Biomass-To-Biofuels' Supply Chain Design And Management

Ambarish Madhukar Acharya

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
https://scholarsjunction.msstate.edu/td/890

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.
BIOMASS-TO-BIOFUELS’ SUPPLY CHAIN DESIGN AND MANAGEMENT

By

Ambarish Madhukar Acharya

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Industrial Engineering
in the Department of Industrial & Systems Engineering

Mississippi State, Mississippi

December 2010
Copyright by

Ambarish Madhukar Acharya

2010
BIOMASS-TO-BIOFUELS’ SUPPLY CHAIN DESIGN AND MANAGEMENT

By

Ambarish Madhukar Acharya

Approved:

Sandra Ekşioğlu
Assistant Professor of
Industrial & Systems Engineering
(Director of Dissertation)

Burak Ekşioğlu
Associate Professor of
Industrial & Systems Engineering
(Committee Member)

Mingzhou Jin
Associate Professor of
Industrial & Systems Engineering
(Committee Member)

Daniel R. Petrolia
Assistant Professor of
Agricultural Economics
(Committee Member)

John Usher
Professor of
Industrial & Systems Engineering
(Graduate Coordinator)

Sarah A. Rajala
Dean of Bagley College of
Engineering
The goal of this dissertation is to study optimization models that integrate location, production, inventory and transportation decisions for industrial products and apply the knowledge gained to develop supply chains for agricultural products (biomass). We estimate unit cost for the whole biomass-to-biofuels’ supply chain which is the per gallon cost for biofuels up till it reaches the markets. The unit cost estimated is the summation of location, production, inventory holding, and transportation costs.

In this dissertation, we focus on building mathematical models for designing and managing the biomass-to-biofuels’ supply chains. The computational complexity of the developed models makes it advisable to use heuristic solution procedures. We develop a Lagrangean decomposition heuristic. In our heuristic, we divide the problem into two sub-problems, sub-problem 1 is a transportation problem and sub-problem 2 is a combination of a capacitated facility location and production planning problem. Sub-problem 2 is further divided by commodities. The algorithm is tested for a number of different scenarios.
We also develop a decision support system (DSS) for the biomass-to-biofuels’ supply chain. In our DSS, the main problem is divided into four easy-to-solve supply chain problems. These problems were determined based on our knowledge of supply chain and discussions with the experts from the biomass and biofuels’ sector. The DSS is coded using visual basic applications (VBA) for Excel and has a simple user interface which assists the user in running different types of supply chain problems and provides results in form of reports which are easy to understand.
DEDICATION

I dedicate this manuscript to the ONE and to my parents…
ACKNOWLEDGEMENTS

I would like to thank the people who helped me complete the work contained in this dissertation. The help of my supervisor Sandra Ekşioğlu was of great value. I would like to thank Sandra Ekşioğlu for her technical advice, encouragement and insightful comments throughout my dissertation work. Her unconditional support in solving many details surrounding this dissertation and her valuable feedbacks are deeply appreciated.

I extend my thanks to the members of my committee Burak Ekşioğlu, Mingzhou Jin and Daniel Petrolia for their constructive criticism concerning the material of this dissertation. I also would like to express my appreciation to all my friends at the ISE department and in Starkville. In particular I would like to thank my room-mates Arun, Saroj and Shyamesh; my friends Shweta and Lorelei; my lab-mates Fatemeh, Engin, Harun, Huseyin, Shu, Gökçe, Li, Daniela, Roni, and Abdullah for lightening up my life every day and making Starkville, Stark Vegas, and graduate school more fun than I am sure it is supposed to be. Special thanks to Mandeep for her unconditional support and constant motivation.

I would like to express my special thanks to my parents Madhavi and Madhukar Acharya, my sister Lopamudra Karnik and my niece Mrunmayee Karnik. Their understanding and faith in me and my capabilities, their love, encouragement, and eternal support have motivated me all the time. Last but not least, I would like to thank my
childhood buddies Nikhil, Punit, Dhawal and Nirav and my friends and family back in India for their continuous support throughout all my years at Mississippi State University.
TABLE OF CONTENTS

DEDICATION .................................................................................................................. ii

ACKNOWLEDGEMENTS ............................................................................................... iii

LIST OF TABLES ........................................................................................................ vii

LIST OF FIGURES ...................................................................................................... viii

EXECUTIVE SUMMARY ............................................................................................ ix

CHAPTER

I. INTRODUCTION ....................................................................................................... 1

   Background ............................................................................................................. 3
   Research Motivations ........................................................................................... 6
   Research Objectives and Contributions ............................................................... 9

II. LITERATURE REVIEW ............................................................................................ 15

   Supply Chain Design ............................................................................................ 16
      Supply Chain Design for Industrial Products ................................................ 16
      Supply Chain Design for Agricultural Products ............................................. 17
   Supply Chain Management .................................................................................. 21
      Supply Chain Management for Industrial Products ........................................ 22
      Supply Chain Management for Agricultural Products ................................. 24
      Farm Planning ................................................................................................... 24
      Crop Rotation Mix ......................................................................................... 26
      Transportation Models .................................................................................... 27
      Supply Chain Management in Agricultural Fields (Holistic Approach) ......... 28
      Supply Chains for Biomass ............................................................................. 30

III. PROBLEM FORMULATION .................................................................................... 35

   Problem Description ............................................................................................ 35
   Problem Formulation ........................................................................................... 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Parameters</td>
<td>37</td>
</tr>
<tr>
<td>Decision Variables</td>
<td>38</td>
</tr>
<tr>
<td>MIP Model</td>
<td>39</td>
</tr>
<tr>
<td>Related Problems</td>
<td>43</td>
</tr>
<tr>
<td>Capacitated Facility Location Problem</td>
<td>43</td>
</tr>
<tr>
<td>Multi-Commodity Network Flow Problem</td>
<td>46</td>
</tr>
<tr>
<td>Production Planning Problem</td>
<td>48</td>
</tr>
<tr>
<td>IV. SOLUTION PROCEDURES</td>
<td>52</td>
</tr>
<tr>
<td>Computational Complexity of the Problem</td>
<td>52</td>
</tr>
<tr>
<td>Review of Solution Procedures</td>
<td>52</td>
</tr>
<tr>
<td>Exact Mathematical Programming Techniques</td>
<td>53</td>
</tr>
<tr>
<td>Heuristic Algorithms</td>
<td>57</td>
</tr>
<tr>
<td>Decomposition Methods</td>
<td>60</td>
</tr>
<tr>
<td>A Heuristic-Based Solution Approach</td>
<td>65</td>
</tr>
<tr>
<td>A Lagrangean Decomposition-Based Solution Approach</td>
<td>67</td>
</tr>
<tr>
<td>Sub-Gradient Optimization</td>
<td>72</td>
</tr>
<tr>
<td>V. A DECISION SUPPORT SYSTEM</td>
<td>74</td>
</tr>
<tr>
<td>Algorithms for Facility Location Problems</td>
<td>75</td>
</tr>
<tr>
<td>The Modified Problem</td>
<td>78</td>
</tr>
<tr>
<td>Problem Parameters</td>
<td>78</td>
</tr>
<tr>
<td>Decision Variables</td>
<td>79</td>
</tr>
<tr>
<td>Modified MIP Model</td>
<td>79</td>
</tr>
<tr>
<td>Different Supply Chain Problems</td>
<td>80</td>
</tr>
<tr>
<td>DSS Algorithm</td>
<td>82</td>
</tr>
<tr>
<td>VI. EXPERIMENTAL RESULTS</td>
<td>86</td>
</tr>
<tr>
<td>Case Study</td>
<td>86</td>
</tr>
<tr>
<td>Input Data</td>
<td>86</td>
</tr>
<tr>
<td>Corn Stover</td>
<td>87</td>
</tr>
<tr>
<td>Forest Residues</td>
<td>88</td>
</tr>
<tr>
<td>C-ethanol</td>
<td>88</td>
</tr>
<tr>
<td>Facilities</td>
<td>88</td>
</tr>
<tr>
<td>Computational Results</td>
<td>89</td>
</tr>
<tr>
<td>MIP Model</td>
<td>89</td>
</tr>
<tr>
<td>Lagrangean Decomposition Heuristics</td>
<td>101</td>
</tr>
<tr>
<td>DSS Model</td>
<td>103</td>
</tr>
<tr>
<td>VII. CONCLUSIONS</td>
<td>111</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>113</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Decision Types for Supply Chain Management .................................................... 5
2. Authors and Fields of Research .......................................................................... 11
3. Different Algorithms and Facility Location Problems ........................................ 77
4. Different Scenarios ............................................................................................ 90
5. Cost Distribution for Different Scenarios ........................................................... 92
6. Biomass Processed with Biomass Supply ........................................................... 95
7. Unit Delivery Cost of C-ethanol with Conversion Rates ..................................... 99
8. CPU Times vs. Conversion Rates ....................................................................... 100
9. Computational Results for CPLEX and LD Heuristics ..................................... 102
10. Average Unit Delivery Cost for DSS and CPLEX ................................ .......... 104
11. Minimum, Average and Maximum Solution Times for DSS and CPLEX ............... 105
12. Minimum, Average and Maximum GAP between DSS and CPLEX Solutions .......................................................... 105
13. Results for Problem S4 from DSS and CPLEX ................................ .......... 108
14. Biorefinery Locations and Sizes for Problem S4 from DSS and CPLEX ............... 109
LIST OF FIGURES

