage capacity against assumption of unlimited storage capacity in the basic IBSAL model. Mobini et al. [85] and Windisch et al. [116] develop discrete event simulation models for a wood-based supply chain.

Traditionally, supply chain analysis has used optimization models to design a supply chain under relevant assumptions, followed by use of simulation models to evaluate its performance against uncertainties. Another modeling approach, which optimizes the objective function via a simulation model is called simulation-optimization. The simulation-optimization approaches are comprehensively reviewed by Amaran et al. [7], Carson and Maria [22], Tekin and Sabuncuoglu [113]. Almeder et al. [6] propose a novel methodology for supply chain design problems using a combination of optimization and discrete event simulation. The simulation model includes stochastic elements, whereas the optimization model represents a simplified version of the problem without considering the uncertainties. The solution of the optimization model defines the decision rules for the discrete event simulation model. The optimization and simulation modules are run iteratively until the termination criteria is reached. This approach enables modeling of real-life problem instances by effectively incorporating dynamism and uncertainties. Bilgen and Çelebi [20] and de Keizer et al. [33] use this modeling methodology for a production scheduling and distribution planning problem in a dairy supply chain and for designing a logistics network for distributing perishable products, respectively. Marques et al. [84] develops a decision support tool for short-term planning of forest operations by combining optimization and simulation methodologies to study a raw material reception problem in a Portuguese pulp mill where the performance of deterministic schedules is analyzed through the simulation module.

Sokhansanj et al. [109] present a simulation-optimization model for the cellulosic biomass supply chain to evaluate different biomass feedstock storage systems. The optimization model is a mixed integer linear programming model that prescribes the optimum number of farms to receive farmgate contracts and their locations, optimal number of storage sites to open and their locations, and optimal assignment of farms to storage sites to minimize harvesting, collection and hauling costs. The output from the optimization model serves as an input to the simulation model. Day-to-day variations like changing moisture content or dry matter loss are not modeled in the optimization module to avoid a complex and large-sized optimization model, they are modeled as a part of the simulation model. The simulation model, which is an extension of IBSAL-MC [41], evaluates the total logistics costs, capacity of storages, number of equipment units required, resource utilization, dry matter losses, and carbon foot print for different storage systems. Ebadian et al. [42] proposes to integrate the tactical and operational planning levels in the biomass supply chain using an optimization and a simulation model. The overall objective of the integrated model is to assure that the daily biomass demand of the biorefinery is satisfied year-round at a minimum delivery cost. The optimization module produces a supply chain design for a five-year planning horizon which minimizes the logistics cost. Given the supply chain design, the simulation module simulates the biomass feedstock flow in the supply chain to meet demand at the bio-refinery. If the bio-refinery demand is not satisfied, outputs from the simulation model is used to update optimization model parameters. The optimization and simulation modules are iteratively run until the termination criteria, like convergence to a desired supply chain design, or a limit on computational time, is reached.

Our modeling methodology is similar to the one presented by Ebadian et al. [42], but we provide a more formal framework for the approach which integrates optimization and simulation methods to solve stochastic models. In addition, we add a new feature of chance constraint, that is common in supply chain problems, to address the issue of minimizing risk. Over the above, we consider a more elaborate problem that considers the routing of load-out equipment in a multi-period framework to improve resource utilization of the load-out equipment and minimize the associated capital investment in the face of uncertainties, a problem that has not been addressed in the literature, to the best of our knowledge. Such a stochastic model, if modeled and solved as a mathematical program with real-life data set, would not be tractable, and hence, a novel modeling methodology using an optimization and simulation model is presented.

4.4 Proposed modeling methodology

We first provide a general framework to explain the modeling methodology. Consider the following optimization problem.

Optimization problem:

$$\text{Minimize} \quad f(x, y, \xi) \tag{4.1}$$

subject to

$$Pr\ [x(\xi) \geq d)] \geq \beta \tag{4.2}$$

$$x(\xi) \leq \rho(\xi)\ y(\xi) \tag{4.3}$$

$$x(\xi) \geq 0 \tag{4.4}$$

$$y(\xi) \in \mathbb{Z}^{+} \tag{4.5}$$

$$\xi \in \Xi \tag{4.6}$$

where f ( ) is a linear function, ρ is stochastic parameter, and Ξ is set of scenarios

The optimization problem considered above is a stochastic integer optimization problem, based on the constraints (4.2), (4.3) and (4.5). The constraints (4.2) need not hold almost surely and instead hold with some probability. Such constraints are refereed to as probabilistic or chance constraints. Solution methodologies include scenario-based optimization, two-stage recourse, sampling methods to approximate expectation, among others. Such stochastic integer programming models, used to model real-life systems are difficult to be solved as they combine difficulty of stochastic programming and integer programming forcing practitioners to model systems as deterministic models by using average values for the stochastic parameters.

We introduce a two-phase approach using optimization and simulation to obtain acceptable solutions, which are more robust to uncertainties, than solutions obtained by solving deterministic models that use average values of stochastic parameters. Phase 1 and Phase 2 problems are as follows.

Phase 1:

$$\text{Minimize} \quad f(x, y) \tag{4.7}$$

subject to

$$x \geq d \tag{4.8}$$

$$x \leq \bar{\rho}\ y \tag{4.9}$$

$$x \geq 0 \tag{4.10}$$

$$y \in \mathbb{Z}^{+} \tag{4.11}$$

Phase 2:

$$\text{Evaluate} \quad Pr\ [x(\xi) \geq d)] \tag{4.12}$$

subject to

$$x(\xi) \leq \rho(\xi)\ \tilde{y} \tag{4.13}$$

$$x(\xi) \geq 0 \tag{4.14}$$

$\tilde{y}$ is the solution from Phase 1 problem and ξ denotes a random scenario or a replication.

In Phase 1, the probabilistic constraint (4.2) is replaced by constraint (4.8), which holds almost surely. The optimization model in Phase 1 is a deterministic mixed integer linear programming model, which yields solutions $\tilde{x}$, and $\tilde{y}$. Phase 2 of the approach contains only an evaluation problem and uses a simulation model to test the solution obtained from Phase 1 against different realizations of the stochastic parameter $\rho(\xi)$ whose probability distribution is assumed to be known. The simulation model is run for a number of replications, where each replication denotes a random scenario ξ. The probability with which the constraint (4.2) holds is evaluated at the end of the simulation run, and if it is less than β, then the average of $\rho(\xi)$ across replications (i.e., $\bar{\rho}$) is evaluated and updated in the Phase 1 problem. The problems in Phases 1 and 2 are iteratively solved until constraints (4.2) in the original problem is satisfied. Next, we show that, this convergence occurs in a finite number of steps.

Proposition 4.4.1. *The probability with which constraint (4.2) holds increases monotonically and the procedure converges to the required state in finite number of steps.*

Proof. While sampling, only the realizations of $\rho(\xi)$ from the previous iteration that are below $\bar{\rho}$ are considered to evaluate the average value ($\bar{\rho}$) for the current iteration. This gives rise to a non-increasing sequence of $\bar{\rho}$ across iterations, i.e., $\bar{\rho}_1 > \bar{\rho}_2 > \bar{\rho}_3 > \bar{\rho}_2 > ..... > \bar{\rho}_n$. The constraints (4.9) is like a capacity-resource constraint, and hence, decreasing $\bar{\rho}$ (capacity) results in an increment of y (resource), i.e., $y_1 \leq y_2 \leq y_3 \leq ..... \leq y_n$. Thus, the probability with which the constraint (4.2) holds monotonically increases, as the value of y is non-decreasing across iterations. Since a finite number of resource is required to satisfy demand, the procedure will converge to the required state in a finite number of steps. ■

Next, we illustrate this modeling methodology with a simple example.

Example : Let N be the set of n jobs to be completed within a due date d_j. The processing time $\tilde{p_j}$ for these n jobs are stochastic and follows a probabilistic distribution that is assumed to be known. We want to determine the minimum number of machines M to be employed to ensure that all n jobs are completed within the due date d_j with a probability greater than or equal to β.

Let set N consist of 6 jobs with processing time $\tilde{p_j}$ be uniformly distributed, U(1.5, 2.5) hours $\forall j \in N$ and let the due date, $d_j = 6$ hours $\forall j \in N$. The number of machines M to be employed can be determined by considering expected values of the processing time which is 2 hours per job. Hence, the total processing time is 12 hours and 2 machines are required to complete the jobs as per the calculation below.

Processing time $\tilde{p_j}$ = Uniform (1.5, 2.5) hours $\quad \forall j \in N$

Expected value of processing time $\bar{p}$ (in hours) = 2

Total processing time (in hours)= 12

Number of machines required $= \dfrac{12}{6} = 2$

The probability of completing all 6 jobs within the due date using 2 machines in the face of uncertainties is evaluated using a simulation model for different realizations of stochastic processing time $\tilde{p_j}$ and is provided in Table 4.1.

Table 4.1: Parameters evaluated from simulation model.

Parameters	**Values**
Probability of completing all 6 jobs	28%
Probability of completing 5 or more jobs	83%
Probability of completing 4 or more jobs	100%
Min. (Max.) total processing time (in hours) across different scenarios	10.23 (14.43)
Average total processing time (in hours) across different scenarios	12.44
Biased average total processing time $^{+}$ (in hours)	12.80

+Average of all random realizations of total processing times above 12 hours

The number of machines M can now be revised using the biased average total processing time; $M = \left\lceil \dfrac{12.80}{6} \right\rceil = 3$

Hence, the minimum number of machines M for completing all 6 jobs within the due date with a probability of β, or greater, is 3. On testing the revised solution again with the simulation model, the probability of completing 6 jobs with a three machine configuration is found to be 100%.

The general optimization model can also be solved as a mathematical program by replacing the probabilistic constraints with its deterministic equivalent and then solving it as a stochastic program. To do this, the distribution of the stochastic variable should be known and the distribution should satisfy some properties, so that the replaced constraints preserve the convexity of the problem. Hence, some complex stochastic optimization problem with chance constraint cannot be solved by the standard mathematical programming-based methodologies, especially when the distribution of stochastic variable is not known *a priori*.

Some open questions about a stochastic optimization methodology are: (1) Can this methodology be used to model a general stochastic problem, (2) If not, what are the types of stochastic problems which can use this modeling methodology, and (3) How exact is the solution, compared to the global optimum solution, among others. In the face of these difficulties, we investigate the use of a simulation-optimization approach for the solution of the stochastic - biomass logistics chain problem (S-BLCP).

Next, we formulate the S-BLCP as a stochastic integer program with chance constraints.

Consider the following notation.

Sets:

F : Set of SSLs.

F^0 : Set of SSLs including the bio-refinery, which is denoted by 0 ($0 \cup F$).

T : Length of the planning horizon $(1, ..., T)$.

T^0 : Length of the planning horizon including time 0, $(0, 1, ..., T)$.

T^+ : Length of the planning horizon including time T+1, $(1, ..., T+1)$.

Ξ : Set of Scenarios (ξ).

Parameters:

c_e : Load-out equipment purchase price ($ per unit).

c_v : Tractor-trailer purchase price ($ per unit).

$\hat{c}_i^t$: Shipping cost from SSL i to the bio-refinery in period t ($ per Mg), $i \in F$.

c_{ij}^t : Mobilization cost to transport load-out equipment from SSL i to SSL j in period t ($ per km), $\forall i \in F$, $\forall j \in F$.

A_i^t : Amount of biomass available at SSL i during time period t (Mg), $\forall i \in F, \quad \forall t = 0, ..., T$.

α : Fraction of biomass inventory lost (Storage loss).

Q : Tractor-trailer capacity (Mg).

$N_i(\xi)$: Number of out-and-back trips a vehicle can perform during a time period if allocated to SSL i for scenario ξ, $\forall i \in F$, $\xi \in \Xi$.

$U(\xi)$: Maximum load-out rate for each equipment set (Mg per time unit) for scenario ξ, $\xi \in \Xi$.

β : Reliability level for delivering biomass to bio-refinery ($0 < \beta \leq 1$).

P^t : Plant requirement during time period t (Mg), $\forall t = 1, ..., T$.

E : Maximum number of equipment sets to allocate to an SSL during a time period.

Decision Variables :

K : Number of load-out equipment sets purchased.

V : Number of tractor-trailers purchased.

$s_{i\xi}^t$: Amount of biomass shipped from SSL i to the bio-refinery during period t for scenario ξ, $\forall i \in F, \quad \forall t = 1, ..., T$, $\xi \in \Xi$.

$y_{i\xi}^t$: Inventory of biomass at the end of time period t at SSL i for scenario ξ. $\forall i \in F, \quad \forall t = 0, ..., T$, $\xi \in \Xi$.

$v_{i\xi}^t$: Number of tractor-trailers assigned to SSL i during time period t for scenario ξ, $\forall i \in F, \quad \forall t = 1, ..., T$, $\xi \in \Xi$.

$z_{i\xi}^t$: Number of equipment sets assigned to SSL i during time period t for scenario ξ, $\forall i \in F, \quad \forall t = 1, ..., T$, $\xi \in \Xi$.

$x_{ij\xi}^t$: Integer variable that indicates the number of equipment sets traversing from SSL i to SSL j at the end of time period t for scenario ξ. $\quad \forall i \in F^0, \quad \forall j \in F^0, \forall t = 0, ..., T$, $\xi \in \Xi$.

Model S-BLCP:

$$\text{Minimize} \;\; c_e K + c_v V + E_\xi\ [\sum_{t=1}^{T+1} \sum_{i \in F^0} \sum_{j \in F^0} c_{ij}^t x_{ij\xi}^{t-1} + \sum_{t=1}^{T} \sum_{i \in F} \hat{c}_i^t s_{i\xi}^t] \tag{4.15}$$

subject to :

$$Pr(\sum_{i \in F} s_{i\xi}^t \geq P^t) \geq \beta, \;\; \forall t = 1, ..., T \tag{4.16}$$

$$y_{i\xi}^t = (1-\alpha)\ y_{i\xi}^{t-1} - s_{i\xi}^t + A_i^t, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.17}$$

$$y_{i\xi}^0 = A_i^0, \;\; \forall i \in F, \;\; \forall \xi \in \Xi \tag{4.18}$$

$$s_{i\xi}^t \leq Q\ N_i(\xi)\ v_{i\xi}^t, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.19}$$

$$\sum_{i \in F} v_{i\xi}^t \leq V, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.20}$$

$$\sum_{i \in F} z_{i\xi}^t \leq K, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.21}$$

$$s_{i\xi}^t = U(\xi)\ z_{i\xi}^t \;\; , \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.22}$$

$$\sum_{j \in F^0} x_{ji\xi}^{(t-1)} - z_{i\xi}^t = 0, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.23}$$

$$\sum_{j \in F^0} x_{ji\xi}^{(t-1)} - \sum_{j \in F^0} x_{ij\xi}^t = 0, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.24}$$

$$x_{i0\xi}^t = 0, \;\; \forall i \in F, \;\; \forall t = 0, ..., T-1, \;\; \forall \xi \in \Xi \tag{4.25}$$

$$s_{i\xi}^t \geq 0, \;\; \forall i \in F, \;\; \forall t = 1, ...T, \;\; \forall \xi \in \Xi \tag{4.26}$$

$$z_{i\xi}^t \;\; \in \{0, 1, ..., E\}, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.27}$$

$$y_{i\xi}^t \geq 0\ , \;\; \forall i \in F, \;\; \forall t = 0, ..., T, \;\; \forall \xi \in \Xi \tag{4.28}$$

$$v_{i\xi}^t \in Z^+, \;\; \forall i \in F, \;\; \forall t = 1, ..., T, \;\; \forall \xi \in \Xi \tag{4.29}$$

$$x_{ij\xi}^t \in \{0, 1, ..., E\}, \;\; \forall i \in F^0, \;\; \forall j \in F^0, \;\; \forall t = 0, ..., T, \;\; \forall \xi \in \Xi \tag{4.30}$$

$$V \in Z^+, \;\; K \in Z^+ \tag{4.31}$$

The objective function (4.15) minimizes the total cost incurred by the purchase of load-out equipment sets, tractor-trailers, and expectation of cost related to mobilization of equipment, and shipping of biomass from SSLs to the bio-refinery across different scenarios. Constraints (4.16) ensure that the bio-refinery demand during each time period t is satisfied with reliability β. Constraints (4.17) capture the inventory balance during each time period for all scenarios. Constraints (4.18) correspond to the initial inventory at SSL i in period 0. Constraints (4.19) ensure that the amount shipped from SSL i to the bio-refinery in each time period t is at most equal to the number of tractor-trailers ($v_j^t(\xi)$) allocated to this SSL times their capacity (Q), and the number of trips ($N_i(\xi)$) for scenario ξ. Constraints (4.20) and (4.21) provide, respectively, an upper bound for the maximum number of tractor-trailers and load-out equipment sets available for each realization of random scenario. Constraints (4.22) ensure that the bales can be shipped from SSL i to the bio-refinery only if there is a load-out equipment set allocated to that SSL. Furthermore, it makes sure that a load-out equipment set is allocated to an SSL during time period t if and only if it is fully utilized during each time unit. Constraints (4.23) enforce mobilization of required number of equipment sets from SSL $j \in F^0$ to SSL $i \in F$ at the end of a time period t for scenario ξ. Note that we permit an equipment set to stay at an SSL if needed. Constraints (4.24) are the standard flow conservation constraints for equipment sets. Constraints (4.25) enforce an equipment set to stay at an SSL and not move to the bio-refinery until the last time period. Constraints (4.26) - (4.31) define the domains of the variables.

The S-BLCP model determines the number of load-out equipment sets and tractor-trailers required for meeting bio-refinery demand with a reliability level β while minimizing cost. The S-BLCP model can be solved as two-stage stochastic integer program, in which the first stage is to determine the fleet sizes of the load-out equipment sets and tractor-trailers and in the second stage, a capacitated lot sizing and routing problem is solved for different realizations of uncertain parameters to minimize the expected value of cost. Constraints (4.16) represent the probabilistic or chance constraint and these constraints dictate feasibility of the first stage decisions. The second-stage problem is a NP-hard integer program and involves significant computational difficulties even for a single scenario. Table 4.2 is a part of the results in Aguayo et al. [1] that show a comparison between direct solution of S-BLCP model for a single scenario by CPLEX and solution by branch-and-price (B&P) method, indicating significant computational effort required for the second-stage problem. Thus, the S-BLCP may not be tractable, if solved as a stochastic integer program, and the proposed modeling methodology is applied.

Table 4.2: A comparison between direct solution by CPLEX and B&P approach [1].

Instance		Root Node			Branch-and-Price			CPLEX	
#	Zlb	ZIP	Opt. Gap	Time	Nodes	ZIP	Opt. Gap	ZIP	Opt. Gap
15	4273208	4445528	3.90%	901.99	8	4382855	2.50%	-	-
16	4102807	4481690	8.50%	750.15	9	4422982	7.20%	-	-

The overall model includes an optimization module (Phase 1), simulation module (Phase 2), database and an excel interface as shown in the Figure 4.1.

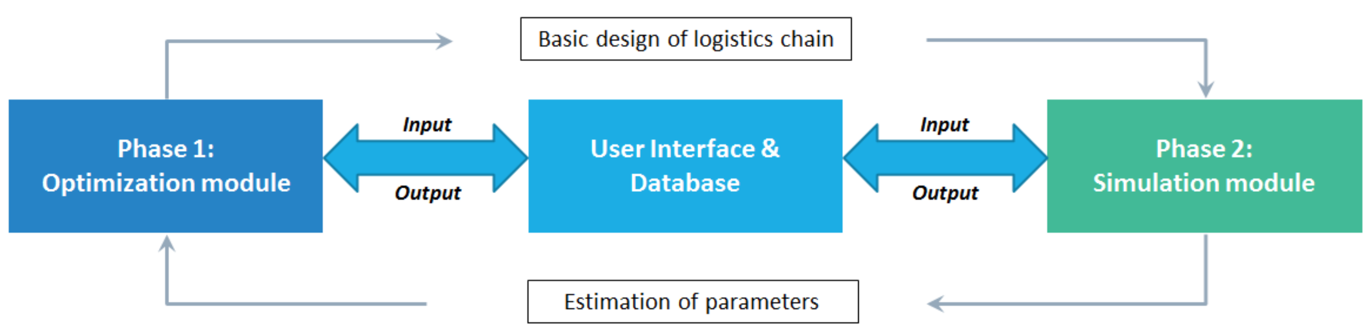


Figure 4.1: Framework of two phase optimization-simulation approach.

Figure 4.2 provides a flowchart of the proposed two-phase optimization-simulation approach applied to the S-BLCP. The Phase 1 problem is a mixed integer linear programming model, which provides the basic design of biomass logistics chain. In Phase 2, a simulation model is used to test the performance of the Phase 1 design under stochasticity and evaluate the probability of meeting demand. If the probability of meeting demand is not met with required reliability level, capacity parameters are estimated from the simulation model and updated in the Phase 1 problem. The process is iteratively carried out until the solution achieves a required reliability level.