1. Ethanol Supply Chain Network ................................................................. 8
2. Characterization of Literature on Supply Chain Optimization ............. 12
3. Biomass Availability for Biofuels .............................................................. 42
4. Various Decomposition Approaches ....................................................... 68
5. Different Types of Biomass-To-Biorefinery Supply Chain Problems .... 81
6. Flow Chart for DSS ............................................................................. 82
7. Changes in Unit Delivery C-ethanol Cost ............................................. 94
8. Changes in Amount of Biomass Types Processed ............................... 96
9. Changes in Cost Parameters with Biomass Supply .............................. 97
10. Average GAP – DSS and CPLEX Solutions for Problem S1 ............... 106
11. Average Time – DSS and CPLEX for Problem S1 ............................... 107
12. Ethanol Unit Delivery Cost – DSS and CPLEX Solutions for Problem S1 108
EXECUTIVE SUMMARY

The main objective of this research is to develop efficient biomass-to-biorefinery supply chains for making biofuels a feasible option. Studies show that ethanol production increased from 1.6 billion gallons in 2000 to 4.9 billion gallons in 2006 and 9.0 billion gallons in 2008 [1]; however, it satisfied only 3% of the total gasoline demand. The literature indicates that the main factors affecting biofuels production are uncertainties in biomass supplies and its logistics. Biomass can be harvested during specific periods in a year, and biofuels should be produced year-round to maintain a steady supply satisfying its demands. Also biomass is bulky and voluminous to transport. Thus, uncertain and seasonal biomass supplies illustrate a need for better inventory management for biomass so as to provide a steady supply to the conversion facilities, and together with higher biomass transportation costs illustrate the need for a well-designed and managed biomass supply network. Considering these factors in biomass supply chain designs would give robust biorefinery locations, which are strategically located and would take into account the uncertainties involved in the whole process before deciding upon a location to choose, and also reduced biofuels costs. A mixed integer linear programming (MILP) model is developed for designing and managing biomass supply chain. The model is applied to a case study. Different scenarios based on changes in problem parameters’ values are
constructed. Finally, numerical experiments are done to measure the performance of the biomass supply chain.

The biomass supply chain problem is formulated as a multi-commodity network design problem (MCNDP). MCNDPs have gained considerable attention in the literature and have been used to solve real world applications. In current literature, many researchers have successfully applied mathematical models and have developed solution algorithms to solve multi-commodity supply chain design and management problems separately [2-7]. There is not much literature available in the area related to multi-commodity problems since very few researchers have studied the problem of supply chain design and management simultaneously. Additionally, most of these studies have considered industrial products. In this research we propose multi-commodity models that reflect the specifics of biomass. We design multi-commodity network flow models for supplying biomass to the conversion facilities and develop a heuristic procedure to solve these problems in an efficient way.

The long term goal of this research is to make biofuels a feasible option. The two objectives of this research are: (1) to develop a mixed-integer linear programming (MILP) model for designing and managing biomass supply chain and (2) to develop solution algorithms for solving the MCNDP optimally. Successful completion of this research would provide efficient design and management tools for the biomass supply chain with strategically located biorefineries, significantly reducing the production and distribution costs. This would ultimately help in reducing global warming and would also help take the nation one step further towards energy independence. On the technical
aspect, the solution algorithms will help researchers to solve the multi-commodity supply chain design and management problems for biorefineries in an efficient way.
CHAPTER I
INTRODUCTION

Supply Chain Design and Supply Chain Management are well studied areas in the field of operations research. Globalization has put a lot of emphasis on development of tools that help with designing and managing supply chains efficiently. For example, to take advantage of the workforce in global markets (like China and India), to minimize delivery costs and maximize profits, more products are being imported into US retail markets. There are many reasons for producing in these countries and importing it to USA: lower labor costs, less intense government regulations, economies of scales, better transportation technology, lower capital costs, etc. With ever growing consumer markets and greater competition, a successful business has to constantly keep pace ahead of its competitors to gain a higher market share. One of the many ways to gain a higher market share is to provide customers with the highest quality products and services at lower costs. As a result, research efforts in the field of supply chain design and management for industrial products has seen a boom in recent years. At the same time, research efforts in the field of supply chain design and management for agricultural products has not developed at the same pace. This is due to the large number of government regulations for agricultural products, like controlling prices of certain agricultural commodities, policies encouraging production of certain agricultural products, regulation on
distribution of certain food commodities, etc.; and also due to the complex nature of their supply chains. Supply chains for most of the industrial products are quite similar in nature but the same is not true for agricultural products. Every agricultural product would have certain characteristics which would differentiate its supply chain. For example, to refill soil of its nutrients and get better corn yields, corn is planted and harvested in rotation with some other grain like wheat or soybean. This means one cannot get corn in every season where as for certain other products like fruits (apple, mango, etc.) can be harvested in every season. Also, for certain agricultural products like grains, pre-processing is a required step in the supply chain, whereas vegetables like potatoes can be directly sent to packaging and shipped to markets.

Interest in the field of supply chains for agricultural products is on the rise because of the growing markets for renewable energy, especially biofuels. The US government is encouraging investments in the biofuels sector and encouraging multi-disciplinary research for increasing biofuels’ production and utilization, by its policies and acts (for example, Energy Policy Act of 2005). Efficient supply chain design and management of biomass (agricultural product/byproduct) would play a pivotal role in efficient production and distribution of biofuels. The aim of our research is to develop such optimization models which would help in designing efficient logistics distribution networks for biomass and minimize delivery costs for biofuels, making them a viable option.
Background

Today it is becoming very important for manufacturers to provide customers with goods at low prices and at higher service levels. Increase in product variety increases the complexity of supply chain design and management. Manufacturers are in constant search for ways of reducing their operating costs. One of the ways is to coordinate decisions about different functions within the company and within the corresponding supply chain. For this purpose, companies can use multi-commodity supply chain optimization models which combine facility location decisions with production and inventory decisions for multiple products. These models help to reduce the costs of managing the supply chain. In addition, coordination between the supply chain activities is also important. Supply chain coordination has gained importance in recent times due to increasing competition among the firms to achieve the utmost level of efficiency of the system in order to minimize the total costs and increase the firms’ profits. Thus it is very important that the flow of information is maintained within the supply chain so as to achieve the required coordination among the supply chain activities.