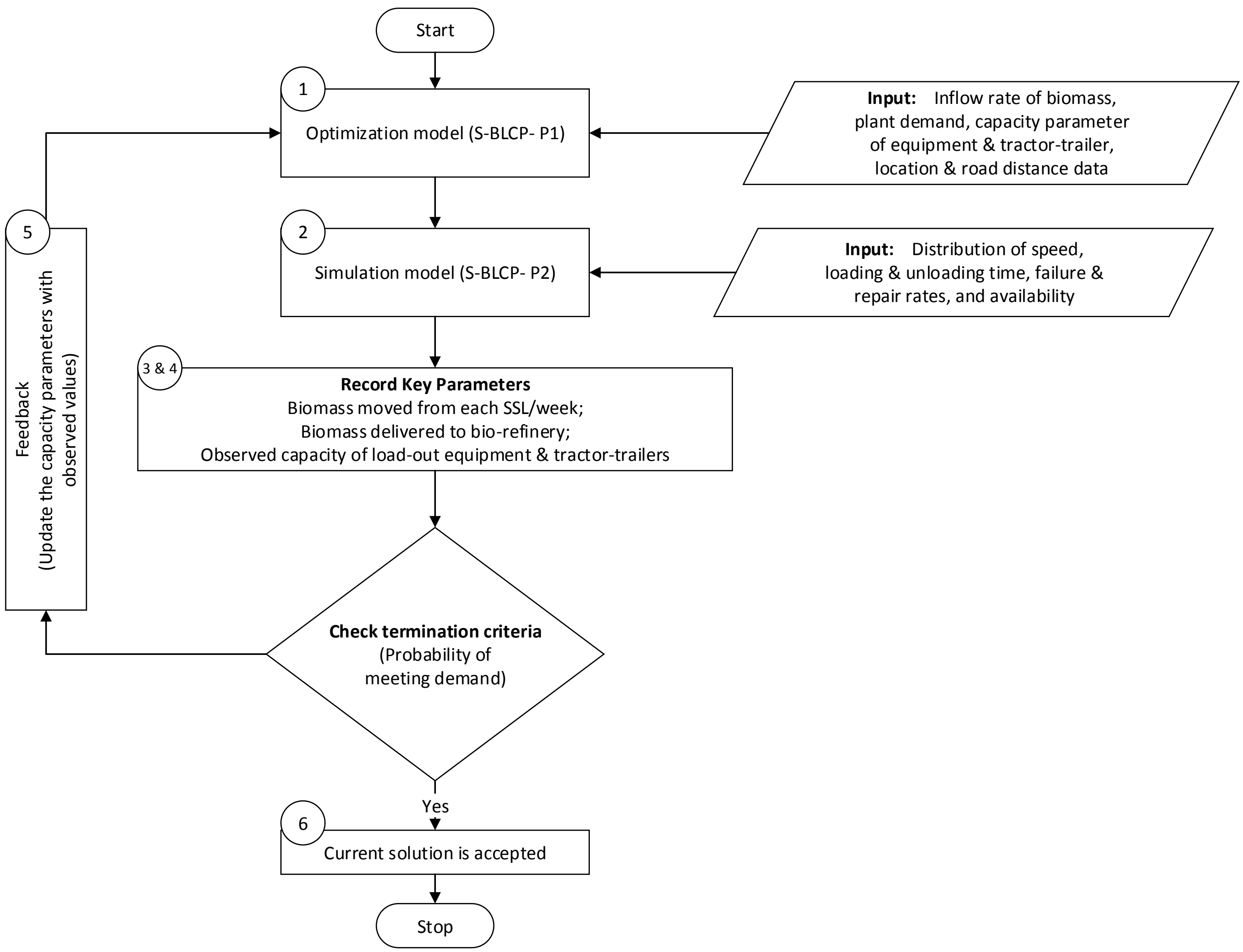


Figure 4.2: A flowchart of the proposed two-phase optimization-simulation approach.

The sequence of steps in the iterative process is as follows:

- **Step 1:** Run the optimization module for average value scenario. To start the iterative procedure, the Phase 1 optimization module is run considering average values of stochastic parameters. The results from the Phase 1 problem like fleet size, biomass flow and routing of equipment are recorded. These details give a preliminary biomass logistics chain design.
- **Step 2:** The Phase 2 problem modeled as a simulation module is run based on Phase 1 problem results and the pre-determined design of logistics chain is tested against stochasticity.
- **Step 3:** Output statistics from the simulation module such as the biomass loaded in the racks, and biomass delivered to the bio-refinery, among others, are recorded in the database.
- **Step 4:** The output statistics recorded in Step 3 is used to re-evaluate capacity parameters used in the optimization module
- **Step 5:** Run the optimization module with updated capacity parameters from Step 5. Repeat steps from 2 to 5 iteratively.
- **Step 6:** Stop the iteration when the required reliability level is achieved.
- **Step 7:** Record the final results. The outputs of the last run of optimization and simulation modules are recorded on an excel spreadsheet.

4.4.1 Phase 1: Optimization module

The optimization module contains a mixed integer programming model (S-BLCP-P1), solution to which prescribes the design of the biomass logistics chain. The mixed integer programming model (S-BLCP-P1) is formulated as below.

Model S-BLCP-P1:

$$\text{Minimize} \;\; c_e K + c_v V + \sum_{t=1}^{T+1} \sum_{i \in F^0} \sum_{j \in F^0} c_{ij}^t x_{ij}^{t-1} + \sum_{t=1}^{T} \sum_{i \in F} \hat{c}_i^t s_i^t \tag{4.32}$$

subject to :

$$\sum_{i \in F} s_i^t \geq P^t, \quad \forall t = 1, ..., T \tag{4.33}$$

$$y_i^t = (1-\alpha) y_i^{t-1} - s_i^t + A_i^t, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.34}$$

$$y_i^0 = A_i^0, \quad \forall i \in F \tag{4.35}$$

$$s_i^t \leq Q N_i v_i^t, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.36}$$

$$\sum_{i \in F} v_i^t \leq V, \quad \forall t = 1, ..., T \tag{4.37}$$

$$\sum_{i \in F} z_i^t \leq K, \quad \forall t = 1, ..., T \tag{4.38}$$

$$s_i^t = U z_i^t \quad , \forall i \in F, \quad \forall t = 1, ..., T \tag{4.39}$$

$$\sum_{j \in F^0} x_{ji}^{(t-1)} - z_i^t = 0, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.40}$$

$$\sum_{j \in F^0} x_{ji}^{(t-1)} - \sum_{j \in F^0} x_{ij}^t = 0, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.41}$$

$$x_{i0}^t = 0, \quad \forall i \in F, \quad \forall t = 0, ..., T-1 \tag{4.42}$$

$$s_i^t \geq 0, \quad \forall i \in F, \quad \forall t = 1, ...T \tag{4.43}$$

$$z_i^t \quad \in \{0, 1, ..., E\}, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.44}$$

$$y_i^t \geq 0 \;, \quad \forall i \in F, \quad \forall t = 0, ..., T \tag{4.45}$$

$$v_i^t \in Z^+, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{4.46}$$

$$x_{ij}^t \in \{0, 1, ..., E\}, \quad \forall i \in F^0, \quad \forall j \in F^0, \quad \forall t = 0, ..., T \tag{4.47}$$

$$V \in Z^+,\ K \in Z^+ \tag{4.48}$$

The Phase 1 (optimization module) determines the optimal number of load-out equipment sets and tractor-trailers, amount of biomass delivered from each SSL and routing of both load-out equipment sets and tractor-trailers during each time period. These outputs are stored in the database to facilitate access by the simulation module. The proposed mixed integer problem contains high-multiplicity multiple traveling salesmen problem (HMmTSP), an extension of the well-known traveling salesman problem (TSP) associated with routing of load-out equipment. The large number of integer decision variables from the routing feature, makes the problem computationally complex to solve with direct application of a commercial solver. Constraints (4.39) is an equality constraint and ensures that the load-out equipment alloted to the SSL is fully utilized, or, in other words, shipment from each SSL is equal to the capacity of load-out equipment sets stationed at that SSL. Hence, the constraints (4.33) and (4.39) would force the total shipment during every time period to be greater than or equal to the bio-refinery requirement and allot additional tractor-trailers to do this. Though, additional tractor-trailers are allotted, this increases the overall reliability of the logistics system to fulfill actual bio-refinery demand. Replacing the constraint (4.39) with an inequality constraint would allow partial shipment from SSLs and thereby lower number of tractor-trailers, but doing so would incur severe computational burden.

A branch-and-price based approach, developed by Aguayo et al. [1] is used to solve this mixed integer programming model. Computational investigation in Aguayo et al. [1] reveals the ability of this approach to obtain near-optimal solutions for large-sized problem instances. The model is implemented in Microsoft Visual Studio 2012, using a branch-and-price algorithm and CPLEX 12.6 solver.

4.4.2 Phase 2: Simulation module

The simulation module (S-BLCP-P2) models the stochasticity associated with the operation of logistics equipment such as load-out equipment and tractor-trailers. Modeling these uncertainties into the optimization module will make it computationally complex, but at the same time, these uncertainties cannot be ignored while determining a cost-effective fleet size of logistics equipment. The simulation module considers the design of the biomass logistics chain dictated by the solution of the Phase 1 optimization module and simulates all logistics activities to load-out from SSLs and deliver feedstock to the bio-refinery to meet their weekly demand schedule. The simulation module is build on an academic version of a simulation software - Simio.

4.4.2.1 Simio Software

Simio is an object-based simulation software, which uses different inbuilt and custom-made objects to build a simulation model. The standard library of these objects (Pegden and Sturrock [92]) are described in Table 4.3. General programming languages such as C++, or Java, among others can be used to develop a simulation module. Using general programming languages to develop a simulation model would provide flexibility and ability to build a completely self-designed model; however, advanced programming expertise is required to do so. Commercially available simulation software like Simio, Arena, and Simul8 provide the easiest and simplest way to develop simulation models but at the expense of flexibility and ability to implement fairly large and complex models. Pegden and Sturrock [92, 93] describe the Simio simulation software, its components, and modeling methodology.

4.4.2.2 Simulation model

The simulation model seeks to study the uncertainties associated with the load-out operation of the SSL and the highway hauling operation (operation of tractor-trailers to deliver the

Table 4.3: Standard library of objects in Simio.

Object	**Description**	**Main Properties**
Source	Generates entity objects of a specified type and arrival pattern.	Entity Arrival Logic, Stopping Conditions, Table Reference Assignments.
Sink	Destroys entities that have completed processing in the model.	Process Logic.
Server	Represents a capacitated process such as a machine or service operation.	Process Logic, Buffer Capacity, Reliability Logic, Secondary Resources.
Workstation	Models a complex workstation with setup, processing, and teardown phases and secondary resource and material requirements.	Process Logic, Buffer Capacity, Reliability Logic, Secondary Resources, Materials and Other Constraints.
Combiner	Combines multiple member entities together with a parent entity (e.g. a rack filled with bales)	Matching Logic, Process Logic, Buffer Capacity, Reliability Logic, Secondary Resources.
Separator	Splits a batched group of entities or makes copies of a single entity.	Separation Logic, Process Logic, Buffer Capacity, Reliability Logic, Secondary Resources.
Resource	A generic object that can be seized and released by other objects.	Resource Logic, Reliability Logic.
Vehicle	A transporter that can follow a fixed route or perform on demand transport pickups/drop offs. Additionally, an On Demand routing type vehicle may be used as a movable resource that is seized and released for non-transport tasks.	Transport Logic, Travel Logic, Routing Logic, Resource Logic, Reliability Logic, Population.
Worker	A movable resource that may be seized and released for tasks as well as used to transport entities between node locations.	Transport Logic, Travel Logic, Routing Logic, Resource Logic, Reliability Logic, Population.
Basic Node	Models a simple intersection between multiple links.	Crossing Logic, Routing Logic, Tally Statistics.
Transfer Node	Models a complex intersection for changing destination and travel mode.	Crossing Logic, Routing Logic, Transport Logic, Tally Statistics.
Connector	A simple zero-time travel link between two nodes.	Routing Logic.
Path	A link over which entities may independently move at their own speeds.	Routing Logic, Travel Logic.
Time Path	A link that has a specified travel time for all entities.	Routing Logic, Travel Logic.
Conveyor	A link that models both accumulating and non-accumulating conveyor devices.	Routing Logic, Travel Logic, Reliability Logic.

feedstock). Hence, the operations from the SSLs to the receiving-facility are considered for modeling S-BLCP-P2. The feedstock bales are considered to be entities in the simulation model. The process logic of the simulation model is depicted in Figure 4.3.

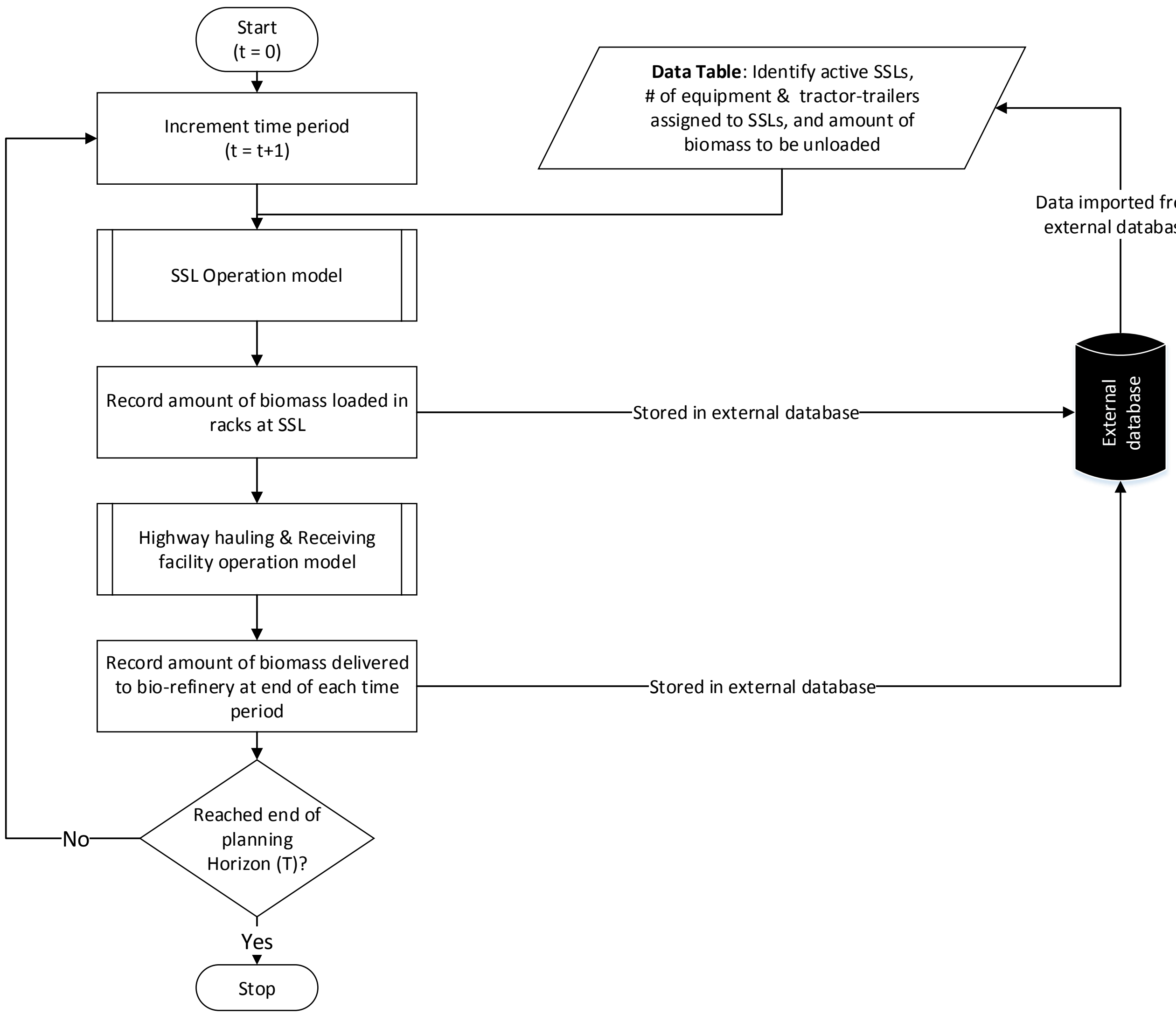


Figure 4.3: Process logic of S-BLCP-P2.

A data table connected to an external database provides details like which of the SSLs are open during a particular time period, the number of load-out equipment sets and tractor-trailers assigned to the open SSLs, and the amount of biomass that needs to be unloaded from each SSL during every time period. The simulation model is run till the end of the planning horizon (T). After each time period, parameters like amount of biomass loaded on racks at each active or open SSL and amount of biomass delivered to bio-refinery are recorded and stored in an external database. An "open" SSL is defined as an SSL where load-out operations are occurring. The observed capacity parameters are estimated from these details to be updated in the Phase 1 optimization problem. The process logic of the SSL operation, highway hauling and receiving facility operation is depicted in Figures 4.4 and 4.5.

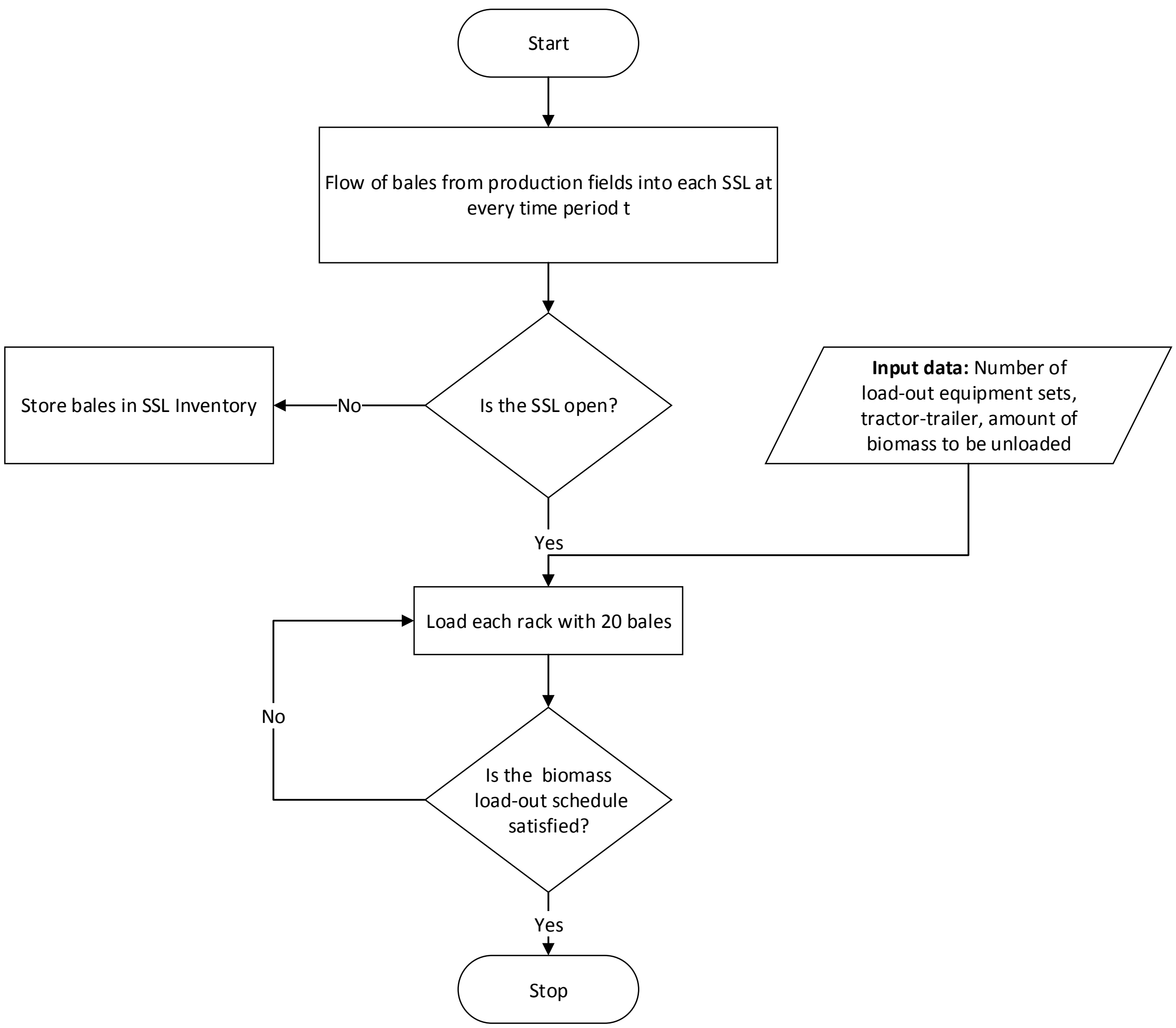


Figure 4.4: Process logic of SSL operation in the simulation model.

The entities representing the incoming inventory of biomass is stored in a queue. The data table is accessed to determine inputs corresponding to a particular time period such as the SSLs that are open, number of load-out equipment sets and tractor-trailers associated with the SSL and amount of biomass that needs to be unloaded from the SSL. The "vehicle object" representing the load-out equipment continuously loads each rack until its capacity (20 bales) is reached. The SSL operation takes place 10 hours a day and six days a week.

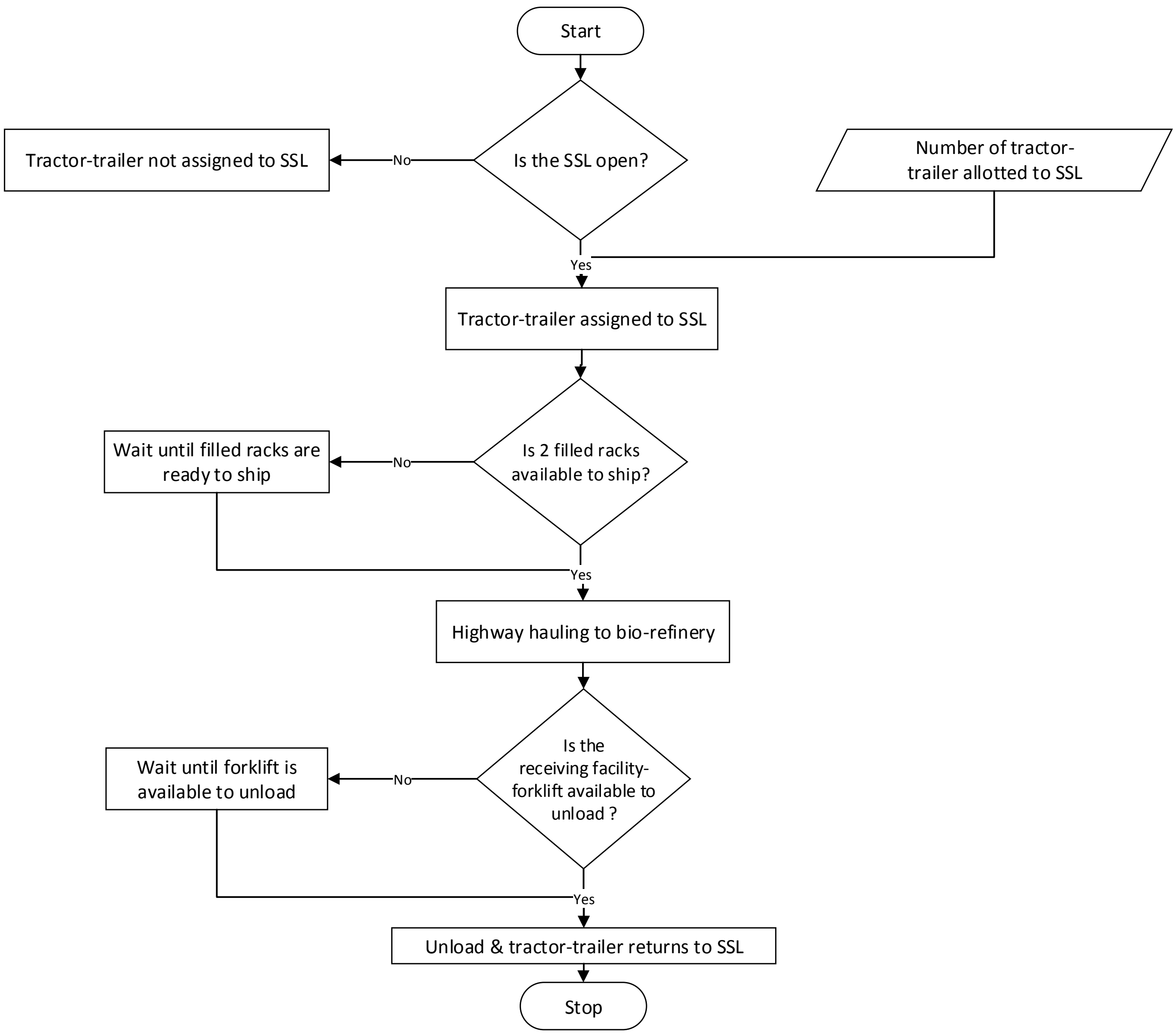


Figure 4.5: Process logic of highway hauling and receiving facility operation in the simulation model.