The multi-commodity network design (MCND) problem has spurred interest of many researchers due to its real world applications. Applications of the MCND problem include telecommunication network design [8, 9], production scheduling and planning [10, 11], logistics and transportation [12-14], etc. For more information the reader is directed to [15]. In most multi-commodity supply chains, major contributors to operating costs are logistics related costs. For example, in the automotive industry, logistics costs are major contributors to the total delivery costs for automotives. The same holds true in the case of biomass supply chains. The logistical costs related to biomass are high
because of several reasons. Biomass is bulky in nature and its bulk density is very low. Thus to improve efficiencies of such supply chains, we need to understand the driving factors which influence the performance of the multi-commodity supply chains. There are three logistical factors: Facilities, Inventory Policies and Transportation [16].

- **Facilities**: The number, size, location and the capacity of a facility greatly influence the logistical costs of the supply chains.

- **Inventory Policies**: The type of inventory policies adopted would in turn influence the flow of a commodity through the chain influencing its distribution costs.

- **Transportation**: The mode of transportation selected and the design of the transportation network would greatly influence the overall logistics costs of the supply chain.

It is very important to design a multi-commodity supply chain such that the location of facilities, inventory policies and transportation routes selected, minimize the operating costs of the supply chain. To optimize such a multi-commodity supply chain, coordination between the decisions among every activity of the supply chain should be established. Thus it is very important to have an efficient supply chain management strategy. The decisions considered during supply chain management can be categorized into three main decision types and are described in Table 1.
<table>
<thead>
<tr>
<th>Decision Type</th>
<th>Time Horizon</th>
<th>Decision</th>
<th>Problem Formulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategic Decision</td>
<td>3 – 5 years</td>
<td>Number of biorefineries and collection facilities to open, Locations of biorefineries and corresponding capacities, etc.</td>
<td>Facility Location Problem</td>
</tr>
<tr>
<td>Tactical Decision</td>
<td>3 – 6 months</td>
<td>Assign biorefineries to blending facilities, how much biomass should be inventoried, where should biomass be stored, etc.</td>
<td>Customer – Facility Allocation Problem, Inventory Management Problem</td>
</tr>
<tr>
<td>Operational Decision</td>
<td>Weekly or Daily</td>
<td>Biofuel production in a given period, shipment quantities between facilities, amount inventoried, etc.</td>
<td>Production Scheduling Problem, Distribution Scheduling Problem</td>
</tr>
</tbody>
</table>

Our model identifies locations and capacities for facilities making the capacitated facility location problem (CFLP) a sub-problem of our main problem. The model also designs a network for flow of biomass making the network design problem a sub-problem of our model. Finally, the model also identifies a production plan and a distribution schedule for biofuels making the production planning problem (PPP) a sub-problem for our model. Overall, we have to design a distribution network for multiple commodities (biomass) making our problem similar to a multi-commodity network flow problem (MCNFP). The relation of our problem to all the above mentioned problems is discussed in detail in further chapters.
Research Motivations

Increasing energy needs and a fossil fuel shortage crisis have increased people’s awareness about the need for alternative sources of energy. Energy Policy Act of 2005 signed by President Bush encourages the production of energy from alternative sources, especially renewable sources like solar energy, tidal energy etc., by providing incentives like loans, subsidies, and tax cuts for those who are willing to start a business. The Act states that the amount of biofuel that is blended with gasoline should be increased to 7.5 billion gallons per year (bgy) in 2012 from 4.0 bgy in 2006 [17]. In 2007, President Bush signed in to law the Energy Independence and Security Act of 2007 which states that the amount of biofuels blended with gasoline be increased to 36 billion gallons by year 2022, and out of 36 billion gallons, 21 billion gallons must be derived from non-cornstarch products [18].

Renewable energy sources consist of tidal, solar, wind and bioenergy. All the above mentioned energy sources can provide energy but are either not as feasible or not efficient options to be considered for transportation fuels as biofuels. Globalization has increased transportation and therefore, the need for fuel to drive our vehicles. Thus biofuels are emerging as one of the most important sources of energy in today’s world. Among biofuels, ethanol is considered to have similar properties to gasoline. A lot of research has been done for improving the efficiency of producing ethanol from corn [19, 20]. A number of researchers are investigating other biomass feedstock sources, such as forest residues and agricultural, municipal, industrial and other types of wastes which are rich in sugars [21, 22]. Processes are being improved, and new technologies are being developed to improve the efficiency of biofuels’ production. However, biofuels’ has not
been able to replace gasoline in the markets significantly. In the USA, only 3% of the gasoline demand is satisfied by ethanol [23]. This is due to the fact that biofuels’ mileage efficiency is less than that obtained from gasoline and the higher costs of producing biofuels. The higher cost of producing biofuels is due to the high logistics costs of supplying biomass to conversion facilities. Around 35-50% of biofuels production costs are feedstock costs, out of which 50-75% of costs are related to the transportation of feedstock from harvesting sites to the conversion facilities [24, 25]. Thus we can say that there is room for improvement in the existing supply chains for biomass. Thus to make biofuels a more viable option as compared to gasoline a lot of effort has to be given to design efficient supply chains so as to lower the costs of producing biofuels. Figure 1 [26] depicts a typical supply chain for ethanol (from corn).

Progress has been made in developing efficient supply chains for industrial products but not much work has been done in designing efficient supply chains in the agricultural field [27, 28]. Further, there are differences in agricultural products and industrial products which limit the use of industrial supply chains to be applied to the agriculture sector. The agricultural products have some distinguished characteristics which hinder the application of industrial supply chains to agricultural products [28]. These characteristics are:

- not all crops can be grown in all places
- amount harvested is restricted by growing process
- yields and amount harvested depends on weather conditions and insect population
- limited availability of land use
harvesting techniques and regional climatic changes influence the quality of produce

Figure 1  Ethanol Supply Chain Network

Biomass feedstock can be referred to as any biological material which can be used for producing biofuels. Biomass feedstock includes plants such as micanthus, switch grass, corn, willow, sugar cane, etc. In addition, biomass can also include biodegradable wastes such as cow dung, wastes from pig and poultry slurry, municipal and industrial wastes etc. According to Browne et al. [29] biomass can mainly be categorized into four main groups:
Forest residue (wood)
Crop residue or Agricultural waste (stover, straw)
Energy crops (micanthus, switch grass)
Municipal and industrial waste

We are interested in developing efficient supply chains for the first three categories of biomass as listed above. Since biomass from the first three categories is either related to agricultural products or forest products, supply chain models built for these products can help us better understand and design supply chain models for biomass feedstocks of interest.

**Research Objectives and Contributions**

Research in the field of supply chain optimization can be divided into two main categories: 1. Supply Chain Design and 2. Supply Chain Management. Supply chain design is concerned with designing a supply chain for a product or a category of products. Supply chain management is related to managing the already established supply chain. Supply chain design involves strategic decisions like locating a facility whereas supply chain management involves operational and tactical decisions like production and distribution planning.

A supply chain can be categorized as either a single commodity supply chain or a multi-commodity supply chain. Single commodity supply chain means there is a flow of just a single product or service through the supply chain and multi-commodity supply chain means more than one commodity flows through the supply chain simultaneously. In multi-commodity supply chain, commodities use the same limited resources. If the
commodities were not to share the same resources, the multi-commodity problem can be easily decomposed into a number of single-commodity problems. However, due to the fact that they share the same resources, the problem cannot be decomposed and should be solved for all commodities together. Therefore, we can divide supply chain optimization into four basic categories: 1. Single-Commodity Supply Chain Design, 2. Single-Commodity Supply Chain Management, 3. Multi-Commodity Supply Chain Design, 4. Multi-Commodity Supply Chain Management. Figure 2 shows the characterization of the present literature available in the field of supply chain optimization.

Supply Chain optimization is a vast field and a lot of research has been done in this field. We can categorize the present research in the field of supply chain optimization into the above mentioned four categories. Much research has been done in categories 1 and 2; and as required by changing times, researchers are now focusing on doing research in categories 3 and 4. Most of the times, the problems in the last two categories are NP-complete or NP-hard, and so basic optimization solvers are not efficient in solving the problems to optimality. Therefore, research has been concentrated in designing solution algorithms for solving these problems to optimality or close to optimality. Although research has been done in categories 3 and 4, not much research has been done at the intersection of category 3 and 4, i.e. not much research has been done for solving multi-commodity supply chain design and management problems. Table 2 shows the sample list of authors and the field of their work for the above four categories.
In short, our area of interest is the intersection of supply chain design and supply chain management problems with multi-commodities, as indicated by the arrow in Figure 2. The problem at hand is a part of the area of interest as shown by a black circle in Figure 2. Our main goal is to develop efficient biomass-to-biorefinery chains, which is a small part of the area of our interest. Thus, we intend to use the knowledge base of existing supply chain models and solution procedures developed for the industrial products to develop a multi-commodity supply chain design and management model for biomass (agricultural products) and solve the model using heuristic procedures. The heuristic procedure developed can also be used to solve the multi-commodity supply chain design and management problems for the industrial products.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Single Commodity</th>
<th>Multi Commodity</th>
<th>Supply Chain Design</th>
<th>Supply Chain Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>S Talluri [30, 31]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. M. Beamon [32]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. C. Cooper et al. [33, 34]</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. M. Lambert et al. [35-37]</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. J. Thomas &amp; P. M. Griffin [38]</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. J. Chen &amp; A. Paulraj [39-41]</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. Min et al. [42]</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. T. Melo et al. [43]</td>
<td>○</td>
<td>○</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohen &amp; Lee [44, 45]</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. C. Arntzen et al. [46]</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V. Jayaraman &amp; H. Pirkul [4]</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. Pirkul &amp; V. Jayaraman [5, 6]</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
</tbody>
</table>
We propose an optimization model for the biomass-to-biorefinery supply chain that considers harvesting sites, collection facilities, biorefineries and blending facilities. The model coordinates location, transportation, production and inventory decisions of the supply chain for a fixed time horizon, $T$. The model minimizes the total delivery costs for biofuels by minimizing fixed costs for investments and variable costs like harvesting, processing, transporting and inventory holding costs. We formulate the problem using a mixed integer programming (MIP) model and solve it using CPLEX 9.0, a commercial
MIP solver. CPLEX 9.0 goes out of memory for larger instances of our problem. Thus, we develop a Lagrangean Decomposition technique to solve our model to near optimality. The problem formulation and solution procedures are discussed in details in further chapters.

The deliverables of our research include an efficient biomass-to-biorefinery supply chain model which takes the whole system’s view in consideration. The model would give us the locations of the biorefineries and collection facilities to be located and the distribution network for biomass and biofuels. The model would also indicate the type and amount of biomass to be harvested for satisfying the demand for biofuels. Some of the questions our model would be answering are:

Questions related to long term decisions:

- How many biorefineries should be opened and what should be their capacities, based on the demand?
- Where should these biorefineries be located?
- How many collection facilities should be opened and what should be their capacities?

Questions related to mid-term and short term decisions:

- What should be the distribution network for biofuels?
- What should be the distribution network for biomass?
- How much inventory should we carry for biomass and where should we store it?

In addition, our research would also provide efficient heuristic solution procedures for solving the multi-commodity supply chain design and management problems which are very common in reality. Most of the MCNFP in present times deal
with either multi-commodities with single pair of origin-destination with known demand for each commodity or multi-commodities with multiple origins and multiple destinations with known demands for each commodity, where commodities cannot be substituted for each other. The problem we are looking at is a MCNFP with multiple origin and multiple destinations for each commodity. The commodities are secondary commodities which are used to produce a primary commodity with known demand. Thus as secondary commodities can be substituted for each other; their demand is not fixed and will be determined by the problem itself.
CHAPTER II
LITERATURE REVIEW

Supply chain is a system whose constituent parts include material suppliers, production facilities, distribution services and customers linked together via the feed-forward flow of material and the feedback flow of information [47]. Supply chain is a system in which we have a flow of products/services in a series from a source (for e.g., a manufacturer) to the destination (for e.g., a customer) via some intermediate points (like warehouses and distribution centers). One can think of a supply chain as a network whose nodes represent customers, retailers, distributors, manufacturers and suppliers and arcs represent flow (transportation) of products between these nodes. The main objective when managing a supply chain is to satisfy the demands of the customers and at the same time maximize the profitability obtained from the revenues generated by satisfying customer demands [16].

A wide variety of literature is available in the field of supply chains. In our literature review we have identified two broad categories: 1) literature related to Supply Chain Design and 2) literature related to Supply Chain Management. Within each one of these categories we have further identified two sub-categories, i.e., applications in industry and agriculture, as supply chains for these products is different. For example, managing the supply chain for agricultural products poses additional challenges due to
certain characteristics such as seasonality, deterioration rate, etc. Finally we review literature specifically related to supply chains for biomass.

Supply Chain Design

“Supply chain design involves the determination of how to structure a supply chain. Design decisions include the selection of partners, the location and capacity of warehouse and production facilities, the products, the modes of transportation, and supporting information systems” [48]. We reviewed many articles related to supply chain design and summarized them into two sub-categories as described below:

Supply Chain Design for Industrial Products

A lot of work has been done to develop supply chains for industrial products. Daskin et al. [49] showed the importance of facility location decisions for supply chain designs. Researchers have developed various types of supply chains based on different types of products. Reiner and Trcka [50] studied a product-specific supply chain for a food industry and showed that supply chain design should be product or company specific. They developed an improvement model to enhance the performance of a specific supply chain. Researchers have also designed multi-echelon supply chains, Hinojosa et al. [51] studied a two-echelon multi-commodity capacitated facility location problem. The objective was to minimize total costs including transportation costs, inventory holding costs and fixed and operating costs for the facilities. They developed a Lagrangean relaxation for the problem and developed a feasible solution for the overall problem based on the solutions of the sub-problems. Qi and Shen [52] studied a 3-tier
supply chain model with one supplier and one or more facilities and retailers. A
Lagrangean relaxation model is developed which in turn is used to develop a solution
algorithm to obtain good quality solutions to the problem. There are models that consider
uncertainty while designing a supply chain. Guillen et al. [53] studied the problem of
supply chain design under uncertainty. They develop a multi-objective model which
evaluates performance of a supply chain based on profit maximization, demand
satisfaction and reducing financial risk. A set of optimal solutions is then identified based
on the choice of the decision maker. Santoso et al. [54] developed a stochastic
programming model along with a solution algorithm to solve realistic supply chain
network design problems under uncertainty. Their solution algorithm includes a sampling
average approximation (SAA) scheme and integrates it with the Benders decomposition
to compute high quality solutions. Qi and Shen [52] embedded supply uncertainties and
its impact on supply chain decisions for a 3-tier supply chain model.