A "combiner object" is used to batch entities, representing bales, into racks. The filled racks are then transported to the bio-refinery, two at-a-time, using a vehicle object that represents tractor-trailer. The tractor-trailer may have to wait to unload the racks at the receiving facility as there might be queue buildup of tractor-trailers waiting to unload racks at the bio-refinery. After unloading the racks, they return to the allotted SSL to pick up the next load, and this process is continued till the end of the time period or till the biomass load-out schedule is completed. Stochasticity is modeled in the SSL and highway hauling operations by considering a distribution for operating parameters such as road speed, loading and unloading time for both trucks and load-out equipment. Breakdown and unscheduled stops are also modeled in the simulation model.

4.4.2.3 Verification and validation

Law [75] describes a seven step approach to design and conduct a successful simulation study. The seven steps include formulating the problem, collecting information to construct a conceptual model, validating the conceptual model, implementing the conceptual model as a computer simulation, verification and validation of the simulation model, and finally conducting experiments and documenting results. Verification is the process of determining whether the conceptual model (process logic) has been correctly translated into a computerized simulation model, whereas validation is the process of determining whether a simulation model is an accurate representation of the system, for the particular objectives of the study. Only a verified and validated simulation model is used to make decisions by conducting experiments on the simulation model. Sargent [100] describes different approaches for verifying and validating a simulation model. Some approaches used to verify and validate S-BLCP-P2 include the following:

1. **Animation:** Simio simulation software supports 2D and 3D animation of the simulation model. The operational behavior of the simulation is studied to verify the model and make necessary changes if necessary.

2. **Comparison to other models:** The integrated approach is used to make tactical decisions on fleet size of load-out equipment sets and tractor-trailers. The decision on the fleet size of equipment was compared with that obtained using the optimization model proposed by Judd et al. [62] and was found to be comparable.

3. **Event validity:** The events occurring in the simulation model are compared with those of the real system to determine if they are similar.

4. **Extreme condition test:** All stochastic parameters are replaced by their averages and results of simulation are compared with results from the deterministic optimization module. Also, parameters are modified to test full capacity of equipment and logistics design before introducing distributions for the stochastic parameters.

5. **Face validity:** Asking individuals knowledgeable about the system whether the model and/or its behavior are reasonable. The process logic of the conceptual model and input-output relationship of the simulation model was evaluated.

6. **Traces:** The behavior of entities in the simulation model is traced through the model to determine if the model's logic is correct and if necessary accuracy is obtained.

7. **Data validity:** Process logic and data are the two most important components in the simulation model. The data set used for validation of the model was initially developed in Resop et al. [96]. Precise data for the equipment parameters were not available, and they were set from reviewing the related literature [29] and from asking individuals knowledgeable about the system. This approach attempts to model the biomass logistics system using the rack concept for multi-bale handling. Several specialized pieces of equipment like the telehandler attachment for picking round bales, racks, rack unloader, among others are being developed concurrently along with the described decision support tool explained in this chapter. As the physical systems are not ready for an extensive data collection and input modeling, we have described the input modeling procedure in Appendix A. The user is expected to collect data and

perform input modeling to use the proposed approach.

8. **Sensitivity analysis:** When a real world system does not exist to compare and validate behavior of the simulation model, sensitivity of different parameters are tested to establish that simulation model behaves as expected.

4.4.2.4 Replication analysis

The number of replications run for the simulation model is decided based on the relative error of the output statistics critical to the analysis. Relative error for the average total shipment of biomass is plotted in Figure 4.6. Each run of the simulation module (a simulation experiment) is run for 30 replications to keep the relative error below 2%, considering the computational burden for performing large number of replications against reduction in relative error of the output statistics.

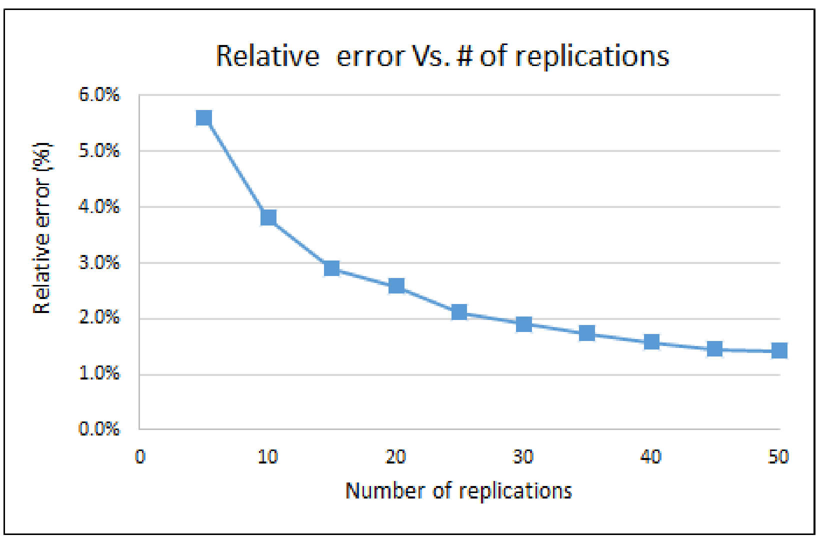


Figure 4.6: Replication analysis.

4.5 Results and Discussion

4.5.1 Results

The dataset used to illustrate the proposed methodology was developed in Resop et al. [96]. A geographical information system (GIS) was used to locate potential fields for switchgrass production. A total of 199 storage locations (SSL) are proposed over a 48-km radius around Gretna, VA as shown in Figure 4.7. The total production area is divided into 5 sections each containing 48, 23, 39, 32, and 57 SSLs, respectively and have approximately the same total production of feedstock. Each production field is assigned to a unique SSL for the entire planning horizon, and one SSL can cater to many production fields.

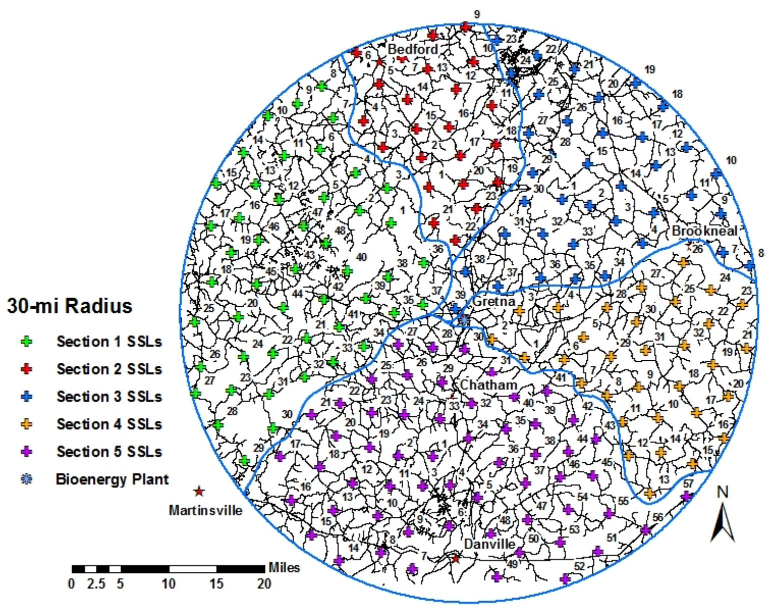


Figure 4.7: Two-dimensional picture displaying SSL(s) and bio-refinery at Gretna, VA. (Source: Liu et al. [79], InTech, Open Access, 2013)

The flow of biomass to each SSL depends on the production yield and the harvest scenario followed. A detailed study on different possible harvesting scenarios is available in Grisso et al. [53]. Here, in-flow rates into each SSL during each week is based on expected harvest schedule, based on typical weather as shown in Table 4.4 ("Typical" weather defines the expected number of days in a given month that harvest operations can proceed). The percentage given are the percent of total annual harvest from production fields assigned to a given SSL that is harvested in the given month. The inflow of biomass feedstock from production fields to SSL occur during all months except April (Month 10), May (Month 11), and June (Month 12). A one year planning horizon is considered with weekly time periods (48 weeks) and biomass inflow into each SSL is assumed to be equally distributed over the weeks of a month.

Table 4.4: Biomass inflow as a percentage of total annual harvest.

Month	1	2	3	4	5	6	7	8	9	10	11	12
% of annual harvest	1%	6%	13%	21%	17%	13%	9%	9%	10%	0%	0%	0%

The optimization module minimizes the total cost which include the purchase price of load-out equipment sets and truck tractors, the cost of mobilizing load-out equipment sets and the truck hauling cost. The purchase price of load-out equipment considered is $94,000 per equipment set. Since the production fields are within a 48-km radius, used truck-tractors with sufficient life can be considered for highway hauling operations. A new truck-tractor can cost about $85,000, whereas an used truck-tractor may cost between $45,000 and $55,000 depending on the left-over service life. In this study, we consider use of second-hand truck tractors with a purchase cost of $50,000. The cost of specially designed trailers and racks are not considered in this study.

A mobilization cost is defined as the cost to move load-out equipment, specifically the telehandler, from one SSL to the next. The following rules were used to estimate mobilization cost: (1). If the next SSL is not more than 5 miles away, the telehandler will be driven to the next SSL. Operational cost of the telehandler is $39.50/h, and the travel time is based on the road speed of 17 mi/h, and (2). The cost incurred for the truck and trailer to haul the telehandler is $ 125/h, which includes the labor cost for the truck operator. Total mobilization time consists of the time required to travel from the bio-energy plant to the current SSL, time to load the telehandler, travel time to the next SSL, time to unload the telehandler, and the travel time back to the bio-energy plant. Similarly, the biomass hauling cost (in $ per km) is calculated based on the equation given in Judd et al. [62], 0.1381 x d_j +1.6667, where d_j is the road distance (in km) from SSL j to bio-refinery. All the other parameters considered for this data set are summarized in Table 4.5.

Table 4.5: Parameters used for simulation of 199 SSL dataset.

Sr. No.	**Parameter**	**Values**
1	Bio-refinery requirement (per week)	6048 Mg
2	Rack capacity	8 Mg
3	Tractor-trailer capacity	16 Mg
4	Tractor-trailer loading time at SSL (avg.)	12.5 Minutes
5	Tractor-trailer unloading time at bio-refinery (avg.)	12.5 Minutes
6	Tractor-trailer speed on highways (avg.)	45 mph (72 km/h)
7	Load-out equipment capacity per 2-bale cycle	0.8 Mg
8	Load-out equipment time to pick bales from SSL (avg.)	45 seconds
9	Load-out equipment time to load bales into rack (avg.)	45 seconds
10	Load-out equipment average speed in SSL (avg.)	4 mph
11	Distance load-out equipment would move within SSL per trip	0.1 miles
12	Weight of bale	0.4 Mg
13	Work hours per day	10 hours
14	Working days per week	6 days
15	Availability of load-out equipment	80%
16	Availability of truck tractors	90%

The unscheduled breaks by operators, refueling, preventive maintenance, breakdowns are accounted in the availability of load-out equipment and tractor-trailers. The parameters like loading, unloading time, average road speed are stochastic. The variability in these parameters affect the overall capacity parameters, and hence, the fleet size of equipment and trucks. The loading/unloading time (in seconds) and speed (in mph) of the telehandler is assumed to follow uniform(30,60) and uniform (3,5) distributions, respectively. Similarly, loading/unloading time (in minutes) and speed (in mph) of the truck is assumed to follow uniform(10,15) and uniform (40,50) distributions, respectively. The user is expected to perform input modeling of these stochastic parameters using the procedure illustrated in Appendix A, to use this methodology rather than using the assumptions mentioned above. The assumptions are made only to illustrate this methodology.

Figure 4.8 gives the probability of meeting bio-refinery demand across each iteration. The solution to the deterministic model considering average values of stochastic parameters corresponds to the iteration 1 and its probability of meeting 100% of bio-refinery demand every week is only 14.8%. This shows that the approximation approach of using average values for stochastic parameters yields solutions that are not robust to stochasticity and dynamics of the real-world system. The reliability of decision variables for meeting bio-refinery demand improves from 14.8% to 100% over three iterations.

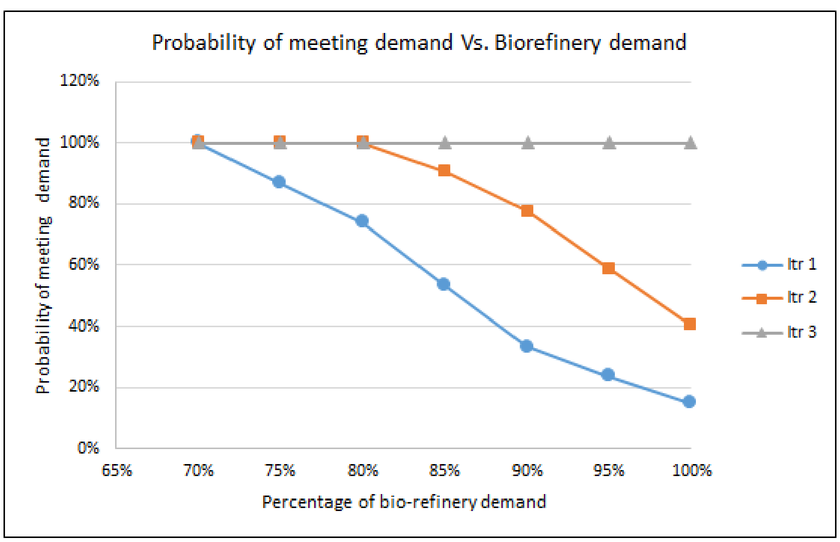


Figure 4.8: Probability of meeting demand.

The Table 4.6 describes the fleet size of the biomass logistics chain during each iteration and its associated capital investment. The graph in Figure 4.9 summarizes reliability of meeting bio-refinery demand and associated capital investment required for different solutions. The approach presents different options for the users with varied risk level and investment requirements. The capital cost associated with each iteration is increasing monotonically for the data set under consideration. The general behavior of cost and probability functions are discussed in Section 4.5.2.

Table 4.6: Fleet size across different iterations.

Parameter	**Iteration 1**	**Iteration 2**	**Iteration 3**
Load-out equipment sets	7	8	9
Truck-tractors	15	19	18
Total capital cost (in USD)	1,408,000	1,702,000	1,746,000

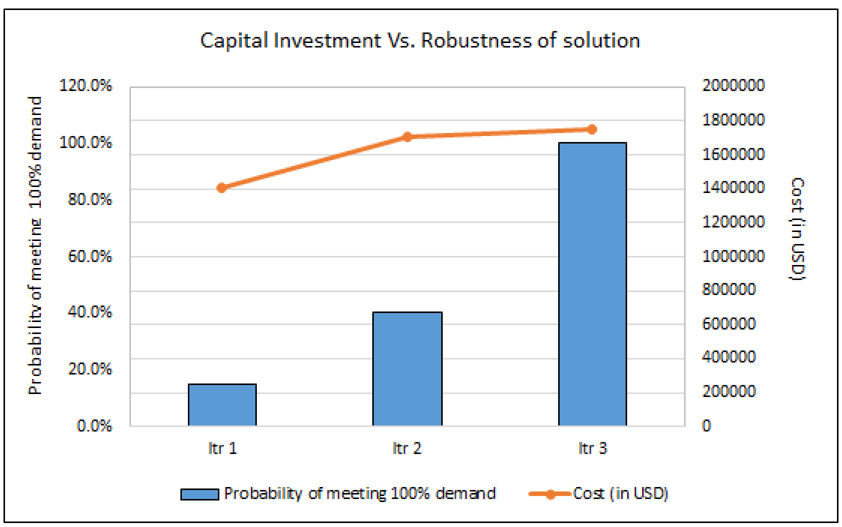


Figure 4.9: Robustness of solution with associated capital investment.

The graph in Figure 4.10, shows the performance of the biomass logistics chain over the entire planning horizon. The solution corresponding to the deterministic model with average values for the stochastic parameters met 86.7% of the annual bio-refinery demand on average across different scenarios.

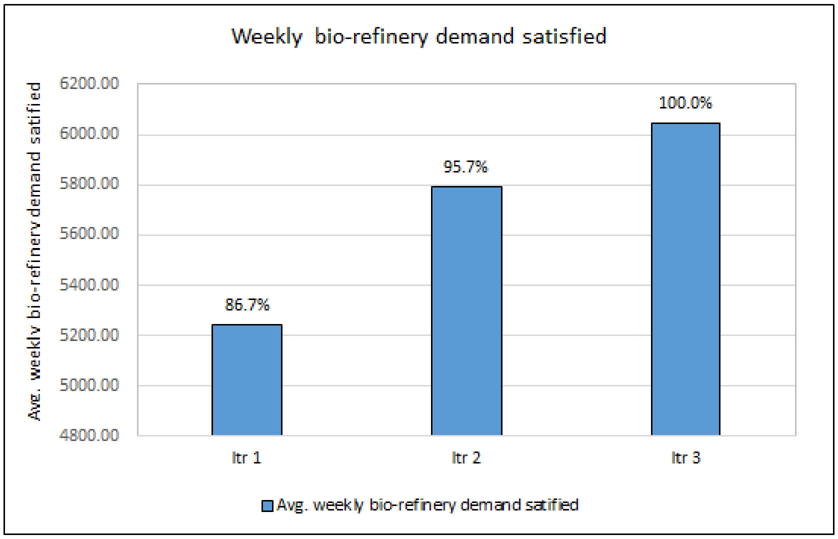


Figure 4.10: Performance of biomass logistics chain over the entire planning horizon.

The weekly performance of the biomass logistics chain under consideration for different iterations is depicted in the graph in Figure 4.11. The weekly bio-refinery demand on average across different scenarios for every week is met only in iteration 3.

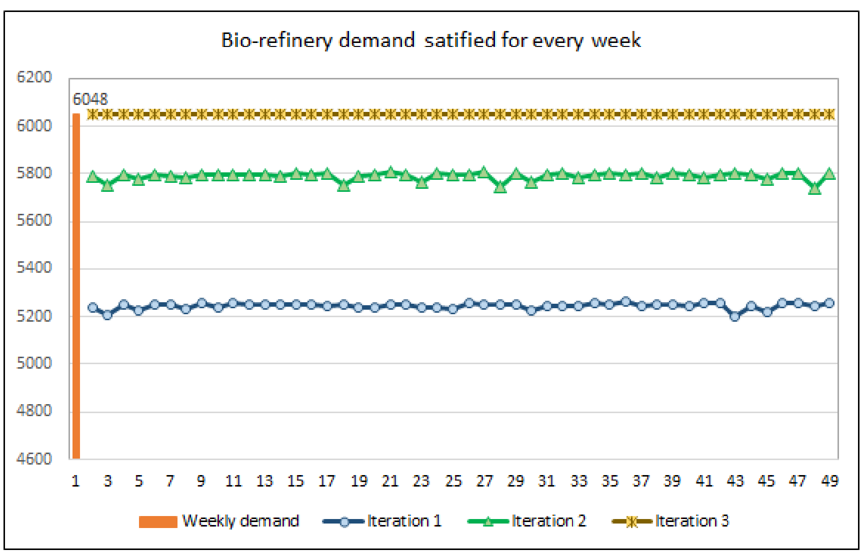


Figure 4.11: Weekly performance of the biomass logistics chain across different iterations.

We define "Productivity factor" as a ratio of observed capacity to the theoretical capacity of a unit of equipment. It is determined using the simulation model by considering stochasticity in operating parameters, breakdown, and unscheduled stops during the operation.

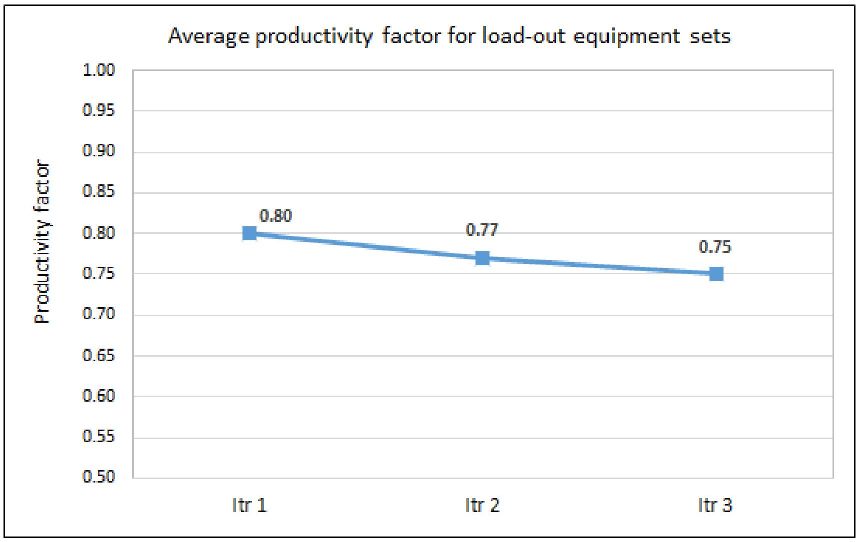


Figure 4.12: Average productivity factor for load-out equipment sets.