To summarize, we can say that a lot of different aspects of the products have been
considered while designing supply chains for industrial products. Also, various
uncertainties have been included in the models which make the models resemble real
world problems, and different techniques have been developed which help us obtain high
quality solutions.

Supply Chain Design for Agricultural Products

Lowe and Preckel [55] provide a brief review of the existing literature in the field
of agribusiness. They mainly reviewed articles related to decision technologies and
supply chains in agribusiness and restricted their research to the food and agribusiness
sectors. In their review, the authors talk about the need of redesigning existing transportation and distribution systems and inventory management systems in the agriculture sector, due to challenges brought up by the increase in technological advancements in the field of agriculture.

We reviewed the location models available in the literature for the agricultural products. Lucas and Chhajed [56] did a review of several articles related to the location analysis problems of 1950’s and 1960’s in agriculture. Most common problems were location-allocation problems, and often routing decisions were added to those problems to create integrated location-allocation-routing problems. Their review covered location problems to locate processing and storage facilities for every field of agriculture; grain industry, fruit and vegetable processing, beef industry, dairy industry etc. Monterosso et al. [57] dealt with plant size-location problem for grain storages in developing countries. They formulated the problem as the capacitated network flow problem and used the out-of-kilter algorithm (OKA) and equilibrated subroutines iteratively to find out the minimum cost network flows. Their results showed that it was better to locate storage with comparatively less capacity near the farms than to locate storage with large capacity far from the farms. A common misconception existed was to build larger storage facilities to take advantage of economies of scale, but transport costs were not considered in that decision. The authors, with their model showed that a reduction in transportation costs exceeded an increase in storage costs by locating a number of smaller sized facilities, thus making it an efficient solution. Hilger et al. [58] developed a mixed integer programming model to find the optimal number of grain terminals to be located in the northwest region of Indiana. They analyzed 19 sites and used Benders Decomposition to
solve the model. It was concluded that having a sub-terminal at any location would help reduce the annual costs rather than not having a sub-terminal; construction of sub-terminals would help in better management of inventories reducing the storage costs, in addition to capacitated country elevators. For optimal cost savings the location of sub-terminals is important, because locations of sub-terminals would dictate transportation costs and thus for lower transportation costs, the sub-terminals would need to be located optimally. Von Oppen and Scott [59] developed an spatial equilibrium model for locating processing plants. The model has two parts, a linear programming model for determining plant size and location and a quadratic programming model for determining the optimal quantities of regional supply of raw materials to be processed and regional demands of the processed products. The two models work iteratively until the optimum solution is reached. Fuller et al. [60] developed a mixed integer programming model to reduce the supply chain costs by determining which existing plants to activate for the season. As the model was too complex to solve by any existing solution softwares, the problem was reformulated as the minimum cost network flow problem. The problem is converted to a minimum cost network flow problem by developing a node-arc structure. The nodes of the structure are production locations, weekly processing at plants, plant locations etc. The arcs are comprised of production flows, marginal costs, etc. The solution procedure of the problem involved fixing the binary variable to an arbitrary value and solving the problem, but this would increase the computational complexity because we have to solve the problem for all the combinations. Thus, an implicit enumeration procedure was applied and results were obtained. Implicit enumeration is a technique in which only those combinations which improve previously enumerated combinations are examined.
For further details please refer to [60]. The results showed that 8 plants out of an existing 14 should be deactivated. Bornstein and de Castro Villela [61] did a study in the southern part of Brazil about the political and economic aspects which influence the warehouse locating problems in developing countries like Brazil and also discussed the computational limitations for formulating such an NP-complete problem. Ladd and Lifferth [62] used a transshipment plant location model to determine the number, size and location of new sub-terminals (sub-terminals are similar to grain elevators but with larger storage capacities, and they usually receive shipments from nearby grain elevators), expansions in the storage capacity of the existing grain elevators and determine the monthly grain flows from source to destinations and the rail network to be used. The main objective of the study was to maximize the profits of corn and soybean producers from their sales. They divided their solution approach into two parts, Part I determined the flow of grains from an origin to the destination for given locations of elevators and sub-terminals with a rail line network such that the revenue is maximized. Part II determined locations of elevators and sub-terminals with a rail network for which the overall net revenue was maximized (subtracting fixed and variable costs from revenue calculated in part I).

The models studied in this section were mostly facility location problems for a variety of agricultural products. Most of the models determined the location and size of storage and processing facilities based on the transportation costs for shipping products to and from facilities. Hilger et al. [58] developed a model which encompassed storage options and time dimension for inventory considerations in determining the location of the facilities. Thus we can say that most of the models studied presented a facility
location problem considering only the transportation distances and costs, but very few models integrated the facility location decisions with inventory management which would be an essential factor in designing efficient supply chains. Also, most of the models considered the facilities as the destination nodes in the network which might not always be the case. In our case; we consider the facilities as intermediate or transshipment nodes. Thus, we can say that existing supply chains in the agriculture sector are not optimized and there is room for more research to make agricultural supply chains more efficient. Two important conclusions can be taken from this review into our model building; the capacities of the facilities should be small but more in numbers and the location of facilities is an important aspect of the supply chain.

Supply Chain Management

Lambert et al. [36] defined supply chain management as “the integration of key business process from end-user through original suppliers that provides the product, service, and information that add value for customers and other stakeholders.” Thus to say in simple words, supply chain management is the management of the flow of information, funds or commodities from the supply point to the demand point in the supply chain.

In today’s world of globalization, it is becoming very important for the manufacturers to provide customers with the goods they want, at low prices with higher service levels. An increase in product variety has also increased the complexity of supply chain design and management. Manufacturers are in constant search for ways of reducing their operating costs. One of the ways is to coordinate decisions about different functions
within the company and within the corresponding supply chain. For this purpose, companies can use multi-commodity supply chain optimization models which combine facility location decisions with production and inventory decisions for multiple products. These models help to reduce the costs of managing the supply chain.

Supply chain coordination has gained importance in recent times due to increasing competition among the firms to achieve utmost level of efficiency of the system in order to minimize the total costs and increase the firms’ profits. Thus it is very important that the flow of information is maintained within the supply chain so as to achieve the required coordination among the supply chain activities. In our review of supply chain management literature, we have again divided it in two categories: literature related to supply chain management for industrial products and agricultural products.

Supply Chain Management for Industrial Products

Li et al. [63] identified five dimensions of supply chain management practices and showed that companies that used supply chain management tools had a competitive advantage and better organizational performance. Sezen [64] did a statistical study about the relative effects of integration, information sharing and supply chain design on the supply chain performance and found out that supply chain design has significant effect on the resource and output performance of supply chain. Effects of integration and information sharing are less significant as compared to the effect of supply chain design. According to Sezen, this is because the supply chains (mostly manufacturing industries) that are already utilizing their resources efficiently may not consider supply chain integration and information sharing as important as other supply chains consider it to be.
Meixell and Gargeya [65] reviewed the present literature on global supply chain and assessed the fit between the existing literature and the practical issues of global supply chain design. They concluded that although a number of models resolved the issue of globalization, only a few models were able to address the practical global supply chain problem entirely. Chen and Paulraj [39] developed a critical framework for understanding supply chain management. They carried out in depth research in the field of supply chain. In their analysis, they studied a number of research papers and found out that the current research in the area of supply chain management is concentrated on one of the aspects of the supply chain, for example, logistics and transportation, purchasing or vendor-buyer relationship. Some people have considered multiple aspects but focused on the performance of only the focal firm, like a supplier or a retailer. In fact, the relative importance and interrelationships of various aspects of supply chain and its effect on overall supply chain performance have not been explored very well. Thus, more needs to be done to fill the gap that exists about models that integrate supply chain design and supply chain management (SCM) decisions.