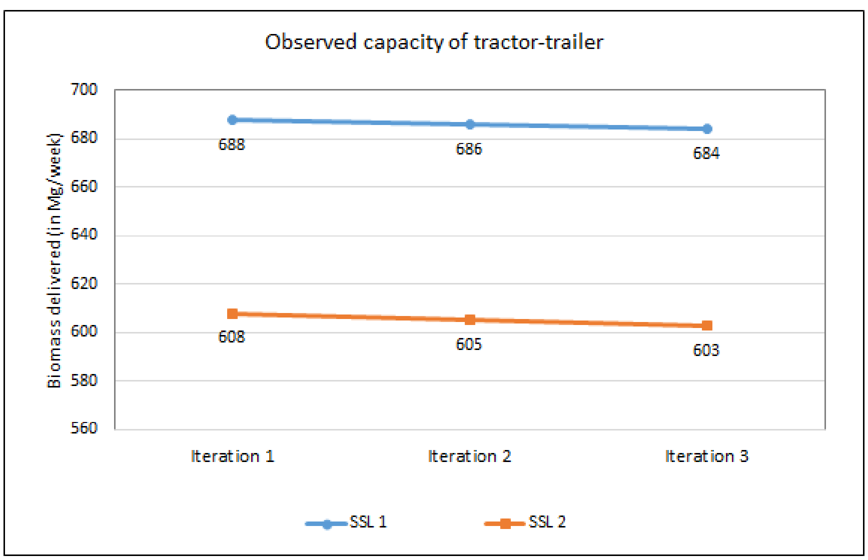


Figure 4.13: Observed capacity of tractor-trailers.

Figure 4.12 shows the productivity factor of the load-out equipment across iterations. The productivity factor of the load-out equipment in iteration 3 is found to be 0.75 in line with expectation of practitioners in the field, which is about 0.7 for a mature operation. Similarly, observed capacity of tractor-trailers are evaluated using the simulation model and is shown in Figure 4.13.

The number of bales loaded on the racks at SSL and delivered to bio-refinery depends on the loading and unloading time, speed at which load-out equipment and truck-tractor operate. While the operation starts, the loading and unloading time would be above average and will improve as the operation matures. In a matured biomass logistics operation at the SSL, the tractor-trailer can be coupled within 10 minutes per trailer and a two-bale loading cycle can be completed within 3 minutes. Table 4.7 provides fleet sizes for cases corresponding to start of operation and a matured operation.

Table 4.7: Comparison of fleet size for start of operation and a matured operation.

Parameters	**Start of operation**	**Matured operation**
Loading/unloading time for tractor-trailer (avg)	15 minutes	10 minutes
2-bale cycle time for load-out equipment (avg)	4.5 minutes	3 minutes
Fleet size of load-out equipment sets	10	7
Fleet size of tractor trailers	19	16

4.5.2 Properties of the solution obtained from the proposed methodology

Property 1. *When all the SSLs have equal inventory and are equidistant from the bio-refinery, the cost is a non-decreasing function of load-out equipment capacity U.*

Consider an example with 10 SSLs (S_1 to S_{10}) equidistant from the bio-refinery P, each with an equal inventory of 700 Mg. Let U (in Mg) be the capacity of the load-out equipment and

Q_i (in Mg), the amount of biomass a tractor-trailer can haul to bio-refinery if allotted to SSL i be 400 Mg. The bio-refinery demand is at least 2100 Mg. Figure 4.14 is a schematic representation of the example under consideration.

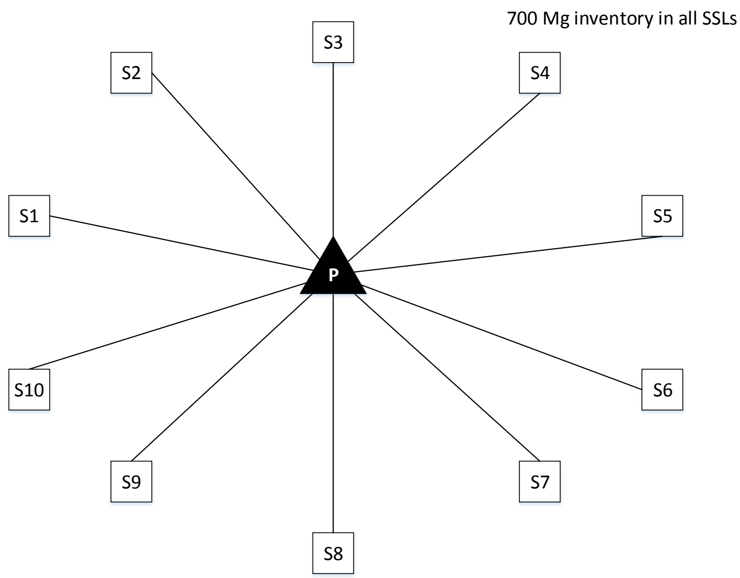


Figure 4.14: SSLs with equal inventory equidistant from bio-refinery.

Let U be varied from 700 Mg to 500 Mg in steps of 50 Mg. The number of load-out equipment sets, and tractor-trailers required, and the cost for each step is summarized in Table 4.8.

Table 4.8: Number of load-out equipment sets and tractor-trailers required for Case 1.

Load-out capacity (U)	**Load-out equipment (K)**	**Tractor-trailers (V)**	**Cost (\$)**
700	3	6	582,000
650	4	8	776,000
600	4	8	776,000
550	4	8	776,000
500	5	10	970,000

The cost function is not strictly increasing due to the discrete nature of number of load-out equipment sets (K) and tractor-trailers (V). Hence, the cost function is neither a convex or a concave function of U.

Property 2. *When the SSLs are located at unequal distances from the bio-refinery and carry unequal inventory, the total cost incurred as function of U is not a monotone function.*

Consider an example with 10 SSLs (S_1 to S_{10}). SSLs S_1 to S_5 are 2 unit distance away from the bio-refinery P and carry an inventory of 700 Mg. SSLs S_6 to S_{10} are 1 unit distance away from the bio-refinery and carry an inventory of 550 Mg. Let Q_i (in Mg), the amount of biomass a tractor-trailer can haul to bio-refinery if allotted to SSL i, be 400 Mg and 800 Mg for SSLs S_1 to S_5 and S_6 to S_{10} , respectively. The bio-refinery demand is at least 2100 Mg. Figure 4.15 is a schematic representation of this case.

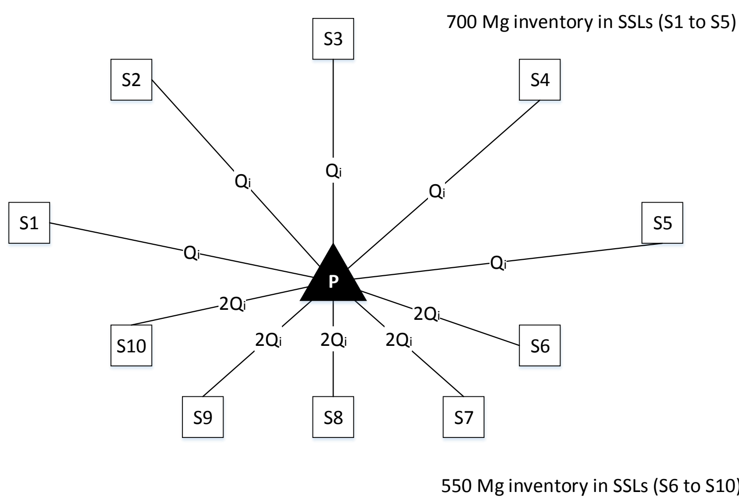


Figure 4.15: SSLs with unequal distance from the bio-refinery with unequal inventory.

Let U be varied from 700 Mg to 500 Mg in steps of 50 Mg. The number of load-out equipment sets, and tractor-trailers required, and the total cost for each step are presented in Table 4.9.

Table 4.9: Number of load-out equipment sets and tractor-trailers required for Case 2.

Load-out capacity (U)	**Load-out equipment (K)**	**Tractor-trailers (V)**	**Cost (\$)**
700	3	6	582,000
650	4	8	776,000
600	4	8	776,000
550	4	4	576,000
500	5	5	720,000

When the load-out capacity U is 700 Mg, only SSLs S_1 to S_5 can be unloaded as the allocated load-out equipment has to be fully utilized. The load-out equipment cannot be allocated to the SSLs S_6 to S_{10}, as they do not have enough inventory for the equipments to be fully utilized. When the load-out capacity U is 550 Mg, the load-out equipment can be allocated to any SSL from S_1 to S_{10}. But unloading from SSLs S_6 to S_{10} would be optimal as they are closer to the bio-refinery and require lesser number of tractor-trailers. The number of load-out equipment sets is a non-decreasing function of U, but the number of tractor-trailers is not a monotonous function of U, and it depends on the inventory in the SSLs and their spatial distribution. The cost function, which depends on the number of load-out equipment sets and tractor-trailers, can hence increase or decrease with respect to change in U.

Property 3. *The probability of meeting bio-refinery demand is monotonically non-decreasing with decrement in the value of U.*

With each iteration, the load-out capacity U monotonically decreases and the number of load-out equipment sets is monotonically non-decreasing. Since with each iteration, the capability of unloading biomass increases or remains same, the probability of meeting bio-refinery demand also must be monotonically non-decreasing. The variation of probability of meeting demand across iterations may not have decreasing slope and hence may not exhibit diminishing marginal return behavior.

4.5.3 Discussion

The methodology iteratively arrives at a robust solution by solving the optimization model in Phase 1 for different capacity parameters. The simulation model considers the stochastic parameters, its interaction in the complex biomass logistics system and evaluates the capacity parameters, which is updated in Phase 1 (optimization problem). Since, the results from the simulation model guides the updating of parameters in Phase 1, the methodology is more effective than randomly changing the capacity parameter. In the proposed modeling methodology, the biomass inflow rate is not treated as a random variable. The number of load-out equipment sets and truck tractors determined using this methodology pertains to a particular scenario of biomass inflow rate. The inflow of biomass to each SSL depends on the harvesting scenario followed and the crop yield, which in turn depend on the weather, and hence, involves uncertainty. By using details from a weather forecast, the expected crop yield for the upcoming harvest season can be estimated. Based on the estimated crop yield and harvest schedule followed by the production field owner, the biomass inflow rate can be estimated and the same modeling methodology can be utilized to adjust fleet sizes of load-out equipment and truck tractors accordingly. The operational-level uncertainties such as delay in delivery of biomass to storage facilities, break-down of equipment, among others can be handled using an effective decision support tool which uses real-time information to make decisions. Such a decision support tool is introduced in the subsequent chapter.

The model under consideration enforces complete utilization of load-out equipment allocated to an SSL for load-out operation by assigning a constraint (4.39). Relaxing this constraint and allowing partial utilization of load-out equipment can yield solutions with smaller fleet size of logistics equipment, but at the same time requires a greater computational effort to determine an optimal solution. Modifying the optimization algorithm to allow partial utilization of load-out equipment and obtain optimal solution within a reasonable computational time is a relevant extension of this work.

Another objective to optimize is the average delivery cost ($/Mg) for hauling switchgrass bales to a bio-refinery, which is a relevant single cost parameter indicating effectiveness of the farmgate and highway hauling operation in biomass feedstock logistics. The average delivery cost considers the truck cost (which includes cost of ownership, fuel cost, and truck operation cost), loading cost (which includes cost of ownership, cost of operation of the load-out equipment and service trucks supporting the load-out equipment sets), and mobilization cost (which is the cost of mobilizing load-out equipment from one SSL to another). The cost of trailers and racks are included in the truck cost. Using average delivery cost in the model would provide meaningful solutions to the user and could be another relevant extension to this work.

4.6 Conclusion

In this chapter, we have introduced the Stochastic-Biomass Logistics Chain Problem (S-BLCP), a new and challenging problem that integrates single-item capacitated lot-sizing problem, high-multiplicity multiple traveling salesman problem, and biomass logistics chain design problem under the uncertainties expected for actual logistics equipment. A two-phase optimization-simulation approach was developed to solve this problem in an iterative fashion. The methodology does not guarantee optimal solution but provides solutions that are robust to uncertainties associated with the logistics equipment in the system. The methodology is applied to a real-life problem instance to show its ability to solve large-sized problem instances. As future research work, we propose to study applicability of the modeling methodology presented in this chapter to stochastic problems in general, and also, further investigate its efficacy.

Chapter 5

An Operational-Level Decision Support Tool for Biomass Feedstock Logistics

5.1 Introduction

The biomass logistics chain comprises a complex collection of logistics activities performed in a stochastic environment. The judicious execution of these activities is imperative to ensure cost-effective biofuel production. Decision support tools (DST) are now being considered to make decisions at strategic, tactical and operational levels in the farmgate-to-bio-refinery supply chain. The operational-level decisions are short term (weekly, daily, or hourly) decisions concentrating on inventory planning, vehicle planning and scheduling to ensure continuous operation of conversion facilities and other processes in the supply chain.

In this chapter, we propose a decision support tool that relies on an optimization model to make operational-level decisions for a switchgrass-based biomass logistics system. The biomass logistics decision support tool (BL-DST) considers the current system status (i.e., inventory level, location of equipment sets, biomass inflow for the immediate time periods) to determine which storage facility to unload biomass from, amount of biomass to unload from each storage facility, and to allocate and route logistics resources such as tractor-trailers and load-out equipment sets in order to minimize the total operational cost.

We have organized the chapter as follows: In Section 5.2, we identify the need for an operational-level decision support tool for the biomass feedstock logistics design proposed in previous chapters. In Section 5.3, we describe and formulate the biomass logistics problem under consideration as a mixed integer programming model. The usefulness of the decision support tool is highlighted in Section 5.4 and concluding remarks are made in Section 5.5.

5.2 Need for an operational-level decision support tool

A decision support tool is a computer-based system, which helps in decision making as system status changes. The fundamental components of a decision support tool includes: (1) database, (2) model, and (3) a user interface as shown in Figure 5.1.

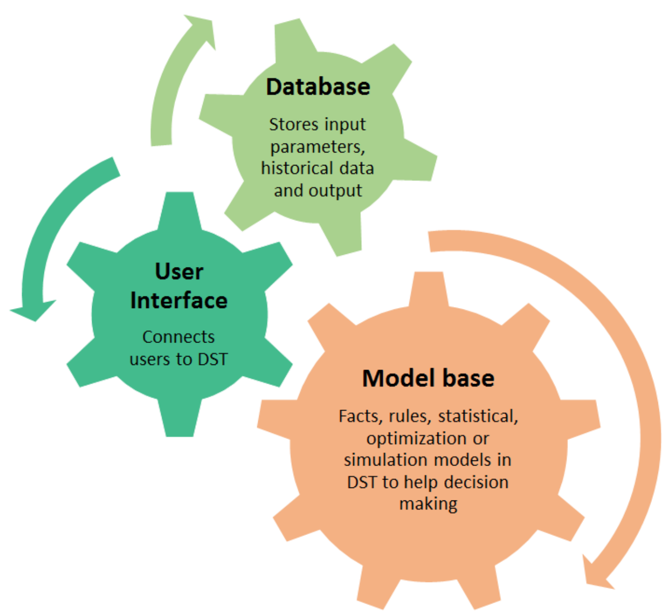


Figure 5.1: Components of a decision support tool.

The decision support tool can be classified based on the way decisions are made. A model-driven decision support tool uses statistical, optimization, or simulation-based models to aid in the decision making process. The proposed decision support tool for biomass feedstock logistics relies on solution to an optimization model.

Judd et al. [62] have presented methods to determine strategic decisions such as number of storage facilities and fleet sizes of load-out equipment and tractor-trailer. They do not take into consideration variation in biomass inflow rate at SSLs that may vary from period to period. Aguayo et al. [1] use a multi-period model to determine tactical decisions such as fleet size of load-out equipment, tractor-trailers and operational-level decisions like amount of biomass to be shipped and routing of load-out equipment sets for every time period, based on the expected inflow of biomass to each storage facility in each period. The biomass inflow to a storage facility is based on the crop yield and the harvest schedule followed by feedstock producers. Both crop yield and harvesting schedule are stochastic in nature, and hence, inflow of biomass to storage facilities cannot be accurately modeled for the entire planning horizon as assumed in Aguayo et al. [1]. Currently, the feedstock manager either use spreadsheet models or rules like shortest distance first and inventory availability to select storage facilities to pickup biomass during every time period. Optimizing routing cost across 100-200 storage facilities is not a trivial problem, and this problem can be mathematically modeled as a mixed integer linear program.

A biomass logistics decision support tool (BL-DST) that relies on solution of a mixed integer programming model is proposed to minimize operational cost. It determines operational decisions like which storage facilities to unload biomass from, amount of biomass feedstock to be unloaded from each storage facility, and allocation and routing of logistics resources such as tractor-trailers and load-out equipment sets. The optimization model considered is similar to the mixed integer programming model proposed in Aguayo et al. [1] but it includes additional constraints associated with system status like current inventory level and current location of load-out equipment.

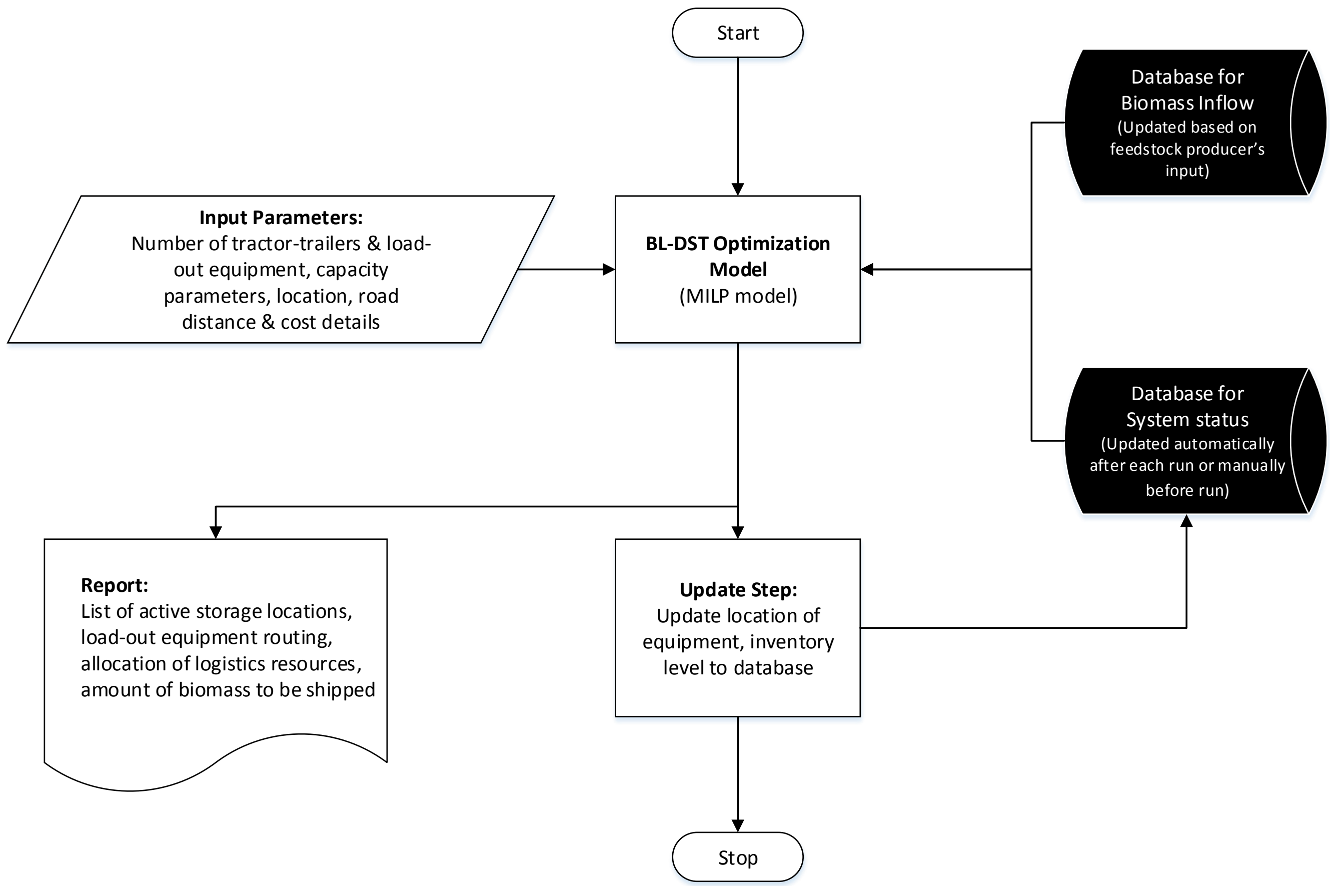


Figure 5.2: Flowchart of the biomass logistics desion support tool.

Two different databases are maintained, one for details on biomass inflow from fields to storage facilities and the other for tracking current system status as shown in Figure 5.2. The feedstock producer's input is used to update the biomass inflow database for the next two weeks. The feedstock producers should be able to provide this information confidently as they may have already harvested the feedstock that is to be delivered within the next two weeks.

5.3 Problem statement and model formulation

Problem statement

Given a set F of pre-located SSLs, inflow of biomass based on inputs from feedstock producers during each time period t at each SSL i, $\hat{A}_i^t$, planning horizon of length T, and operational parameters of logistics equipment, current inventory levels y_i^0, and location of logistics equipment, determine the amount of biomass to be shipped from SSLs to the bio-refinery, routing of logistics equipment during each time period to share them across different demand points, allocation of tractor-trailers to SSLs during each time period, so as to minimize the mobilization and shipment costs.

Formulation for the BL-DST

Next, we present a mixed integer programming model for the BL-DST. Consider the following notation.

Sets:

- F : Set of SSLs.
- F^0 : Set of SSLs including the bio-refinery, which is denoted by 0 $(0 \cup F)$.
- L : Set of SSLs where the load-out equipment sets are stationed at $t = 0$.
- T : Length of the planning horizon $(1, ..., T)$.
- T^0 : Length of the planning horizon including time 0, $(0, 1, ..., T)$.
- T^+ : Length of the planning horizon including time T+1, $(1, ..., T + 1)$.

Parameters:

- l_i : Number of load-out equipment sets stationed at SSL i at $t = 0$, $\forall i \in L$.

$\hat{c}_i$: Shipping cost from SSL i to the bio-refinery (\$ per Mg), $i \in F$.

c_{ij} : Cost to transport an equipment from SSL i to SSL j (\$ per mile), $\forall i \in F^0$, $\forall j \in F^0$.

$\hat{A}_i^t$: Amount of biomass available at SSL i during time period t (Mg), $\forall i \in F, \quad \forall t = 0, ..., T$.

N_i : Number of out-and-back trips a vehicle can perform during a time period if allocated to SSL i, $\forall i \in F$.

α : Fraction of biomass inventory lost (Storage loss).

Q : Tractor-trailer capacity (Mg).

U : Maximum load-out rate for each equipment set (Mg per time unit).

K : Number of load-out equipment sets available.

V : Number of tractor-trailers available.

P^t : Plant requirement during time period t (Mg), $\forall t = 1, ..., T$.

E : Maximum number of equipment sets to allocate to an SSL during a time period.

Decision Variables:

s_i^t : Amount of biomass shipped from SSL i to the bio-refinery during period t, $\forall i \in F, \quad \forall t = 1, ..., T$.

y_i^t : Inventory of biomass at the end of time period t at SSL i. $\forall i \in F, \quad \forall t = 0, ..., T$.

v_i^t : Number of tractor-trailers assigned to SSL i during time period t, $\forall i \in F, \quad \forall t = 1, ..., T$.

z_i^t : Number of equipment sets assigned to SSL i during time period t, $\forall i \in F, \quad \forall t = 1, ..., T$.

x_{ij}^t : Integer variable that indicates the number of equipment sets traversing from SSL i to SSL j at the end of time period t. $\forall i \in F^0$, $\forall j \in F^0$, $\forall t = 0, ..., T$.

Model BL-DST:

$$\text{Minimize} \quad \sum_{t=1}^{T} \sum_{i \in F^0} \sum_{j \in F^0} c_{ij} x_{ij}^{t-1} + \sum_{t=1}^{T} \sum_{i \in F} \hat{c}_i s_i^t \tag{5.1}$$

subject to :

$$\sum_{i \notin L} \sum_{j \in F^0} x_{ij}^0 = 0 \tag{5.2}$$

$$\sum_{j \in F^0} x_{ij}^0 \leq l_i, \quad \forall i \in L \tag{5.3}$$

$$\sum_{i \in F} s_i^t \geq P^t, \quad \forall t = 1, ..., T \tag{5.4}$$

$$y_i^t = (1-\alpha) y_i^{t-1} - s_i^t + \hat{A}_i^t, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{5.5}$$

$$y_i^0 = \hat{A}_i^0, \quad \forall i \in F \tag{5.6}$$

$$s_i^t \leq Q N_i v_i^t, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{5.7}$$

$$\sum_{i \in F} v_i^t \leq V, \quad \forall t = 1, ..., T \tag{5.8}$$

$$\sum_{i \in F} z_i^t \leq K, \quad \forall t = 1, ..., T \tag{5.9}$$

$$s_i^t \leq U z_i^t \quad , \forall i \in F, \quad \forall t = 1, ..., T \tag{5.10}$$

$$\sum_{j \in F^0} x_{ji}^{(t-1)} - z_i^t = 0, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{5.11}$$

$$\sum_{j \in F^0} x_{ji}^{(t-1)} - \sum_{j \in F^0} x_{ij}^t = 0, \quad \forall i \in F, \quad \forall t = 1, ..., T-1 \tag{5.12}$$

$$x_{i0}^t = 0, \quad \forall i \in F, \quad \forall t = 0, ..., T \tag{5.13}$$

$$s_i^t \geq 0, \quad \forall i \in F, \quad \forall t = 1, ...T \tag{5.14}$$

$$z_i^t \quad \in \{0, 1, ..., E\}, \quad \forall i \in F, \quad \forall t = 1, ..., T \tag{5.15}$$

$$y_i^t \geq 0 \ , \quad \forall i \in F, \quad \forall t = 0, ..., T \tag{5.16}$$

$$v_i^t \in \mathbb{Z}^+, \quad \forall j \in F, \quad \forall t = 1, ..., T \tag{5.17}$$

$$x_{ij}^t \in \mathbb{Z}^+, \quad \forall i \in F^0, \quad \forall j \in F^0, \quad \forall t = 0, ..., T-1 \tag{5.18}$$

The objective function (5.1) minimizes the total operating cost incurred by mobilization of equipment and shipping of biomass from SSLs to the bio-refinery. Constraints (5.2) and constraints (5.3) capture the mobilization of load-out equipment sets at time period 0. Constraints (5.4) assure that the demand of the bio-refinery during each time period t is satisfied. Constraints (5.5) capture the inventory balance during each time period. Constraints (5.6) correspond to the inventory at SSL i in time period 0. Constraints (5.7) ensure that the amount shipped from SSL i to the bio-refinery in each time period t is at most equal to the number of tractor-trailers (v_j^t) allocated to this SSL times their capacity (Q), and the number of trips (N_i). Constraints (5.8) and (5.9) capture, respectively, the maximum number of tractor-trailers and load-out equipment sets available. Constraints (5.10) assure that bales can be shipped from SSL i to the bio-refinery only if there is load-out equipment allocated to that SSL. The amount of biomass shipped is at most equal to the number of load-out equipment sets (z_j^t) allocated to this SSL times their capacity (U). Constraints (5.11) enforce mobilization of required number of equipment sets from SSL $j \in F^0$ to SSL $i \in F$ at the end of a time period t. Note that we permit an equipment set to stay at an SSL if needed. Constraints (5.12) are the standard flow conservation constraints for equipment sets. Constraints (5.13) enforce that an equipment stay at an SSL and not move to the bio-refinery. Constraints (5.14) - (5.18) define the domains of the variables.

Modification of the BL-DST

Note that the BL-DST model will be infeasible by constraints (5.4) when sufficient biomass or logistical resources are not available to meet the bio-refinery demand. A modification of BL-DST is presented by allowing shortages with a penalty cost of γ (\$ per Mg). We designate this problem as BL-DST-E1.

Model BL-DST-E1:

$$\text{Minimize} \quad \sum_{t=1}^{T}\sum_{i\in F^0}\sum_{j\in F^0} c_{ij}x_{ij}^{t-1} + \sum_{t=1}^{T}\sum_{i\in F}\hat{c}_i s_i^t + \sum_{t=1}^{T}\gamma[P^t - \sum_{i\in F} s_i^t] \tag{5.19}$$

subject to :

$$\sum_{i\in F} s_i^t \leq P^t, \quad \forall t = 1, ..., T \tag{5.20}$$

Constraints (5.2) & (5.3)

Constraints (5.5) - (5.18)

This model will minimize shortages at the bio-refinery due to non-availability of feedstock at the SSLs. If the amount of biomass shipped is short due to lack of logistical resources, options like leasing truck tractors and load-out equipment sets can be explored. Leasing option for load-out equipment is always not practical and viable especially for short durations as there is a huge cost involved in renting and transporting these equipment sets from rental locations in large cities to production fields mostly in rural areas. A more realistic option is to add a back-up load-out equipment set in addition to the fleet of equipment set that are continuously used. The back-up equipment set is typically not a brand new machine, and hence, has a lower cost of ownership. This approach is used to address a spike in demand and unexpected breakdowns in other agriculture-related industries like sugar mills, among others.

5.4 Implementation and results

The delivery of biomass feedstock from the production fields to the SSLs is governed during a specific time window by farm gate contracts. The harvesting and delivery of biomass feedstock is a highly stochastic process, and hence, deviations from a pre-planned delivery schedule is natural. The BL-DST contains an operational level model which determines load-out schedule and routing for short time periods (weekly or daily) considering the current status of the biomass logistics system. The model can be run by the feedstock manager either at fixed intervals like the end of week or whenever the value of a parameter in the system changes.

The time-expanded network in Figure 5.3 illustrates the features of the BL-DST model.

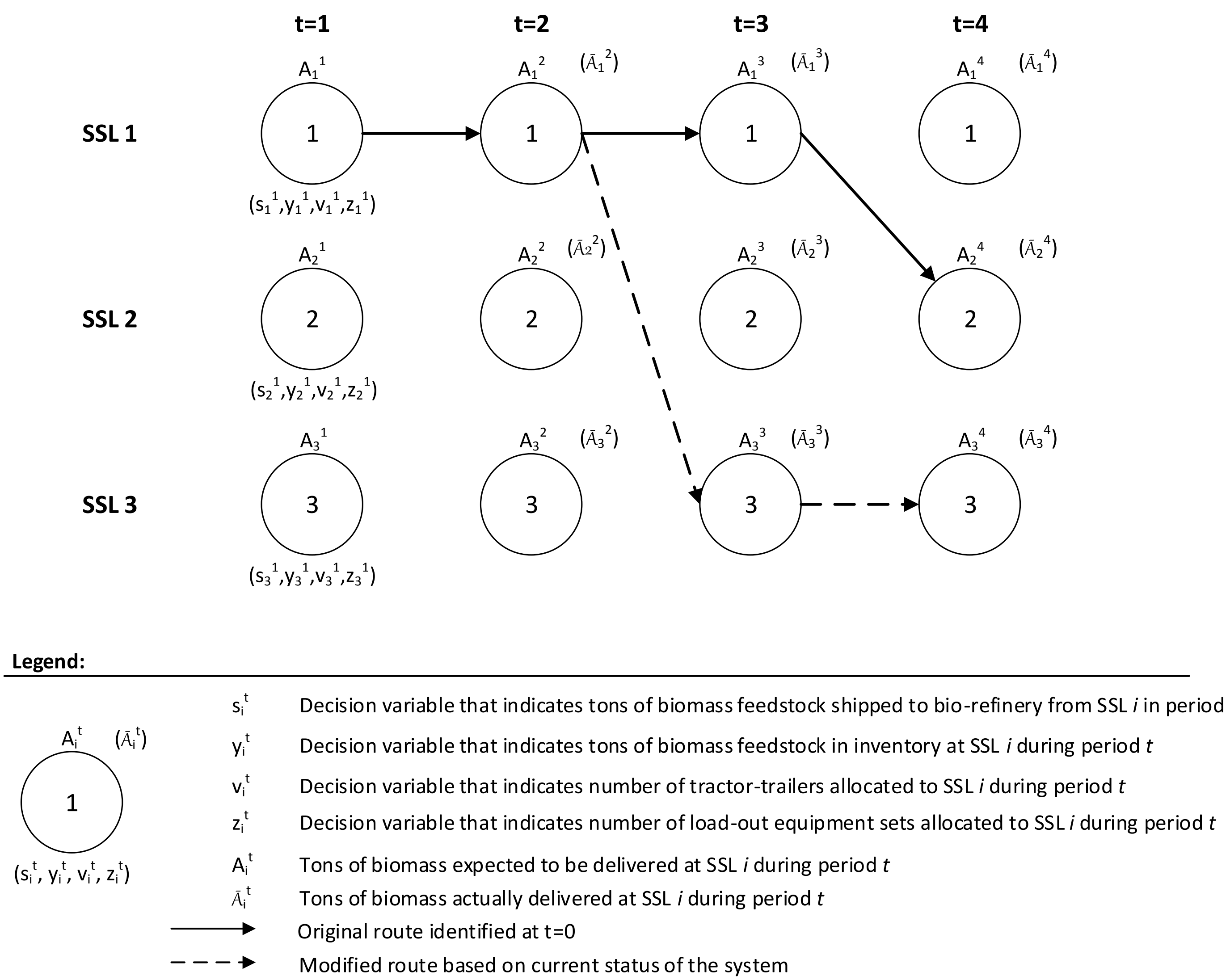


Figure 5.3: Features of the BL-DST in a time-expanded network.

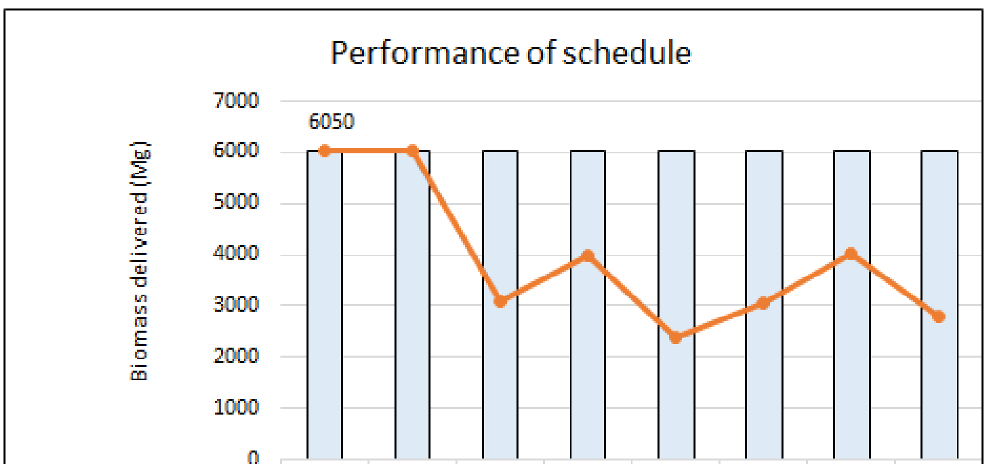
It depicts three pre-located SSLs and a planning horizon of 4 time periods. A_i^t is tons of biomass expected to be delivered at SSL i during the time period t based on longterm forecasts. $\hat{A}_i^t$ is the firm amount of biomass in tons that will be delivered to SSL i for immediate time periods t based on inputs from feedstock producers. The routing decisions and load-out schedule is modified as inflow of biomass ($\hat{A}_i^t$) changes.

Table 5.1: Other parameters and their values used in BL-DST.

Sr. No.	**Parameter**	**Values**
1	Number of SSLs	199
2	Number of trucks	23
3	Number of load-out equipment sets	10
4	Plant Requirement (weekly)	6050 Mg
5	Truck capacity per trip	16 Mg
6	Truck loading (unloading) time at SSL (bio-refinery)	12.5 Minutes
7	Truck speed on highways	45 mph
8	Telehandler capacity per week	605 Mg

Usefulness of the BL-DST model is evaluated using the dataset described in Resop et al. [96]. A total of 199 SSL(s) distributed over 48-km radius around Gretna, VA is considered. The expected inflow of biomass is calculated based on typical harvesting schedule and crop yield. The mobilization cost is defined as the cost to move load-out equipment sets from one SSL to another. The following rules were used to estimate mobilization cost: (1) If the next SSL is not more than 5 miles away, the load-out equipment will be driven to the next SSL. Operational cost of the load-out equipment is $ 39.50/h, and the travel time is based on the road speed of 17 mi/h; and (2) The cost incurred for the truck and equipment trailer to haul the load-out equipment is $ 125/h, which includes the labor cost for the truck operator. Total mobilization time consists of the time required to travel from the bio-refinery to the current SSL, time to load the equipment, travel time to the next SSL, time to unload the equipment, and the travel time back to the bio-refinery. Similarly, the biomass hauling cost (\$ per km) is calculated based on the equation given in Judd et al. [62], 0.1381 x d_j +1.6667, where d_j (km) is the distance from SSL j to bio-refinery. All the other parameters considered for this data set are summarized in Table 5.1.

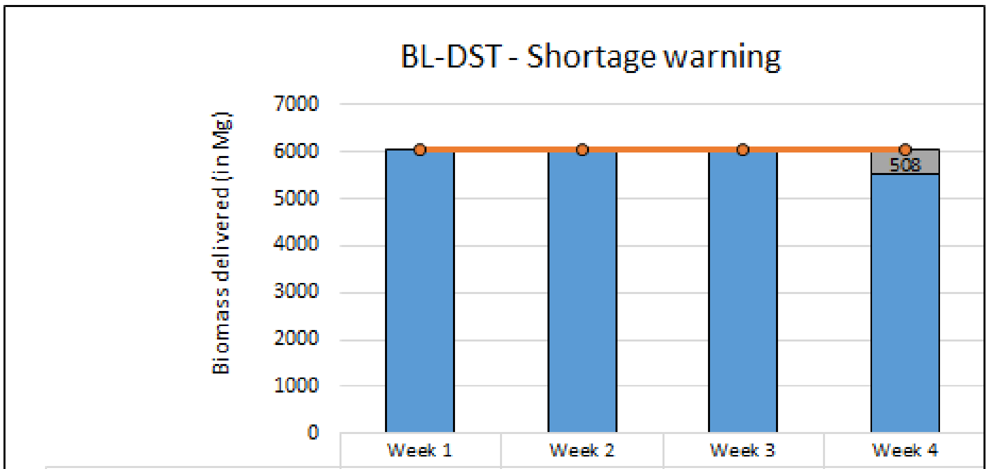
Figure 5.4 compares performances of different load-out and routing schedules. We compare the biomass delivered to the bio-refinery for the original schedule determined based on the expected inflow rate of biomass and the schedule from the biomass logistics decision support tool for a period of 8 weeks. An initial schedule is developed for a period of 2 months considering the expected biomass inflow into the SSLs. The actual biomass inflow every week may vary and the initial schedule may no longer satisfy the bio-refinery demand. The database of the BL-DST is updated with firm amount of biomass that can be delivered for next two weeks, and this is used to develop load-out schedule and routes for the equipment sets. The feedstock producers are expected to contact the feedstock manager's office to update exact details of biomass feedstock delivery to the SSLs. By re-adjusting the schedule every week or once in two weeks, the BL-DST schedules ensure that the bio-refinery demand is satisfied every week.



	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
BL-DST schedule	6050	6050	6050	6050	6050	6050	6050	6050
Original schedule	6050	6050	3080	3977	2389	3041	4019	2786

Figure 5.4: Biomass deilvered to bio-refinery for orginial and BL-DST schedules.

The BL-DST-E1 model can warn the feedstock manager of a potential shortage in the coming weeks, to help his team expedite biomass delivery to SSLs or to make alternate arrangements. The incoming inventory of the dataset with 199 SSLs, is modified to deliberately cause shortages to show this feature of the decision support tool. The planning horizon considered is four weeks. Figure 5.5 shows the possibility of a shortage of 508 Mg of biomass in week 4. The feedstock manager can now asses risk and try to expedite delivery of biomass from production fields to storage locations to mitigate risk of not meeting bio-refinery demand.



	Week 1	Week 2	Week 3	Week 4
Shortage (in Mg)	0	0	0	508
Biomass delivered (in Mg)	6050	6050	6050	5542
Demand (in Mg)	6050	6050	6050	6050

Figure 5.5: Biomass deilvered to bio-refinery for orginial and BL-DST schedules.

A sample of output report of BL-DST is shown in Figure 5.6. The output report gives the cost details, load-out schedule, allocation of load-out equipment sets and trucks to the SSL, and routing details.

Output Statistics Report
****

Total cost : $ 81707
Shipping cost : $ 79860
Mobilization cost : $ 1847

Load-out Schedule
****

Time Period	SSL#	Shipping Quantity
1	9	605
1	53	605
1	54	605
1	87	605
1	117	1210
1	118	1210
1	183	1210
2	6	605
2	52	605
2	95	1210
2	118	605
2	126	605
2	132	1210
2	181	1210

Equipment Allocation
****

Time Period	SSL#	# of Equipments
1	9	1
1	53	1
1	54	1
1	87	1
1	117	2
1	118	2
1	183	2
2	6	1
2	52	1
2	95	2
2	118	1
2	126	1
2	132	2
2	181	2

Truck Allocation
****

Time Period	SSL#	# of Trucks
1	9	3
1	53	2
1	54	3
1	87	3
1	117	4
1	118	4
1	183	4
2	6	3
2	52	2
2	95	5
2	118	2
2	126	3
2	132	5
2	181	3

Equipment Routing
****

Time Period	Initial_Location	Final_Location	#_of_Equipments
0	0	9	1
0	0	53	1
0	0	54	1
0	0	87	1
0	0	117	2
0	0	118	2
0	0	183	2
1	9	6	1
1	53	52	1
1	54	95	1
1	87	95	1
1	117	118	1
1	117	126	1
1	118	132	2
1	183	181	2

Figure 5.6: A sample output report for BL-DST.

5.5 Conclusion

In this chapter, the need for a decision support tool for managing operational-level decisions in a biomass feedstock logistics chain is emphasized. A mixed integer program-based decision support tool (BL-DST) is developed in Microsoft Visual Studio 2012, using CPLEX 12.6 solver. A dataset for a hypothetical bio-refinery in Gretna, VA with SSLs distributed in a 48-km radius around the bio-refinery was considered for implementation and to showcase its usefulness. The long-term schedules based on forecasting models or expected values become invalid as the system deviates from the expected behavior. The BL-DST considers changes in system status to adjust load-out schedule and routing to minimize operational cost, which includes cost to mobilize load-out equipment sets and shipping of biomass. As future work, a graphical user interface can be attempted to communicate output of the model to the users in a lucid fashion. Furthermore, features like prioritizing the unloading of specific SSLs can be included to address additional features encountered in practice.

Chapter 6

Bale Collection Problem

6.1 Introduction

The need for big data analytics and visualization is being recognized by the agricultural sector. These practices are being used to improve efficiency by developing sustainable and data-driven practices such as precision agriculture. The cost-effective growing, harvesting, transporting and processing of energy crops is needed to make biofuel a viable option. Data-driven approaches and analytics have increasingly become popular for biomass feedstock logistics decision making at strategic, tactical and operational levels. The biomass feedstock logistics related costs account for 35%-60% of biofuel retail price (Fales et al. [46]).

Bale collection is an operation in which bales are collected and transported from a field to a storage location situated for ready access by highway-hauling trucks. Determining a sequence for collecting these bales that minimizes the operating time, or cost, is termed as the *Bale Collection Problem* (BCP). Though, some argue that an experienced operator does a good job in judging the sequence in which to pick the bales in order to minimize the travel time between bales, the problem is not trivial when multiple capacitated storage depots or multiple vehicles or both are involved. The plan calls or the bale location to be recorded when it falls from the baler. The resulting location map is then input to the in-field hauling vehicle and route is computed. The routes generated could support round-the-clock operation of the in-field hauling vehicle and thereby improve utilization of this vehicle.