Few efforts have been made to close the gap that exists, to integrate supply chain design and management decisions. Manzini et al. [66] provided a conceptual framework for solving Production Distribution Logistic System Design (PDSD) which would integrate strategic decisions (supply chain design) with tactical and operational decisions (SCM) for effective supply chain management. Campbell and Sankaran [67] developed a framework for supply chain integration. The framework, known as SCIEF (Supply Chain Integration Enhancement Framework) was developed for enhancing the participation of small and medium enterprises in the supply chain. Felix and Chan [68] also present a
review of the existing literature in the field of supply chain and found out that most of the research was focused on supply chain segments like production planning, inventory management, warehouse management, etc. rather than solving supply chain problems at the systems’ level considering all the segments together at a time. They also propose a multi-agent system approach to solve the problems related to the supply chain network taking a holistic approach. Thus from these articles it can be seen that the intersection area of the supply chain design and supply chain management is now being explored by researchers.

Supply Chain Management for Agricultural Products

Supply chain management for agricultural products is mainly related to three aspects of agriculture; farm planning, crop rotation/mix and transportation. A few researchers have also considered the problem of managing the whole supply chain for a few agricultural products. Thus, our literature review on supply chain management for agricultural products is divided into four sub-categories and is discussed in details below.

Farm Planning

Farm planning can be considered as the first stage of any agricultural supply chain which determines the amount of product to be harvested at any given time. Recio et al. [69] developed a decision support system, which they called AgriSupport II, using a mixed integer programming (MIP) model to minimize the total costs of harvesting. The outputs for the model were schedules for field tasks. Biswas and Pal [70] used the Fuzzy Goal Programming method for land use planning to optimize the production of seasonal
crops. Vitoriano et al. [71] used integer programming models, one which assumes that the time horizon is divided into discrete time periods and another which consists of continuous time variable, to determine schedules of carrying tasks with minimum cost based on precedence among tasks, resource availability and time constraints for performing tasks. Jiao et al. [72] developed a linear programming model to maximize the sugar content of cane (CCS) by developing a harvesting schedule for the farms in the sugar mill region in Australia. Abdulkadri and Ajibefun [73] developed a linear programming model with an objective of maximizing gross margins of farm profits subjected to land, labor and operating expenses constraints, for Ondo state in Nigeria. The model generated plans for crop plantations. Ferrer et al. [74] used a mixed integer linear programming (MILP) model to generate optimal schedules for harvesting wine grapes taking into consideration both the operational costs as well as grape quality. Caixeta-Filho [75] developed a linear programming model to generate optimal harvesting schedules for maximizing the profits and optimizing the quality of orange juice obtained by placing chemical, biological and transportation constraints. Jones et al. [76] developed a two stage dynamic linear programming model for Syngenta, Inc., a company producing seed-corn, that manages their seed-corn production and matches seed-corn demands. The dynamic programming model helped Syngenta Inc. to increase land utilization, improve their inventory problems, and increase their revenue margins. Zuo et al. [77] developed a mathematical programming model for managing the production planning of a seed corn production company. The model was used to allocate products to facilities and transport products from production facilities to customer demand points. The model was similar to a linear programming model but with the additional constraint which suggested that the
production can either be zero or a very large number (it is called either-or constraint). The model was modified and a heuristic was developed to solve the modified model. Sensitivity analysis was performed to provide insights of the system’s performance and to help managers at the company to select the best course needed.

These models help us understand how the land utilization, optimal use of machinery, resource allocations, harvesting schedules etc. influence production output. These models give us insights about how to distribute available land among different biomass, when to harvest a biomass to obtain maximum yield, and how to efficiently allocate harvesting machineries among different biomass types. These will help us develop efficient supply chains where a continuous biomass flow is maintained for the conversion facilities.

**Crop Rotation Mix**

This sub-category reviews models which were developed for achieving cost minimization or profit maximization by rotating crops during plantation over a certain time period. El-Nazer and McCarl [78] developed a linear programming model for designing optimal long run crop rotation policies with the objective of maximizing profits. Perry *et al.* [79] developed a MIP for identifying optimal participation in governmental programs and crop mix with an objective of maximizing the net present value (NPV) for the present and future returns from the crops. The participation in government programs has many options and influences the crop mix decisions, thus the model developed would help to optimize the crop mix strategies. Ekman [80] developed a discrete stochastic sequential programming (DSSP) model for optimizing the
combination of crop mix together with machinery chosen under the effects of weather variability.

The above mentioned models help us understand the requirements needed to improve the biomass supply chains where a constant supply of biomass will be maintained to the conversion facilities. By using crop rotation policies biomass yield can be improved and a steady flow of biomass can be maintained even under specific harvesting periods and weather uncertainties.

**Transportation Models**

There have been very few transportation models developed for agricultural products. Caixeta-Filho [75] in his linear programming model put an additional transportation constraint for shipping oranges from harvesting sites to the processing plants. Ladd and Lifferth [62] designed a transshipment problem in which they were able to determine the flow of grains from a source to a destination for given locations of country elevators and sub-terminals with its rail network such that it maximized the revenues by minimizing the transportation and grain handling costs. These models show the importance of having transportation constraints in our biomass models when optimizing its supply chain. These models will lead to reducing the transportation costs within the chain. D’Souza [81] studied the structure of the soybean processing industry in the US in 1990. His model was a simple linear programming model with the objective of minimizing the processing and transportation costs. He considered the supply of soybeans as a parameter and also had the locations of processing units as inputs in his model.
Supply Chain Management in Agricultural Fields (Holistic Approach)

Higgins et al. [82] reviewed the literature present in the value chain in the sugar industry and concluded that most of the research is focused on the logistics aspect of the chain. The value chain in this industry lacks research related to integration and coordination of activities among different members of the supply chain. The concept of the value chains is similar to that of a supply chain, however, unlike supply chains, the value chains capture the value added to the product at each stage of the chain. Higgins et al. also mentioned that although sizeable research is being done in sugar value chains (or say in agriculture), unlike manufacturing chains, its practicality has been limited. Apaiah and Hendrix [83] developed a simple linear programming model for the supply chain of novel protein foods (NPFs). The supply chain network consisted of locations for growing and harvesting peas, pea processing and markets as nodes and transportation arcs for pea and pea products among these nodes for Dutch markets. The objective of their linear programming model was to minimizing the total manufacturing costs for NPFs while finding the optimal locations for the facilities and modes of transportation. Gigler et al. [84] used a dynamic programming (DP) approach for optimizing supply chains for agricultural products. They divided supply chains into two categories: Agri chains, which are designed for agricultural products whose quality changes as it moves ahead in the chain, and Non-agri chains, which are designed for other products, whose quality doesn’t change as it moves ahead in supply chain. The authors successfully applied their DP approach to a case study for developing the supply chain for willow biomass to an energy plant. However, this study only incorporated a 2-stage supply chain consisting of farms and energy plants. Widodo et al. [28] developed a mathematical model to reduce the loss
in mass of fresh agricultural products and to maximize the demand satisfaction level. They developed the maturing curve and loss function in order to capture the plant growing process and deterioration of the fresh product. Their model considered one farmland where the product was grown and harvested, and one retailer where the product was shipped to and customer demand was satisfied.

These models help us understand the product requirements when designing the supply chains for biomass. Note that these models mostly analyze the logistics aspects of the supply chain, which deal with transportation of the products. The existing literature does not consider the coordination among strategic and tactical decisions of the supply chain. Coordination among these decisions is important for efficient performance of the supply chains.