The chapter is organized as follows: In Section 6.2, we provide a brief description of the bale collection problem. In Section 6.3, we formally define the problem under consideration and review literature related to the bale collection problem and capacitated vehicle routing problem. In section 6.4, we introduce a mathematical formulation for the BCP and its extensions. Sections 6.5 and 6.6 provide details of solution methodologies explored for the BCP and some preliminary results respectively. Finally, concluding remarks are made in Section 6.7.

6.2 System description

The ownership and management of equipment for farmgate operations, which include cutting, baling and storage of biomass feedstock, play an active role in achieving cost-effective solutions. Two models widely used in agriculture, that could be emulated in farmgate operations, are: (1) the feedstock producer operates all the equipment for harvesting and storage (active owner), and (2) a custom harvest company is contracted to harvest and store the biomass (passive owner). A major drawback of having each feedstock producer operate his/her own equipment is that the majority of the required equipment is very expensive and must be operated over a large acreage to ensure competitive cost (\$/Mg). The passive ownership model, where the feedstock producer only grows the crop and does not own or operate equipment is a more viable option for a small farmer. A harvesting contractor harvests many production fields in the schedule that are ready for harvest. This model provides better resource utilization as the equipment is shared among many production fields. The fields are first cut, then baled, and finally the bales are transported to a farm-side storage location. The process of collecting and moving biomass bales from different locations on the field to a farm-side storage location is called bale collection, or in-field hauling. The bales are collected using a specialized vehicle that is capable of lifting and locating bales on a flatbed trailer. The equipment has a pick up arm for self loading and a pusher arm to place the collected bales on the trailer. Figure 6.1 shows one in-field hauling option, a tractor with a specially designed wagon for collecting bales from a production field. The capacity of this equipment set can vary from six to twelve bales depending on size of the trailer. The bale collection equipment moves the bales from field to SSL where they are placed in ambient storage as a single layer. The in-field hauling equipment is designed to be operated in the field as well as on paved roads. This equipment has a limited road speed making it less efficient than a tractor-trailer truck for highway hauling. Hence, only short distance hauling to a storage site is done with the in-field hauling equipment. The SSL is a transfer point in the logistics chain to uncouple in-field hauling and highway hauling.

Figure 6.1: Bale collection. (*Source: Antares Group*)

The bale collection problem is relevant in both active and passive ownership models. In a passive ownership model, a given set of equipment, which the contractor owns, is routed to different production fields to complete the farmgate operations within a time window. Hence, the cost and duration of operation are very important metrics for these contractors.

The bales are assumed to be uniformly distributed across the field in this study as shown in the Figure 6.2. However, the real distribution of bales need not be uniform throughout the production field due to the variability in crop yield, which in-turn depends on the interactions between soil, plant and atmosphere.

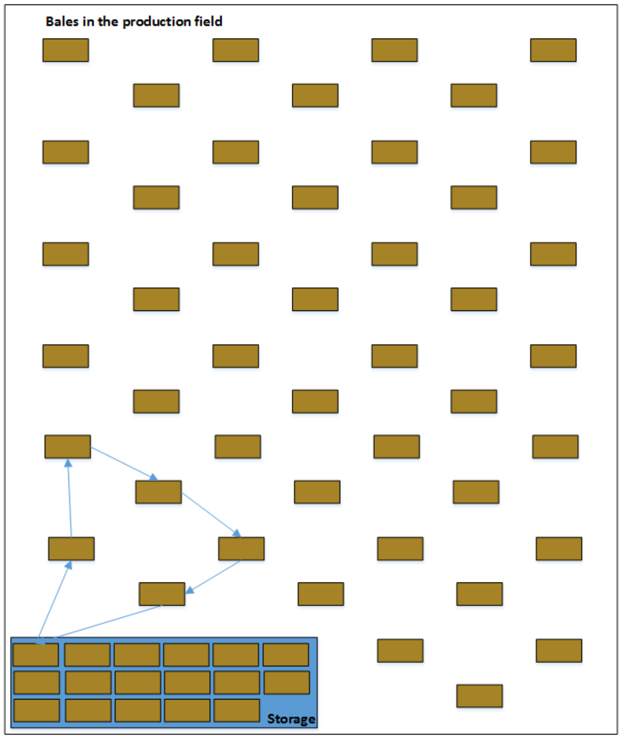


Figure 6.2: An ariel depiction of bale distribution and field-side storage in a production field.

6.3 Problem Statement and literature review

The BCP can be concisely defined as follows: Given a set, N, of bale locations in a production field and a set D, of farm-side storage depots, a non-negative cost $C_{i,j}$ associated with traveling from locations i to j, capacity c of the in-field hauling vehicle, determine an optimal sequence to collect bales from the production field in order to minimize total cost incurred. We assume $C_{i,j} = C_{j,i}$ making this problem a symmetric capacitated vehicle routing problem with unit demand (SCVRP-UD). A vehicle routing problem (VRP) is concerned with determining optimal routes to visit a set of customers by a fleet of vehicles that are based at one or more depots. The capacitated vehicle routing problem with unit demand (CVRP-UD) is a special case of the basic VRP in which there exists a limit on the maximum number of customers that can be visited in each trip. Figure 6.3 depicts features of the bale collection problem. The field contains nine bales and a field-side storage (depot). The capacity of the hauling wagon, c is 3, and hence, a total of three trips are necessary to collect all nine bales from the field.

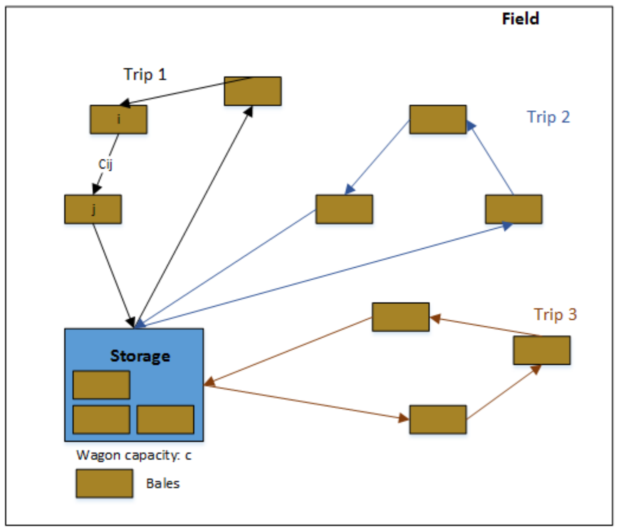


Figure 6.3: Features of the bale collection problem.

Literature review

The vehicle routing problem (VRP) has been extensively studied in the literature through many exact algorithms and heuristics approaches. The VRP is a NP-hard problem as it includes the traveling salesman problem (TSP) as a special case. It is considered to be more difficult to solve than a TSP of same size. The TSP involving hundreds or even thousands of nodes can be solved using advanced branch-and-cut algorithms whereas for VRP, sophisticated exact algorithms can only solve instances of up to about 100 nodes. Though VRP is a difficult problem to solve, it is widely studied due to its applicability to many real-world problems. The capacitated vehicle routing problem (CVRP) is a special case of VRP in which the vehicles have a capacity which cannot be violated. The objective is to visit all the customers while minimizing the total routing cost without violating capacity constraints. Laporte and Nobert [73] presents an extensive survey on exact methods for CVRP. Other literature surveys on exact methods for CVRP include Magnanti [82], Laporte [70], and Baldacci et al. [13]. Laporte [72] summarizes known results for CVRP for solution methodologies including exact algorithms, classical heuristics, and meta heuristics.

The exact methods for CVRP can be classified mainly into the following categories: branch-and-cut, dynamic programming, and set-partitioning methods. Laporte et al. [74] present an exact method for CVRP based on a cutting plane approach. Augerat et al. [11] propose a branch-and-cut algorithm which includes a class of valid inequalities, such as comb and extended comb inequalities, generalized capacity constraints, and hypotour inequalities. The branch-and-cut methodology proposed in Augerat et al. [11] could solve problem instances with up to 135 customers. Lysgaard et al. [80] propose another branch-and-cut-based approach for CVRP. Dynamic programming has been used to solve CVRP or to obtain tight lower bounds. Christofides et al. [26] present three formulations for CVRP and introduce the state space relaxation method for relaxing the dynamic programming recursions to obtain valid lower bounds. Set-partitioning-based methodologies have been studied by Balinski and Quandt [15], Hadjiconstantinou et al. [56], and Baldacci et al. [14]. Araque et al. [9] study a special case of CVRP with unit demand (CVRP-UD) by introducing several facet defining inequalities. CVRP-UD constitutes an inherent feature of the BCP.

To the best of our knowledge only Cundiff et al. [28] and Gracia et al. [51] have studied the bale collecting problem. Cundiff et al. [28] introduced the problem with a simple instance in which 34 bales scattered over the field are collected with a vehicle of capacity 6. Gracia et al. [51] propose a hybrid genetic algorithm-based approach to solve a problem instance consisting of 200 bales that are uniformly distributed over a production field. The objective is to minimize total cost. We introduce extensions to the basic bale collection problem which incorporates features encountered in practice.

6.4 Model formulation

We formulate two versions of the BCP as mixed integer programming models based on the following objective functions:

1. minimize cost of operation (BCP-C).
2. minimize total time duration when multiple vehicles are used (BCP-T).

Consider the following notation.

<u>Sets:</u>

N : Set of locations containing bales.

N^0 : Set of locations including the storage location, which is denoted by 0 ($0 \cup N$).

V : Set of vehicles $(1, ..., V)$.

<u>Parameters:</u>

t_{ij} : Travel time from location i to location j, $\forall i \in N^0$, $\forall j \in N^0$.

C_{ij} : Cost incurred for traveling from location i to location j, $\forall i \in N^0$, $\forall j \in N^0$.

tl : Loading time at each location i, $\forall i \in N^0$.

tu : Unloading time at depot.

V : Number of vehicles available for bale collection operation.

c : Capacity of vehicle (Number of bales it can carry on each trip).

<u>Decision Variables:</u>

x_{ij}^v : Binary variable indicating travel from location i to location j, $\forall i \in N^0$, $\forall j \in N^0$, $\forall v = 1, ..., V$.

y_{ij}^v : Flow variable indicating flow between location i to location j, $\forall i \in N^0$, $\forall j \in N^0$, $\forall v = 1, ..., V$.

T : Total time duration across all vehicles.

Model BCP-C:

$$\text{Minimize} \quad \sum_{i \in N^0} \sum_{j \in N^0} C_{ij} x_{ij} \tag{6.1}$$

subject to :

$$\sum_{i \in N} x_{0i} \geq \lceil N/c \rceil \tag{6.2}$$

$$\sum_{i \in N} x_{i0} \geq \lceil N/c \rceil \tag{6.3}$$

$$\sum_{j \in N^0} x_{ij} = 1, \quad \forall i \in N \tag{6.4}$$

$$\sum_{j \in N^0} x_{ji} = 1, \quad \forall i \in N \tag{6.5}$$

$$x_{ij} = 0, \quad \forall i \in N^0, i = j \tag{6.6}$$

$$y_{ij} \leq c\, x_{ij}, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.7}$$

$$\sum_{i \in N^0} y_{ij} - \sum_{i \in N^0} y_{ji} = 1, \quad \forall j \in N \tag{6.8}$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.9}$$

$$y_{ij} \geq 0, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.10}$$

The BCP-C can be modeled without keeping track of the vehicle by using a two index formulation as presented above. The objective function (6.1) minimizes the total cost of operations. Constraints (6.2) & (6.3) provides a lower bound on the number of trips required to visit all locations. Constraints (6.4) and (6.5) ensure that all locations are visited exactly once. Constraints (6.6) ensure that a self-loop by a vehicle at a location is not permitted. Constraints (6.7) guarantee that the vehicle capacity is not exceeded. Constraints (6.8) are flow conservation equations and also eliminate sub-tours. Constraints (6.9) and (6.10) define the domains of the variables.

Two relevant extensions of BCP-C are as follows:

1. Heterogeneous fleet with different capacities and cost (BCP-C-HF).

2. Multiple depots with a limited storage capacity for each depot (BCP-C-MD).

Next, we present formulations for both of these extensions of BCP-C.

Model BCP-C-HF:

$$\text{Minimize} \quad \sum_{v \in V} \sum_{i \in N^0} \sum_{j \in N^0} C_{ij}^v x_{ij}^v \tag{6.11}$$

subject to :

$$\sum_{v \in V} \sum_{i \in N} x_{0i}^v \geq \lceil N/c_{max} \rceil \tag{6.12}$$

$$\sum_{v \in V} \sum_{j \in N^0} x_{ij}^v = 1, \quad \forall i \in N \tag{6.13}$$

$$\sum_{v \in V} \sum_{j \in N^0} x_{ji}^v = 1, \quad \forall i \in N \tag{6.14}$$

$$x_{ij}^v = 0, \quad \forall i \in N^0, i = j, \quad \forall v \in V \tag{6.15}$$

$$y_{ij}^v \leq c^v \ x_{ij}^v, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v \in V \tag{6.16}$$

$$\sum_{v \in V} \sum_{i \in N^0} y_{ij}^v - \sum_{v \in V} \sum_{i \in N^0} y_{ji}^v = 1, \quad \forall j \in N \tag{6.17}$$

$$x_{ij}^v \in \{0, 1\}, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v \in V \tag{6.18}$$

$$y_{ij}^v \geq 0, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v \in V \tag{6.19}$$

BCP-HF features a heterogeneous fleet of vehicle with different capacities and cost. Constraints (6.12) provide a lower bound on the total number of trips required to visit all locations by considering vehicle with maximum capacity. Constraints (6.16) guarantee that the capacity of specific vehicle v is not exceeded. All the other constraints are the same as those for the BCP-C.

Consider the following additional notations that is used in the model for BCP-C-MD.

D	:	Set of storage depots.
N^D	:	Set of locations including the storage locations, which is denoted by set D ($D \cup N$).
q_d	:	Maximum number of full truck loads that can be accommodated at depot d.

Model BCP-C-MD:

$$\text{Minimize} \quad \sum_{i \in N^D} \sum_{j \in N^D} C_{ij} x_{ij} \tag{6.20}$$

subject to :

$$\sum_{i \in D} \sum_{j \in N} x_{ij} \geq \lceil N/c \rceil \tag{6.21}$$

$$\sum_{j \in N^D} x_{ij} = 1, \quad \forall i \in N \tag{6.22}$$

$$\sum_{j \in N^D} x_{ji} = 1, \quad \forall i \in N \tag{6.23}$$

$$x_{ij} = 0, \quad \forall i \in N^D, i = j \tag{6.24}$$

$$\sum_{i \in N} x_{id} \leq q_d, \quad \forall d \in D \tag{6.25}$$

$$y_{ij} \leq c \; x_{ij}, \quad \forall i \in N^D, \quad \forall j \in N^D \tag{6.26}$$

$$\sum_{i \in N^D} y_{ij} - \sum_{i \in N^D} y_{ji} = 1, \quad \forall j \in N \tag{6.27}$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in N^D, \quad \forall j \in N^D \tag{6.28}$$

$$y_{ij} \geq 0, \quad \forall i \in N^D, \quad \forall j \in N^D \tag{6.29}$$

BCP-C-MD considers a homogeneous fleet of vehicles. Constraints (6.25) ensure that capacity of storage depot is not exceeded. All the other constraints are the same as those for the

BCP-C.

A formulation for the minimization of total time required to pick up all the bales is as follows:

Model BCP-T:

$$\text{Minimize} \;\; T \tag{6.30}$$

subject to :

$$\sum_{v \in V} \sum_{j \in N^0} x_{ij}^v = 1, \quad \forall i \in N \tag{6.31}$$

$$\sum_{v \in V} \sum_{j \in N^0} x_{ji}^v = 1, \quad \forall i \in N \tag{6.32}$$

$$x_{ij}^v = 0, \quad \forall i \in N^0, i = j, \quad \forall v = 1, ..., V \tag{6.33}$$

$$y_{ij}^v \leq c^v \; x_{ij}^v, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v \in V \tag{6.34}$$

$$\sum_{v \in V} \sum_{i \in N^0} y_{ij}^v - \sum_{v \in V} \sum_{i \in N^0} y_{ji}^v = 1, \quad \forall j \in N \tag{6.35}$$

$$T \geq \sum_{i \in N^0} \sum_{j \in N} (t_{ij} + tl) x_{ij}^v + \sum_{k \in N} (t_{k0} + tu) x_{k0}^v, \quad \forall v = 1, ..., V \tag{6.36}$$

$$x_{ij}^v \in \{0, 1\}, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v = 1, ..., V \tag{6.37}$$

$$y_{ij}^v \geq 0, \quad \forall i \in N^0, \quad \forall j \in N^0, \quad \forall v = 1, ..., V \tag{6.38}$$

$$T \geq 0 \tag{6.39}$$

The objective function (6.30) minimizes the maximum total operating duration across all the vehicles. Constraints (6.36) capture the operating duration across all the vehicles. All the other constraints are the same as those for the BCP-C. Next, we present the solution methodologies explored for the BCP-C.

6.5 Methodolodgies explored

The solution methodologies explored for the BCP-C include direct application of CPLEX to the above formulation, column generation, and a cutting plane method where we iteratively refine the feasible region by adding linear inequalities or cuts from integer solutions.

6.5.1 Column generation

The column generation approach does not consider all the variables explicitly, and instead considers only those variables which have the potential of improving the objective function value. This approach is useful for large-scale models. The original problem is decomposed into a master problem and a sub-problem as follows. The master problem is the original problem with only a subset of variables being considered, whereas the sub-problem, which is also called the pricing problem, identifies variables that can improve the objective function of the master problem. First the master problem is solved and the values of dual variables associated with the constraints of the master problem are determined and used in formulating the objective function of the pricing problem. For a minimization problem, when the sub-problem yields a negative objective function value, there exists a variable with negative reduced cost that can improve the original objective function in the master problem. This variable is added to the master problem, and it is solved again. This process is iteratively carried out until no negative reduced cost variables are identified in the sub-problem, thereby indicating that the solution in the master problem is optimal.

Consider the following notation.

Decision Variables:

R : Set of feasible routes.

C_r : Cost of route $r \in R$.

a_{ir} : $a_{ir} = 1$ if node i is visited in route r, $a_{ir} = 0$, otherwise.

λ_r : Binary variable indicating if route r is utilized.

Master Problem:

$$\text{Minimize} \quad \sum_{r \in R} C_r \lambda_r$$

subject to :

$$\sum_{r \in R} a_{ir} \lambda_r \geq 1, \quad \forall i \in N \qquad (\pi_i) \tag{6.40}$$

$$\sum_{r \in R} \lambda_r \geq \left\lceil \frac{N}{c} \right\rceil \qquad (\beta) \tag{6.41}$$

$$\lambda_r \in \{0, 1\}, \quad \forall r \in R \tag{6.42}$$

Sub-problem:

$$\text{Minimize} \sum_{i \in N^0} \sum_{j \in N^0} (C_{ij} - \pi_i) x_{ij} - \beta \tag{6.43}$$

$$\sum_{j \in N} x_{1j} = 1 \tag{6.44}$$

$$\sum_{j \in N} x_{j1} = 1 \tag{6.45}$$

$$y_{ij} \leq c x_{ij}, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.46}$$

$$\sum_{i \in N^0} y_{ij} - \sum_{i \in N^0} y_{ji} = \sum_{i \in N^0} x_{ij}, \quad \forall j \in N \tag{6.47}$$

$$\sum_{i \in N^0} x_{ij} - \sum_{i \in N^0} x_{ji} = 0, \quad \forall j \in N \tag{6.48}$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.49}$$

$$y_{ij} \geq 0, \quad \forall i \in N^0, \quad \forall j \in N^0 \tag{6.50}$$

We solve the master problem by relaxing the integer constraints (6.42) to obtain a dual solution for use in the sub-problem. The values of a_{ir} for a route r are determined from the corresponding x_{ij} values obtained by solving the sub-problem during an iteration. After the relaxed master problem is solved to optimality, all the master problem variables are used to re-solve the master problem with integer constraints. The solution thus obtained need not be an optimal solution.

6.5.2 Cutting plane method

In this approach, we iteratively refine the feasible region by adding linear inequalities called cuts generated from integer solutions. The capacity and sub-tour elimination constraints (6.7) and (6.8) of BCP-C are relaxed and the resulting optimization problem is solved to integer optimality. All violated capacity and sub-tour elimination constraints are identified and added before re-solving the problem with these additional constraints. The process is repeated until a feasible solution to the original problem, i.e., a solution without any sub-tour or capacity violation is obtained.

The model under consideration is as follows:

Model BCP-CPM:

Minimize (6.1)

subject to :

Constraints (6.2) - (6.6)

Constraints (6.9) - (6.10)

The violation of capacity and sub-tour constraints are identified by evaluating constraints (6.51) for the integer optimal solution of BCP-CPM. The violated constraints are then added and the model is solved again.

$$\sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S| - \left\lceil \frac{|S|}{c} \right\rceil \tag{6.51}$$

where $S \subseteq N$

Constraints (6.51) is said to be facet defining inequalities to the capacitated vehicle routing problem with identical demand in Araque et al. [9].

6.6 Results and discussion

The design and operation of the biomass feedstock logistics system largely depends on the crop and physiographic region of the operation. A typical feedstock production field in the Southeastern United States is of an average size 10 ha with a yield of 6.7 Mg/ha. Though, biomass is generally baled into both rectangular and round bales, the latter is found to be more advantages for the operation in the Southeast. Round switchgrass bales of size 1.5 m diameter x 1.2 m long, each weighing about 0.4 Mg are commonly used. Hence a typical production field would have about 160 round bales. The distribution of bales in the production field depends on the baler and variation of yield across the production field. From practice, it is safe to assume that the distribution of bales are more or less uniform across the production field. The capacity of infield hauling equipment varies from 6 to 12 bales per trip, depending on the size of trailer or wagon attached to the in-field hauling vehicle.

The bale collection problem can be mathematically modeled as a capacitated vehicle routing problem (CVRP) with unit demand, but standard benchmark instances for the CVRP or the traveling salesman problem (TSP) cannot be used as the nodes are not uniformly distributed. Figure 6.4 shows the scatter plot of a standard benchmark instance with 48 nodes used to test CVRP or TSP related algorithms .For BCP, we generated instances by assuming a yield of 16-17 bales per hectare for a field with bales uniformly distributed over the production field. A problem instance is defined by the capacity constraint of the vehicle (c) and by the $n+1$ exact locations in a field: n corresponding to the location of bales to be collected, and 1 corresponding to the SSL (starting/ending point of the tour). Figure 6.5 shows the scatter plot of the data instance with 48 nodes specifically designed for the bale collection problem.

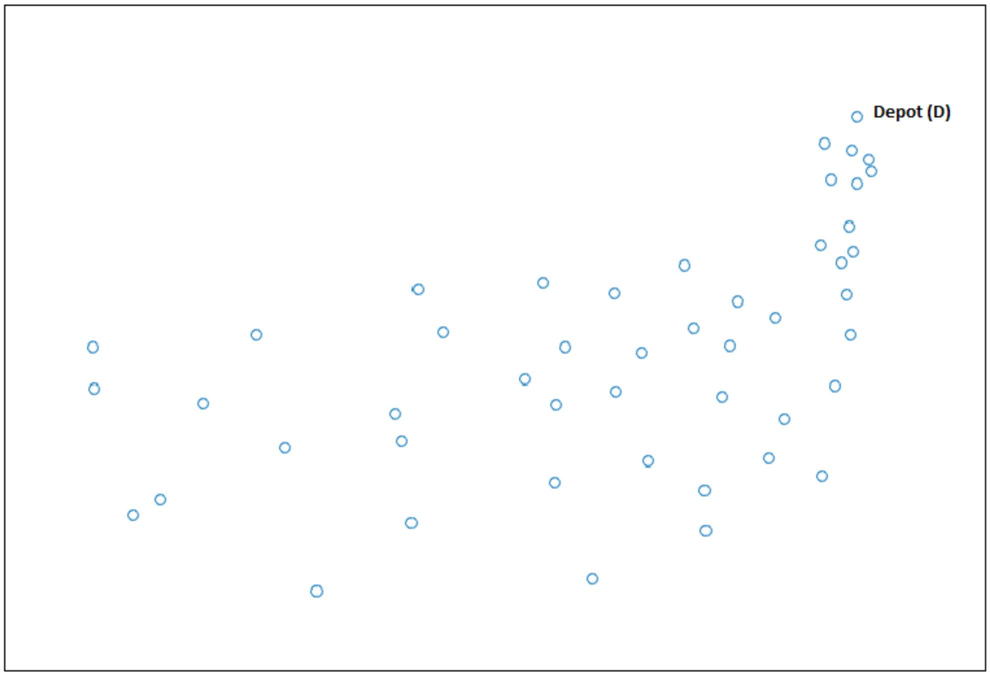


Figure 6.4: Standard benchmark data set with 48 nodes.

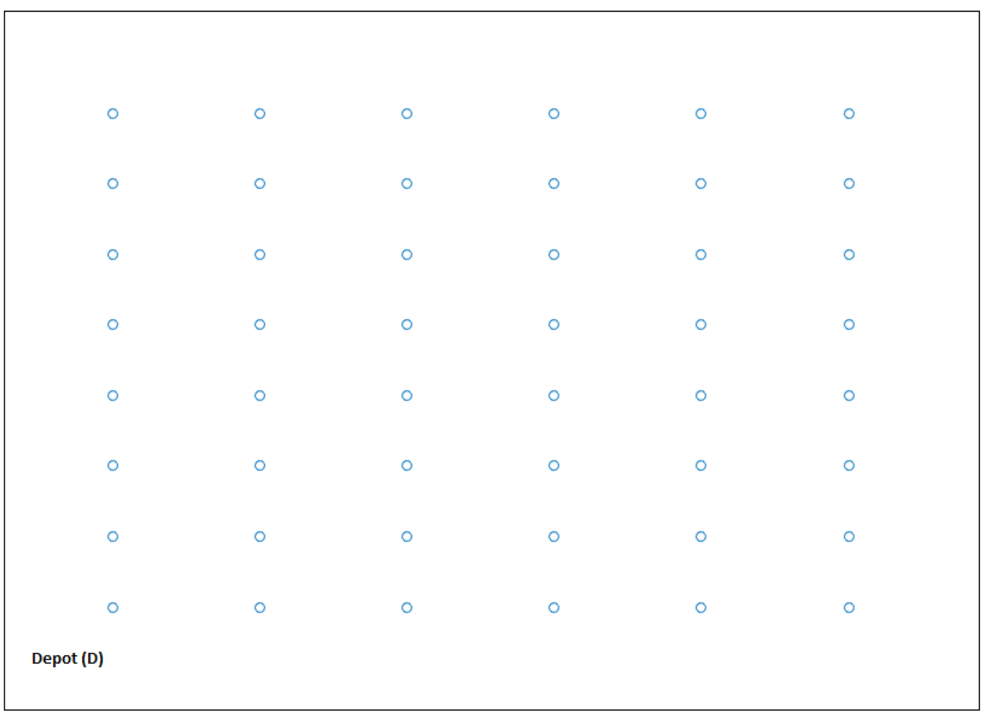


Figure 6.5: BCP data set with 48 nodes.

Table 6.1 provides some problem instances specifically for the bale collection problem, which are used for testing some of the solution approaches under study.

Table 6.1: BCP problem instances.

Instance	**Total number of nodes** (N^0)	**Number of bales** (N)	**Vehicle capacity** (c)
1	18	17	6
2	18	17	8
3	18	17	12
4	28	27	6
5	28	27	8
6	28	27	12
7	30	29	6
8	30	29	8
9	30	29	12
10	48	47	6
11	48	47	8
12	48	47	12
13	72	71	6
14	72	71	8
15	72	71	12
16	84	83	6
17	84	83	8
18	84	83	12

We compare the performances of formulations BCP-C, proposed in section 6.4, when implemented directly on OPL and solved CPLEX 12.6 with default settings. A two index formulation by removing the index for vehicle in the binary variables reduces the problem to a capacitated TSP. All runs were made on a 2.5 GHz computer having 1 GB of RAM with a maximum CPU time limit of 3600 seconds. The direct use of CPLEX did not solve the problem instances to optimality for node size above 50 within the CPU time limit. Table 6.2 presents results obtained when BCP-C instances are run directly on CPLEX with default settings.

Table 6.2: Computational results for direct application of CPLEX on BCP-C instances.

Instance	N^0	N	c	**Time(Sec)/Optimality gap (%)**	**LB**$^+$	**OFV**$^{++}$
1	18	17	6	16.5	-	28.0645
2	18	17	8	2.66	-	24.537
3	18	17	12	0.64	-	21.361
4	28	27	6	1.64%	50.1886	51.0274
5	28	27	8	542.95	-	42.6608
6	28	27	12	76.44	-	35.4787
7	30	29	6	853.06	-	52.304
8	30	29	8	1460.5	-	44.661
9	30	29	12	19.08	-	37.877
10	48	47	6	5.81%	97.6484	103.6693
11	48	47	8	8.71%	78.9243	86.4519
12	48	47	12	3.56%	62.4836	64.7897
13	72	71	6	8.37%	171.5517	187.2136
14	72	71	8	13.32%	136.1387	157.0611
15	72	71	12	13.10%	103.9599	119.6356
16	84	83	6	8.95%	224.1322	246.1521
17	84	83	8	11.97%	176.3288	200.3071
18	84	83	12	13.41%	129.8938	150.0116

LB$^+$ Lower bound, OFV^{++} Objective function value

A column generation and cutting plane-based methodology presented in section 6.5 were explored to see if problem instances with more than 50 nodes can be solved effectively, but they were found to be ineffective. The column generation-based methodology was not effective for any problem instance, but smaller problem instances (1,2,3,and 6) performed better using the cutting-plane-based methodology. Problem instances 5 and 8 reached the optimal solution within 82 and 32 seconds, respectively, but could not prove optimality in the former case and took 1380 seconds to prove optimality for the latter case.

Table 6.3: Computational results for cutting plane-based methodology on BCP-C instances.

Instance	N^0	N	c	**Time (Sec)/Opt. gap (%)**	**LB**$^+$	**OFV**$^{++}$	**Cuts**
1	18	17	6	1.05	-	28.0645	333
2	18	17	8	0.36	-	24.5373	54
3	18	17	12	0.47	-	21.361	10
4	28	27	6	7.57%	47.1884	51.0515	14978
5	28	27	8	1.58%	41.9848	42.6607	14068
6	28	27	12	9.84	-	35.4787	738
7	30	29	6	6.98%	49.8734	53.6148	16149
8	30	29	8	1379.47	-	44.6608	5422
9	30	29	12	18.55	-	37.877	831
10	48	47	6	34.74%	84.2955	129.1673	14366
11	48	47	8	18.76%	73.1556	90.053	15157
12	48	47	12	13.16%	61.237	70.5184	10598
13	72	71	6	49.12%	120.5816	236.9766	13274
14	72	71	8	50.15%	97.4296	195.4494	13951
15	72	71	12	42.45%	89.2459	155.0668	14039
16	84	83	6	54.39%	140.1231	307.2264	13986
17	84	83	8	55.61%	121.6023	273.935	12104
18	84	83	12	52.38%	100.6844	211.4114	13425

LB$^+$ Lower bound, OFV^{++} Objective function value

Column generation-based methodology can be further improved my devising ways to solve the pricing problem effectively and also by introducing a good starting solution. For the cutting plane-based methodology, sequential introduction of the violated cuts can be explored, instead of adding them as lazy constraints. Different families of cuts can be explored to evaluate if that improves the effectiveness of the cutting-plane-based methodology. Also, structure of the bale collection problem and the uniform distribution of nodes can be studied to develop efficient solution approaches.

6.7 Conclusion

In this Chapter, we introduce a relevant operational-level problem in the biomass feedstock logistics called the bale collection problem (BCP). The BCP and some of its relevant extensions were mathematically modeled as a capacitated vehicle routing problem with unit demand. Since, the standard benchmark instances for CVRP and TSP cannot be used for BCP, we generated problem instances with nodes uniformly distributed, which is in line with characteristics of bale distribution in the actual production field. The solution approaches explored includes direct application of CPLEX, and column generation and cutting plane-based methods. The methodologies explored were not found to be very effective on large-sized instances. For future research, we recommend study of the problem structure to develop effective, valid inequalities and exploring methodologies like "Cluster first and route next" to obtain good heuristic solutions to this problem with larger number of nodes in reasonable computational time. Modifications of clustering algorithms like k-means clustering can be used to obtain clusters of size equal to vehicle capacity. Then, a traveling salesman problem can be solved for each cluster to find sequence of nodes to be visited within a cluster. Exact algorithms using a branch-and-price-based approach may be another viable approach to explore for this problem.

Chapter 7

Concluding Remarks and Directions for Future Research

In this thesis, we have addressed the design and analysis of tactical and operational-level decision making for the biomass feedstock logistics system. In Chapter 1, we have presented the various problems encountered in determining these decisions. The key decisions involved are: fleet size of load-out equipment sets and tractor-trailers, allocation of tractor-trailers to SSLs during every period, routing of equipment sets among the SSLs for unloading biomass to transport to bio-refinery, among others. In Chapter 2, we have provided a detailed description of biomass feedstock logistics system for the Southeast United States to help understand the underlying problems and challenges, and also, to set a background for mathematical and simulation models described later.

In Chapter 3, we have provided a taxonomic review of literature pertaining to the biofuel supply chain based on fundamental operations research problems, modeling methodologies, and solution approaches used in order to help the practitioners in identifying the underlying operations research problems and in taking advantage of the state-of-the-art modeling and solution approaches from extensive research on these well-studied problems.

In Chapter 4, we propose a novel modeling approach to determine tactical decisions like fleet sizes of equipment and trucks by using optimization and simulation models. The application of the proposed methodology proves its ability to provide an implementable robust solution in the face of uncertainties. A general framework of the proposed methodology is provided. For future research, we propose to investigate if this modeling approach can be used for solving stochastic programs in general, and if not, to identify the class of stochastic programming-based models that can use the proposed approach.

In Chapter 5, we have presented a decision support tool for biomass feedstock logistics based on an optimization model. The proposed decision support tool considers current system status, like inventory, current location of equipment sets, inflow of biomass from fields to storage locations, and determines the storage locations to unload biomass from, amount of biomass to be shipped, allocation and routing of tractor-trailers and load-out equipment sets in order to minimize operational cost. The mixed integer programming-based model is implemented in Visual Studio 12.0, and solved using CPLEX 12.6. For future research, a graphic user interface of routes can be explored to provide output to the user in a more lucid manner. Also, additional features like prioritizing unloading of specific SSLs can be included to make the decision support tool more realistic.

In Chapter 6, we introduce an operational-level problem in biomass feedstock logistics called the bale collection problem (BCP). The problem was mathematically modeled as a capacitated vehicle routing problem with unit demand. The solution approaches explored include direct application of CPLEX, and column generation and cutting plane-based methods, and were not found to be very effective on large-sized instances.

Bibliography

[1] M.M. Aguayo, S.C. Sarin, and J.S.Cundiff. A branch and price algorithm for biomass logistic supply chain design. Research Manuscript, 2015.

[2] R. K. Ahuja, J. B. Orlin, S. Pallottino, M. P. Scaparra, and M. G. Scutella. A multi-exchange heuristic for the single-source capacitated facility location problem. *Management Science*, 50(6):749–760, 2004.

[3] Ozlem Akgul, Andrea Zamboni, Fabrizio Bezzo, Nilay Shah, and Lazaros G. Papageorgiou. Optimization-based approaches for bioethanol supply chains. *Industrial & Engineering Chemistry Research*, 50(9):4927–4938, 2011.

[4] W. Alex Marvin, Lanny D. Schmidt, Saif Benjaafar, Douglas G. Tiffany, and Prodromos Daoutidis. Economic Optimization of a Lignocellulosic Biomass-to-Ethanol Supply Chain. *Chemical Engineering Science*, 67(1):68–79, 2012.

[5] D. Alfonso, C. Perpi, A. Prez-Navarro, E. Pealvo, C. Vargas, and R. Crdenas. Methodology for optimization of distributed biomass resources evaluation, management and final energy use. *Biomass and Bioenergy*, 33(8):1070 – 1079, 2009.

[6] Christian Almeder, Margaretha Preusser, and Richard F. Hartl. Simulation and optimization of supply chains: alternative or complementary approaches? *OR Spectrum*, 31(1):95–119, 2008.

[7] Satyajith Amaran, Nikolaos V. Sahinidis, Bikram Sharda, and Scott J. Bury. Simulation optimization: a review of algorithms and applications. *4OR*, 12(4):301–333, 2014.

[8] Heungjo An, Wilbert E. Wilhelm, and Stephen W. Searcy. Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass and Bioenergy*, 35(9): 3763–3774, October 2011.

[9] J. Araque, L. Hall, and T. Magnanti. Capacitated trees, capacitated routing, and associated polyhedra. Core discussion papers, Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE), 1990.

[10] A. Armentano, Vinícius, Paulo M. Franca, and Franklina M.B. de Toledo. A network flow model for the capacitated lot-sizing problem. *Omega*, 27(2):275 – 284, 1999.

[11] P. Augerat, J. Belenguer, E. Benavent, A. Corberan, D. Naddef, and G. Rinaldi. Computational results with a branch and cut code for the capacitated vehicle routing problem. 1995.

[12] Zeinab Azarmand and Ensiyeh Neishabouri. *Facility Location: Concepts, Models, Algorithms and Case Studies*, chapter Location Allocation Problem, pages 93–109. Physica-Verlag HD, 2009.

[13] R. Baldacci, E. Hadjiconstantinou, and A. Mingozzi. An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation. *Oper. Res.*, 52(5):723–738, 2004.

[14] Roberto Baldacci, Nicos Christofides, and Aristide Mingozzi. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*, 115(2):351–385, 2007.

[15] M. L. Balinski and R. E. Quandt. On an integer program for a delivery problem. *Oper. Res.*, 12(2):300–304, 1964.

[16] R. Baos, F. Manzano-Agugliaro, F. G. Montoya, C. Gil, A. Alcayde, and J. Gmez. Optimization methods applied to renewable and sustainable energy: A review. *Renewable and Sustainable Energy Reviews*, 15(4):1753–1766, May 2011.

[17] Imre Barany, Tony J. Van Roy, and Laurence A. Wolsey. Strong formulations for multi-item capacitated lot sizing. *Management Science*, 30(10):1255 – 1261, 1984.

[18] J. Barceló and J. Casanovas. A heuristic lagrangean algorithm for the capacitated plant location problem. *European Journal of Operational Research*, 15(2):212 – 226, 1984.

[19] J.E. Beasley. Lagrangean heuristics for location problems. *European Journal of Operational Research*, 65(3):383 – 399, 1993.

[20] Bilge Bilgen and Yelda Çelebi. Integrated production scheduling and distribution planning in dairy supply chain by hybrid modelling. *Annals of Operations Research*, 211(1):55–82, 2013.

[21] Jack Brimberg, Pierre Hansen, Nenad Mladenovi, and Eric D. Taillard. Improvements and comparison of heuristics for solving the uncapacitated multisource weber problem. *Operations Research*, 48(3):444–460, 2000.

[22] Yolanda Carson and Anu Maria. Simulation optimization: Methods and applications, 1997.

[23] Dirk Cattrysse, Johan Maes, and Luk N. Van Wassenhove. Set partitioning and column generation heuristics for capacitated dynamic lot-sizing. *European Journal of Operational Research*, 46(1):38 – 47, 1990.

[24] Chia-Ho Chen and Ching-Jung Ting. Combining lagrangian heuristic and ant colony system to solve the single source capacitated facility location problem. *Transportation Research Part E: Logistics and Transportation Review*, 44(6):1099 – 1122, 2008.

[25] W.-H. Chen and J.-M. Thizy. Analysis of relaxations for the multi-item capacitated lot-sizing problem. *Annals of Operations Research*, 26(1):29 – 72, 1990.

[26] Nicos Christofides, A. Mingozzi, and P. Toth. State-space relaxation procedures for the computation of bounds to routing problems. *Networks*, 11(2):145–164, 1981.

[27] Michele Conforti, Gerard Cornuejols, and Giacomo Zambelli. *Integer Programming*. Springer Publishing Company, Incorporated, 2014.

[28] J. S. Cundiff, R. Grisso, and D. H. Vaughan. Investigating machinery management parameters with computer tools. Number 071030. presented at the ASAE, St. Joseph, Mi, 2007.

[29] John S. Cundiff. *Herbaceous Biomass Logistics*, pages 219–249. John Wiley & Sons, Ltd, 2014.

[30] John S. Cundiff, Neil Dias, and Hanif D. Sherali. A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource Technology*, 59(1):47 – 55, 1997.

[31] Matteo Dal-Mas, Sara Giarola, Andrea Zamboni, and Fabrizio Bezzo. Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty. *Biomass and Bioenergy*, 35(5):2059 – 2071, 2011.

[32] Claudia D'Ambrosio, Andrea Lodi, and Silvano Martello. *Combinatorial Traveling Salesman Problem Algorithms*. John Wiley & Sons, Inc., 2010.

[33] Marlies de Keizer, Ren Haijema, Jacqueline M. Bloemhof, and Jack G.A.J. van der Vorst. Hybrid optimization and simulation to design a logistics network for distributing perishable products. *Computers & Industrial Engineering*, 88:26 – 38, 2015.

[34] Daniel De La Torre Ugarte, Burton English, and Kim Jensen. Sixty billion gallons by 2030: Economic and agricultural impacts of ethanol and biodiesel expansion. *American Journal of Agricultural Economics*, 89(5):1290–1295, 2007.

[35] Annelies De Meyer, Dirk Cattrysse, Jussi Rasinmki, and Jos Van Orshoven. Methods to optimise the design and management of biomass-for-bioenergy supply chains: A review. *Renewable and Sustainable Energy Reviews*, 31:657–670, March 2014.

[36] Zeger Degraeve and Raf Jans. A new dantzig-wolfe reformulation and branch-and-price algorithm for the capacitated lot-sizing problem with setup times. *Operations Research*, 55(5):909–920, 2007.

[37] M. Diaby, H.C. Bahl, M.H. Karwan, and S. Zionts. Capacitated lot-sizing and scheduling by lagrangean relaxation. *European Journal of Operational Research*, 59(3):444 – 458, 1992.

[38] J. A. Díaz and E. Fernández. A branch-and-price algorithm for the single source capacitated plant location problem. *The Journal of the Operational Research Society*, 53(7):728–740, 2002.

[39] Marco A. Duran and Ignacio E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming*, 36(3):307–339, 1986.

[40] Silke Van Dyken, Bjorn H. Bakken, and Hans I. Skjelbred. Linear mixed-integer models for biomass supply chains with transport, storage and processing. *Energy*, 35(3):1338 – 1350, 2010.

[41] Mahmood Ebadian, Taraneh Sowlati, Shahab Sokhansanj, Mark Stumborg, and Lawrence Townley-Smith. A new simulation model for multi-agricultural biomass logistics system in bioenergy production. *Biosystems Engineering*, 110(3):280 – 290, 2011.

[42] Mahmood Ebadian, Taraneh Sowlati, Shahab Sokhansanj, Lawrence T. Smith, and Mark Stumborg. Development of an integrated tactical and operational planning model for supply of feedstock to a commercial-scale bioethanol plant. *Biofuels, Bioproducts and Biorefining*, 8(2):171–188, 2014.

[43] Sandra Ekiolu, Song Li, Shu Zhang, Shahabaddine Sokhansanj, and Daniel Petrolia. Analyzing Impact of Intermodal Facilities on Design and Management of Biofuel Supply Chain. *Transportation Research Record: Journal of the Transportation Research Board*, 2191:144–151, 2010.

[44] Sandra D. Ekiolu, Ambarish Acharya, Liam E. Leightley, and Sumesh Arora. Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers & Industrial Engineering*, 57(4):1342–1352, 2009.

[45] Gary D. Eppen and R. Kipp Martin. Solving multi-item capacitated lot-sizing problems using variable redefinition. *Operations Research*, 35(6):832 – 848, 1984.

[46] S.L. Fales, J. Richard Hess, and W.W. Wilhelm. Convergence of agriculture and energy: producing cellulosic biomass for biofuels. *Engineering & Technology for a Sustainable World*, 2008.