Vast amounts of literature are available in the field of supply chain designs and supply chain management for industrial products. The literature deals with coordinating decisions to optimize supply chains. These decisions are location of a plant and its distribution network, inventory management and transportation decisions, production planning and distribution decisions ([7, 85-88]). Thomas and Griffin [38] did a review of available literature in the field of supply chain coordination in which models presenting two types of planning were reviewed: operational planning and strategic planning.

A detailed literature review was performed in order to see if similar research has been done in the areas of supply chain designs and management for agricultural products (biomass). Studies have been done in developing harvesting schedules for biomass, studying effects of crop rotation/mix, inventory management of biomass, transportation
issues of biomass, and locations of biomass processing facilities, but very few researchers have taken a holistic approach which would incorporate all of these factors together [89].

**Supply Chains for Biomass**

The recent energy crises have been the motivation for research in the field of energy production from renewable sources. Biomass is a key factor for energy production; therefore, many researchers have started exploring this new field of interest. There are two main streams of research in the field of biomass: biomass-to-biofuel conversion processes and transportation and logistics of biomass. Although both of the categories are important, we concentrate on the second category because it is relevant to the supply chain design of biomass. Researchers have developed various models for minimizing biomass transportation and logistics costs; comparing various transportation modes for biomass, developing logistics chains for biomass delivery to the conversion plants, integrating harvesting schedules with biomass transportation, integrating biomass storage and transportation, *etc.* In the remaining section we review articles related to the biomass logistics.

Many researchers have used simulations as an efficient tool for analyzing the existing logistics network for the biomass. Nilsson [90, 91] developed SHAM (Straw Handling Model). The model assumes fixed conversion facility locations and is used to determine the locations of storage sites and handling operations for straw. They incorporated weather uncertainties as well as yield uncertainties in their model. The incorporation of weather and yield uncertainties makes the simulation model more dynamic and close to real-world scenarios, because these uncertainties are always present
in the real world and influence straw harvesting and handling operations. The objective of the model was to analyze various delivery systems to improve systems performance and reduce costs and energy needs for straw handling. Sokhansanj et al. [92] developed IBSAL (Integrated Biomass Supply Analysis and Logistics). The objective was to simulate the collection, storage and transportation of large quantities of biomass to predict the delivery costs. Mol et al. [93] combined simulation with mathematical programming. They developed a simulation model called BioLogiCS (BIOmass LOGIstics Computer Simulation) to calculate costs and energy consumption for the biomass logistics. The input to BioLogiCS was a biomass supply network which was obtained by solving the MIP model.

Researchers have also explored other methods for designing and managing efficient biomass supply chains. Gigler et al. [94] studied a few supply chain strategies for supplying willow to the conversion facilities and compared them to determine the minimum cost supply chain strategy. Another such study was done by Brown et al. [29]. They developed five supply chain strategies, mainly based on the method of harvesting, for four different kinds of biomass and then compared them to determine the best supply chain strategy. In their study they concluded that delivery costs of biomass with intermediate storage were higher as compared to biomass with onsite storage. This is because for supply chains with intermediate storage, biomass is to be transported first to the storage facilities from farms/forests and then from storage facilities to the plants incurring the transportation costs twice, whereas for supply chains with onsite storage this transportation cost is incurred only once (to transport biomass from farms/forests to plants). Forsberg [95] used the Life Cycle Inventory (LCI) method to investigate the
environmental impacts of different bioenergy supply chains. Forsberg considered five types of supply chains for just one type of biomass (forest residues) and compared the supply chains based on their environmental impact. Forsberg found out in his study that the environmental impacts of shipping biomass to a conversion facility in a different country are in no means different than shipping the biomass to a local conversion facility, i.e., the environmental impact of a biomass supply chain is the same irrespective of the distance.

Petrolia [96], in his study for the state of Minnesota, developed cost functions to estimate transportation costs of corn stover based on distance traveled. Sokhansanj et al. [97] developed a stochastic model capturing variations in crop yield, bale density and other various factors. The model was then used to estimate the transportation costs for transporting corn stover to the intermediate storage facilities. Tatsiopoulos and Tolis [98] studied different logistic networks for the cotton-stalk biomass. They proposed a linear programming model with the objective function of minimizing the total costs subject to the flow conservation constraints for biomass. Cundiff et al. [99] developed a linear programming model for designing a herbaceous biomass delivery system. The goal of the model was to optimize the schedules for shipping biomass from the producer to the plant while minimizing the overall costs including transportation and storage costs for biomass. Weather uncertainties play an important role in determining the production quantities and harvesting processes for the agricultural products. Weather uncertainties were incorporated in the model by introducing weather-related factors affecting production. Two such factors were introduced, one related to weather during the growing season and another related to weather during the harvesting period. Both the factors were classified
as either “good” or “poor”, thus four different weather scenarios were developed by a combination of the above two factors. Probabilities were assigned to each of the possible four scenarios. A two-stage linear programming model with recourse was developed to solve the problem. Jenkins and Arthur [100] used network analysis and dynamic programming techniques to select optimal handling and transportation methods for biomass (rice straw). Tembo et al. [89] considered an integrated approach for designing the biomass supply chain. They developed a MIP model and solved it using commercial software. The results from the model were the location and size of the biorefinery based on the biomass supply available in that region.

The models reviewed in this section dealt with biomass and biofuels. Most of the model studies considered only tactical and operational decisions [90, 91]. Some of the models considered strategic decisions like facility locations but did not coordinate with the tactical and operational decisions [90, 92]. Most of the models assumed that the conversion plant was centrally located and then evaluated different transportation modes or handling machinery to use so as to reduce logistics costs. Other models ([29, 93, 94, 99]), compared different types of supply chains for different biomass and suggested the best supply chain to be used for a particular biomass. Thus, none of these models studied above considered the holistic approach of designing and efficiently managing the biomass-to-biorefinery chain, integrating each and every component of the supply chain like, harvesting, storage, transportation, location of conversion plants, etc. reducing the overall system’s costs. The approach by Tembo et al. [89] can be by far considered as the closest to achieve this efficiency. Our model can be considered as an extension of their
model, in which we include the decisions like determining location and size of collection facilities and biofuel distribution to markets.
Our problem is a supply chain design and management problem. We develop (design) an efficient biomass supply chain and optimize (manage) the supply chain over a time horizon. The main objective of our supply chain is to produce and distribute biofuels to the blending facilities in an efficient way by reducing the biofuels’ production and distribution costs.

**Problem Description**

The biomass-to-biorefinery supply chain consists of four echelons: 1. Harvesting Sites, 2. Collection Facilities, 3. Biorefineries, and 4. Blending Points. Biomass is produced at the harvesting sites and is harvested in specific time periods which depends on the type of biomass. Biomass is then transported to collection facilities where it is stored and supplied to biorefineries. Biorefineries process biomass and convert it to biofuels which are then supplied to blending facilities to be mixed with petroleum fuel for production of E-10 and E-85. These mixtures are currently used as gas by vehicles. Biomass can be stored at harvesting sites, collection facilities or biorefineries. Biofuels can be stored at biorefineries. The main objective of the research is to optimize collection, distribution and processing of biomass and biofuels.
Biofuels can be produced by using more than one type of biomass feedstock. Different feedstocks can only be harvested in a specific time period over a time horizon. The deterioration of biomass should also be taken into consideration since biomass losses matter with time. Some of the costs associated with biomass are harvesting costs, transportation costs, inventory holding costs and biomass-to-biofuels conversion costs. Costs associated with biofuels include transportation costs and inventory holding costs. Investment costs for opening collection facilities and biorefineries should be considered as well. We took all these factors into consideration and developed a supply chain model to optimize the biomass and biofuel flow through the supply chain.