[47] Matteo Fischetti, Andrea Lodi, and Paolo Toth. *The Traveling Salesman Problem and Its Variations*, chapter Exact Methods for the Asymmetric Traveling Salesman Problem, pages 169–205. Springer US, Boston, MA, 2007.

[48] F. Frombo, R. Minciardi, M. Robba, F. Rosso, and R. Sacile. Planning woody biomass logistics for energy production: A strategic decision model. *Biomass and Bioenergy*, 33(3):372 – 383, 2009.

[49] A. M. Geoffrion and G. W. Graves. Multicommodity distribution system design by benders decomposition. *Management Science (pre-1986)*, 20(4):822–845, 1974.

[50] Dijin Gong, Mitsuo Gen, Genji Yamazaki, and Weixuan Xu. Hybrid evolutionary method for capacitated location-allocation problem. *Computers & Industrial Engineering*, 33(34):577 – 580, 1997.

[51] C. Gracia, B. Diezma-Iglesias, and P. Barreiro. A hybrid genetic algorithm for route

optimization in the bale collecting problem. *Spanish Journal of Agricultural Research*, 11(3):603–614, 2013.

[52] Alexander Grigoriev and Joris van de Klundert. On the high multiplicity traveling salesman problem. *Discrete Optimization*, 3(1):50 – 62, 2006.

[53] Robert D. Grisso, Devita McCullough, John S. Cundiff, and Jason D. Judd. Harvest schedule to fill storage for year-round delivery of grasses to biorefinery. *Biomass and Bioenergy*, 55:331 – 338, 2013.

[54] G. Guastaroba and M.G. Speranza. A heuristic for {BILP} problems: The single source capacitated facility location problem. *European Journal of Operational Research*, 238 (2):438 – 450, 2014.

[55] Helene Gunnarsson, Mikael Rnnqvist, and Jan T Lundgren. Supply chain modelling of forest fuel. *European Journal of Operational Research*, 158(1):103 – 123, 2004.

[56] Eleni Hadjiconstantinou, Nicos Christofides, and Aristide Mingozzi. A new exact algorithm for the vehicle routing problem based on q-paths and k-shortest paths relaxations. *Annals of Operations Research*, 61(1):21–43, 1995.

[57] Patrick T. Hester and Kevin MacG. Adams. *Systemic Thinking - Fundamentals for Understanding Problems and Messes*. Springer International Publishing, 2014.

[58] K. S. Hindi. Efficient solution of the single-item, capacitated lot-sizing problem with start-up and reservation costs. *The Journal of the Operational Research Society*, 46 (10):1223–1236, 1995.

[59] Kaj Holmberg, Mikael Ronnqvist, and Di Yuan. An exact algorithm for the capacitated facility location problems with single sourcing. *European Journal of Operational Research*, 113(3):544 – 559, 1999.

[60] Yongxi Huang, Chien-Wei Chen, and Yueyue Fan. Multistage optimization of the supply chains of biofuels. *Transportation Research Part E: Logistics and Transportation Review*, 46(6):820 – 830, 2010.

[61] Jason D. Judd. *Modeling and Analysis of a Feedstock Logistics Problem*. PhD thesis, Virginia Polytechnic Institute and State University, 2012.

[62] Jason D. Judd, Subhash C. Sarin, and John S. Cundiff. Design, modeling, and analysis of a feedstock logistics system. *Bioresource Technology*, 103(1):209 – 218, 2012.

[63] Imdat Kara and Tolga Bektas. Integer linear programming formulations of multiple salesman problems and its variations. *European Journal of Operational Research*, 174 (3):1449 – 1458, 2006.

[64] B. Karimi, S.M.T. Fatemi Ghomi, and J.M. Wilson. The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 31(5):365 – 378, 2003.

[65] Jinkyung Kim, Matthew J. Realff, and Jay H. Lee. Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty. *Computers & Chemical Engineering*, 35(9):1738–1751, 2011.

[66] Jinkyung Kim, Matthew J. Realff, Jay H. Lee, Craig Whittaker, and Ludwig Furtner. Design of biomass processing network for biofuel production using an {MILP} model. *Biomass and Bioenergy*, 35(2):853 – 871, 2011.

[67] John G. Klincewicz and Hanan Luss. A lagrangian relaxation heuristic for capacitated facility location with single-source constraints. *The Journal of the Operational Research Society*, 37(5):495–500, 1986.

[68] Robert E. Kuenne and Richard M. Soland. Exact and approximate solutions to the multisource weber problem. *Mathematical Programming*, 3(1):193–209, 1972.

[69] Amit Kumar and Shahab Sokhansanj. Switchgrass (Panicum vigratum, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresource Technology*, 98(5):1033–1044, 2007.

[70] Gilbert Laporte. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59(3):345 – 358, 1992.

[71] Gilbert Laporte. The traveling salesman problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59(2):231 – 247, 1992.

[72] Gilbert Laporte. What you should know about the vehicle routing problem. *Naval Research Logistics (NRL)*, 54(8):811–819, 2007.

[73] Gilbert Laporte and Yves Nobert. Exact algorithms for the vehicle routing problem. volume 132, pages 147 – 184. 1987.

[74] Gilbert Laporte, Yves Nobert, and Martin Desrochers. Optimal routing under capacity and distance restrictions. *Oper. Res.*, 33(5):1050–1073, 1985.

[75] Averill M. Law. Designing a simulation study: How to conduct a successful simulation study. In *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation*, WSC '03, pages 66–70. Winter Simulation Conference, 2003.

[76] JannyM.Y. Leung, ThomasL. Magnanti, and Rita Vachani. Facets and algorithms for capacitated lot sizing. *Mathematical Programming*, 45(1-3):331–359, 1989.

[77] Tao Lin, Luis F Rodrguez, Yogendra N Shastri, Alan C Hansen, and KC Ting. Gis-enabled biomass-ethanol supply chain optimization: model development and miscanthus application. *Biofuels, Bioproducts and Biorefining*, 7(3):314–333, 2013.

[78] Tao Lin, Luis F. Rodrguez, Yogendra N. Shastri, Alan C. Hansen, and K.C. Ting. Integrated strategic and tactical biomassbiofuel supply chain optimization. *Bioresource Technology*, 156:256–266, 2014.

[79] Jude Liu, Robert Grisso, and John Cundiff. *Harvest Systems and Analysis for Herbaceous Biomass*. InTech, 2013.

[80] Jens Lysgaard, N. Adam Letchford, and W. Richard Eglese. A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming*, 100(2):423–445, 2003.

[81] Fereshteh Mafakheri and Fuzhan Nasiri. Modeling of biomass-to-energy supply chain operations: Applications, challenges and research directions. *Energy Policy*, 67:116–126, April 2014.

[82] T. L. Magnanti. Combinatorial optimization and vehicle fleet planning: Perspectives and prospects. *Networks*, 11(2):179–213, 1981.

[83] Alan S. Manne. Programming of economic lot sizes. *Management Science*, 4(2):115–135, 1958.

[84] Alexandra F. Marques, Jorge P. de Sousa, Mikael Rnnqvist, and Ricardo Jafe. Combining optimization and simulation tools for short-term planning of forest operations. *Scandinavian Journal of Forest Research*, 29:166–177, 2014.

[85] Mahdi Mobini, Taraneh Sowlati, and Shahab Sokhansanj. Forest biomass supply logistics for a power plant using the discrete-event simulation approach. *Applied Energy*, 88(4):1241 – 1250, 2011.

[86] R.M.De. Mol, M.A.H. Jogems, P. Van Beek, and J.K. Gigler. Simulation and optimization of the logistics of biomass fuel collection. *Netherlands Journal of Agricultural Science*, 45:219 – 228, 1997.

[87] Alan T. Murray and Richard L. Church. Applying simulated annealing to location-planning models. *Journal of Heuristics*, 2(1):31–53, 1996.

[88] S.N. Naik, Vaibhav V. Goud, Prasant K. Rout, and Ajay K. Dalai. Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 14(2):578 – 597, 2010.

[89] Barry Nelson. *Foundations and Methods of Stochastic Simulation: A First Course*. Springer Publishing Company, Incorporated, 2013.

[90] M. Ohlemller. Tabu search for large location-allocation problems. *The Journal of the Operational Research Society*, 48(7):745–750, 1997.

[91] Temel Öncan, İ. Kuban Altnel, and Gilbert Laporte. A comparative analysis of several asymmetric traveling salesman problem formulations. *Computers & Operations Research*, 36(3):637 – 654, 2009.

[92] C. Dennis Pegden and David T. Sturrock. Introduction to simio. In *Proceedings of the Winter Simulation Conference*, WSC '11, pages 29–38. Winter Simulation Conference, 2011.

[93] C. Dennis Pegden and David T. Sturrock. Recent innovations in simio. In *Proceedings of the Winter Simulation Conference*, WSC '12, pages 437:1–437:12. Winter Simulation Conference, 2012.

[94] P.P. Ravula, R.D. Grisso, and J.S. Cundiff. Comparison between two policy strategies for scheduling trucks in a biomass logistic system. *Bioresource Technology*, 99(13):5710 – 5721, 2008.

[95] Jingzheng Ren, Alessandro Manzardo, Sara Toniolo, Antonio Scipioni, Shiyu Tan, Lichun Dong, and Suzhao Gao. Design and modeling of sustainable bioethanol supply chain by minimizing the total ecological footprint in life cycle perspective. *Bioresource Technology*, 146:771 – 774, 2013.

[96] J. P. Resop, J. S. Cundiff, and C. D. Heatwole. Spatial analysis to site satellite storage

locations for herbaceous biomass in the piedmont of the southeast. *Applied Engineering in Agriculture*, 27(1):25–32, 2011.

[97] Roberto Roberti and Paolo Toth. Models and algorithms for the asymmetric traveling salesman problem: an experimental comparison. *EURO Journal on Transportation and Logistics*, 1(1):113–133, 2012.

[98] Ralph M. Rotty. Growth in global energy demand and contribution of alternative supply systems. *Energy*, 4(5):881 – 890, 1979.

[99] Murali Sambasivan and Salleh Yahya. A lagrangean-based heuristic for multi-plant, multi-item, multi-period capacitated lot-sizing problems with inter-plant transfers. *Computers & Operations Research*, 32(3):537 – 555, 2005.

[100] Robert G. Sargent. Verification and validation of simulation models. In *Proceedings of the 30th Conference on Winter Simulation*, WSC '98, pages 121–130. IEEE Computer Society Press, 1998.

[101] Subhash C. Sarin, Hanif D. Sherali, and Liming Yao. New formulation for the high multiplicity asymmetric traveling salesman problem with application to the Chesapeake problem. *Optimization Letters*, 5(2):259–272, 2011.

[102] Subhash C. Sarin, Hanif D. Sherali, Jason D. Judd, and Pei-Fang (Jennifer) Tsai. Multiple asymmetric traveling salesmen problem with and without precedence constraints: Performance comparison of alternative formulations. *Comput. Oper. Res.*, 51:64–89, November 2014.

[103] Nazanin Shabani and Taraneh Sowlati. A mixed integer non-linear programming model for tactical value chain optimization of a wood biomass power plant. *Applied Energy*, 104:353–361, 2013.

[104] Nazanin Shabani, Shaghaygh Akhtari, and Taraneh Sowlati. Value chain optimization

of forest biomass for bioenergy production: A review. *Renewable and Sustainable Energy Reviews*, 23:299 – 311, 2013.

[105] Shahriar Shafiee and Erkan Topal. When will fossil fuel reserves be diminished? *Energy Policy*, 37(1):181 – 189, 2009.

[106] B. Sharma, R. G. Ingalls, C. L. Jones, and A. Khanchi. Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, 24:608–627, August 2013.

[107] Bhavna Sharma, Ricki G. Ingalls, Carol L. Jones, Raymond L. Huhnke, and Amit Khanchi. Scenario optimization modeling approach for design and management of biomass-to-biorefinery supply chain system. *Bioresource Technology*, 150:163 – 171, 2013.

[108] S Sokhansanj, A Kumar, and A Turhollow. Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy*, 30(10):838–847, 2006.

[109] Shahab Sokhansanj, Mahmood Ebadian, Taraneh Sowlati, Lawrence Townley-Smith, and Mark Stumborg. Modeling and analysing storage systems in agricultural biomass supply chain for cellulosic ethanol production. *Applied Energy*, 102:840 – 849, 2013.

[110] R. Sridharan. A lagrangian heuristic for the capacitated plant location problem with single source constraints. *European Journal of Operational Research*, 66(3):305 – 312, 1993.

[111] C. S. Sung. A single-product parallel-facilities production-planning model. *International Journal of Systems Science*, 17(7):983–989, 1986.

[112] Hamdy A. Taha. *Operations Research: An Introduction (8th Edition)*. Prentice-Hall, Inc., 2006.

[113] Eylem Tekin and Ihsan Sabuncuoglu. Simulation optimization: A comprehensive review on theory and applications. *IIE Transactions*, 36(11):1067–1081, 2004.

[114] J. M. Thizy and L. N. Van Wassenhove. Lagrangean relaxation for the multi-item capacitated lot-sizing problem: A heuristic implementation. *IIE Transactions*, 17(4): 308 – 313, 1985.

[115] B. Velazquez-Marti and E. Fernandez-Gonzalez. Mathematical algorithms to locate factories to transform biomass in bioenergy focused on logistic network construction. *Renewable Energy*, 35(9):2136 – 2142, 2010.

[116] Johannes Windisch, Kari Vtinen, Perttu Anttila, Mikko Nivala, Juha Laitila, Antti Asikainen, and Lauri Sikanen. Discrete-event simulation of an information-based raw material allocation process for increasing the efficiency of an energy wood supply chain. *Applied Energy*, 149:315 – 325, 2015.

[117] Wayne L. Winston. *Operations Research. Applications and Algorithms.* Brooks/Cole, 4th edition, 2004.

[118] Fei Xie, Yongxi Huang, and Sandra Eksioglu. Integrating multimodal transport into cellulosic biofuel supply chain design under feedstock seasonality with a case study based on california. *Bioresource Technology*, 152:15 – 23, 2014.

[119] Dajun Yue, Fengqi You, and Seth W. Snyder. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, 66:36–56, July 2014.

[120] Willard I. Zangwill. A deterministic multiproduct, multifacility production and inventory model. *Operations Research*, 14(3):486–507, 1966.

[121] Jun Zhang, Atif Osmani, Iddrisu Awudu, and Vinay Gonela. An integrated optimization model for switchgrass-based bioethanol supply chain. *Applied Energy*, 102: 1205–1217, 2013.

[122] Xiaoyan Zhu and Qingzhu Yao. Logistics system design for biomass-to-bioenergy industry with multiple types of feedstocks. *Bioresource Technology*, 102(23):10936 – 10945, 2011.

Appendix A

Input modeling

Input data and process logic are the most important components in a simulation model. Verification and validation procedures are carried out to check if these two components are modeled correctly. Input modeling comprise of three main steps as follows:

- Step 1: Collection of data
- Step 2: Selecting a distribution to fit the collected data
- Step 3: Conduct goodness of fit test

Collection of Data: The process of collecting data is onerous, time consuming and expensive. Also, a clear understanding of what data is needed before data collection exercise and hence it is recommended to collect data after building the simulation model. The data collected will provide insights about the randomness in the process being modeled.

Selection of a distribution: The collected data is used to fit a distribution, and the simulation model will use random samples from the distribution and not the collected data directly. The distribution selection hierarchy is as follows:

1. Use collected data to fit the distribution using an input modeling software

2. Use raw data and load discrete points into a custom distribution (i.e., empirical CDF)

3. Use the distribution suggested by the nature of the process or underlying physics

4. Assume a simple distribution and apply reasonable limits when lacking data (e.g., Pert, Triangular)

Several softwares are available to fit distributions for collected data, e.g., @Risk, Matlab, Minitab, SAS are some among others. Statistical methods like maximum likelihood estimation is used to estimate parameters for the fitted distributions.

Goodness of fit test: After the distribution and its parameters are selected, it can be compared to the collected data through visual inspection and statistical tests. Visual methods includes histograms, comparison of fitted CDF versus the data based CDF, Q-Q plots and P-P plots as shown in Figure A.3. The statistical tests are based on classical statistical testing in which a null hypothesis is compared to an alternative. Three commonly used goodness of fit tests are the Chi-Squared test, the Kolmogorov-Smirnov test, and the Anderson Darling test. The figures below show parameter estimation, goodness of fit test, and Q-Q plot using @Risk software.

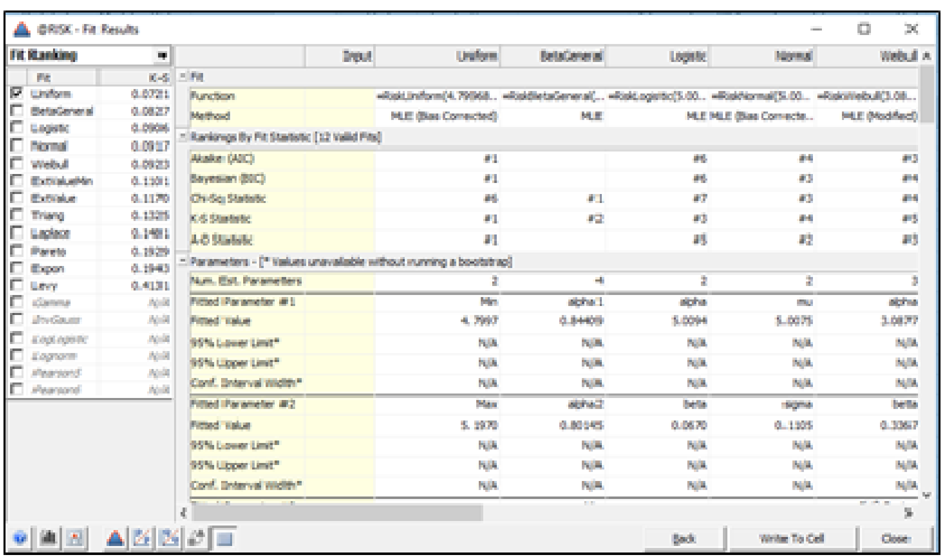


Figure A.1: Parameter estimation.

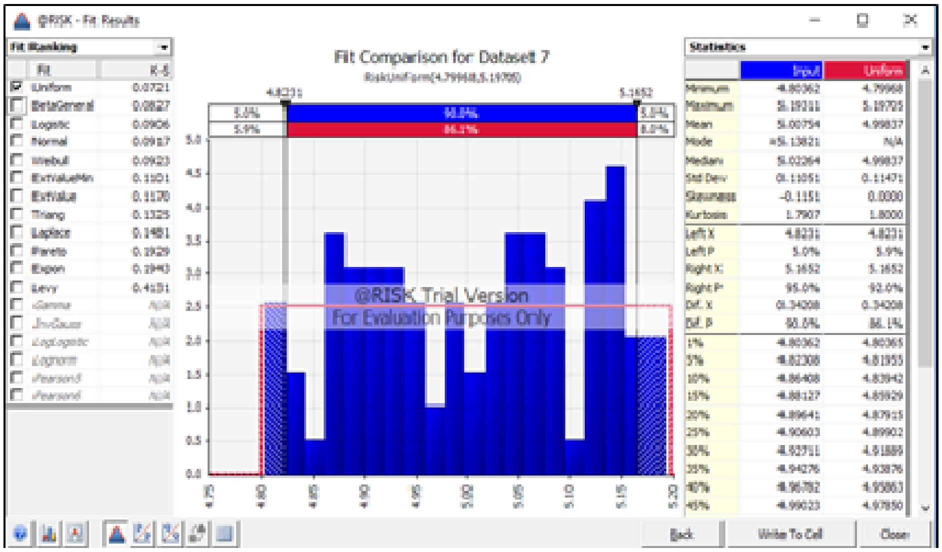


Figure A.2: Goodness of fit test.

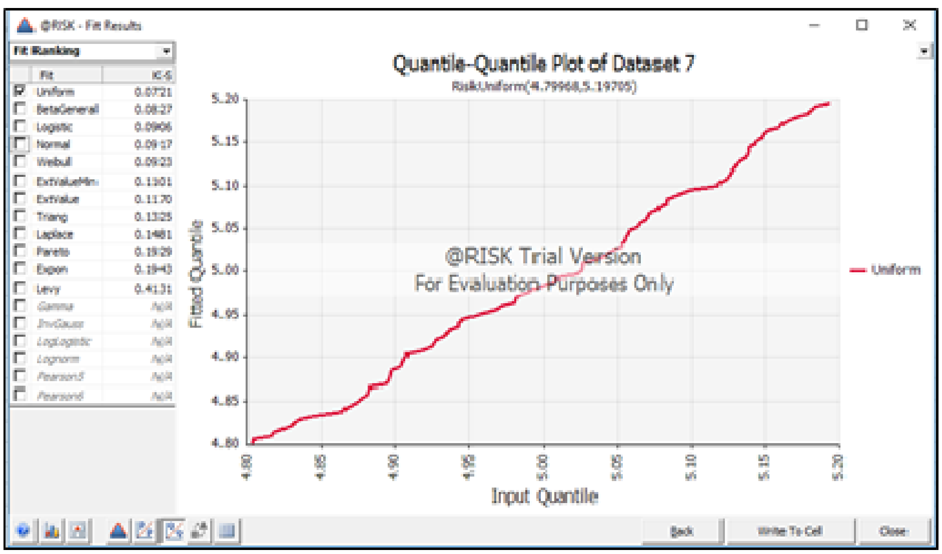


Figure A.3: Q-Q plot.

The Chi-Squared (χ^2) test is based on comparing the observed number of observations in a histogram cell with the number expected, if the fitted distribution was the input model. χ^2 test can be applied to discrete distributions like binomial and Poisson distributions. The Kolmogorov-Smirnov (K-S) test is a non-parametric test that looks at the largest (vertical) distance between the fitted CDF and the data-based CDF. Anderson Darling (A-D) test is a generalization of K-S test, but puts more weight in the tails of the difference unlike K-S test which uses equal weights. The tail of the fitted distribution is very important as the variation in the distribution is largest in tails and it is the tails that affects queuing, utilization and flow time statistics. Both K-S and A-D tests are restricted to continuous distributions. For more detailed description on input modeling for simulation, readers are directed to Nelson [89].