Our models help managers and decision makers with the following decisions: identifying biorefinery and collection facility locations, selecting a harvesting site and identifying type and amount of biomass collected from that harvesting site in a particular time period, identifying type and amount of biomass processed at a biorefinery in a given time period, identifying distribution network for biomass and biofuels, and identifying inventory levels for biomass and biofuels.

**Problem Formulation**

We formulate the problem as a mixed integer linear programming (MILP) problem. The objective function for our MILP is to minimize the overall delivery costs for biofuels. The problem parameters, decision variables, objective function and constraints are defined as follows:
Problem Parameters

$t$ index of time period. $T$ denotes the length of the planning horizon (year)

$p_b$ unit price for planting, growing and harvesting biomass type $b$ ($/ton)

$h_b$ unit inventory holding cost for biomass type $b$ ($/tons/year)

$h_e$ unit inventory holding cost for biofuels ($/gallons/year)

$c_{b k j}^1$ cost of transporting one unit biomass type $b$ from harvesting site $k$ to collection facility $j$ ($/tons)

$c_{b j i}^2$ cost of transporting one unit biomass type $b$ from collection facility $j$ to biorefinery $i$ ($/tons)

$c_{i m}^3$ cost of transporting one unit of biofuels from biorefinery $i$ to market $m$ ($/gallon)

$\omega_b$ unit cost for processing biomass type $b$ ($/tons)

$\psi_{id}^l$ amortized fixed investment cost for a biorefinery of size $l$ at a location $i$ ($$/year)

$\psi_{jn}^i$ amortized fixed investment cost for a collection facility of size $n$ at a location $j$ ($$/year)

$\alpha^1$ deterioration rate for outdoor storage of biomass (e.g., at a harvesting site) (%)

$\alpha^2$ deterioration rate for indoor storage (e.g., at collection facility & biorefinery) of biomass (%)

$\beta_b$ conversion rate of biomass type $b$ to biofuels (gallons/tons)

$S_{nCF}^1$ storage capacity of a collection facility of size $n$ (tons/month)

$S_{iBR}^2$ storage capacity of a biorefinery of size $l$ (tons/month)

$d_{m t}$ demand for biofuels at market $m$ in time $t$ (gallons)
\( C_i \) \hspace{1em} \text{production capacity of a biorefinery of size } l \text{ (gallons/month)}

\( z_{kbt} \) \hspace{1em} \text{amount of biomass type } b \text{ available at site } k \text{ in time } t \text{ (tons)}

**Decision Variables**

\( \phi_{kbt} \) \hspace{1em} \text{amount of biomass type } b \text{ harvested at site } k \text{ in time period } t \text{ (tons)}

\( y_{kjbt}^1 \) \hspace{1em} \text{amount of biomass type } b \text{ shipped from harvesting site } k \text{ to collection facility } j \text{ in time period } t \text{ (tons)}

\( y_{jibt}^2 \) \hspace{1em} \text{amount of biomass type } b \text{ shipped from collection facility } j \text{ to biorefinery } i \text{ in time period } t \text{ (tons)}

\( y_{limt}^3 \) \hspace{1em} \text{amount of biofuels shipped from biorefinery } i \text{ to market } m \text{ in time period } t \text{ (gallons)}

\( z_{kbt}^1 \) \hspace{1em} \text{amount of biomass type } b \text{ stored at harvesting site } k \text{ in time period } t \text{ (tons)}

\( z_{jbt}^2 \) \hspace{1em} \text{amount of biomass type } b \text{ stored at collection facility } j \text{ in time period } t \text{ (tons)}

\( z_{jibt}^3 \) \hspace{1em} \text{amount of biomass type } b \text{ stored at biorefinery } i \text{ in time period } t \text{ (tons)}

\( z_{it}^4 \) \hspace{1em} \text{amount of biofuels stored at biorefinery } i \text{ in time period } t \text{ (gallons)}

\( w_{ibt} \) \hspace{1em} \text{amount of biomass type } b \text{ processed in biorefinery } i \text{ in time period } t \text{ (tons)}

\( e_{it} \) \hspace{1em} \text{amount of biofuels produced in biorefinery } i \text{ in time period } t \text{ (gallons)}

\( x_{il} \) \hspace{1em} \text{binary variable equals 1 if a biorefinery of size } l \text{ is opened at a location } i, 0 \text{ otherwise}
binary variable equals 1 if a collection facility of size \( n \) is opened at location \( j \), 0 otherwise

MIP Model

\[
\begin{align*}
\min_{P} & \quad p_{b} \phi_{kbt} + h_{b} \{ z_{kbt}^1, z_{jbt}^2, z_{ibt}^3 \} \\
\text{subject to:} & \\
& \phi_{kbt} + \xi_{kbt} \quad \forall k \in \{1, \ldots, K\}, b \in \{1, \ldots, B\}, t \in \{1, \ldots, T\} (2) \\
& \phi_{kbt} (1 - \alpha^1) z_{kbt}^1 \quad \forall k \in \{1, \ldots, K\}, b \in \{1, \ldots, B\}, t \in \{1, \ldots, T\} (3) \\
& y_{kht}^1 (1 - \alpha^2) z_{kbt}^2 \quad \forall k \in \{1, \ldots, K\}, b \in \{1, \ldots, B\}, t \in \{1, \ldots, T\} (4) \\
& y_{ibt}^3 (1 - \alpha^2) z_{ibt}^3 \quad \forall i \in \{1, \ldots, I\}, b \in \{1, \ldots, B\}, t \in \{1, \ldots, T\} (5) \\
& e_{it}^4 \beta_{bt} w_{ibt} \quad \forall i \in \{1, \ldots, I\}, t \in \{1, \ldots, T\} (6) \\
& e_{it}^4 z_{ibt}^4 \quad \forall i \in \{1, \ldots, I\}, t \in \{1, \ldots, T\} (7) \\
& z_{ibt}^2 \quad \forall i \in \{1, \ldots, I\}, t \in \{1, \ldots, T\} (8)
\end{align*}
\]
\]
\begin{align*}
B & \quad \sum_{b=1}^{B} z_{ibt} \quad \sum_{l=1}^{L} S_{lbh} x_{il} \quad \forall i \quad 1, \ldots, I, t \quad 1, \ldots, T \quad (9) \\
I & \quad \sum_{i=1}^{I} y_{imt} \quad d_{mt} \quad \forall m \quad 1, \ldots, M, t \quad 1, \ldots, T \quad (10) \\
1 & \quad \sum_{i=1}^{I} e_{il} \quad C_{l} x_{il} \quad \forall i \quad 1, \ldots, I, t \quad 1, \ldots, T \quad (11) \\
1 & \quad \sum_{i=1}^{I} x_{il} \quad 1 \quad \forall i \quad 1, \ldots, I \quad (12) \\
N & \quad \sum_{j=1}^{N} x_{jn} \quad 1 \quad \forall j \quad 1, \ldots, J \quad (13) \\
& \quad \sum_{i=1}^{I} \sum_{b=1}^{B} \sum_{k=1}^{K_{b}} z_{kib0} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{b=1}^{B} z_{jib0} \sum_{i=1}^{I} \sum_{b=1}^{B} z_{ib0} \sum_{i=1}^{I} x_{ib0} \sum_{i=1}^{I} x_{ib3} \sum_{i=1}^{I} x_{ib4} \sum_{i=1}^{I} 0 \quad \forall k \quad 1, \ldots, K, j \quad 1, \ldots, J, \\
& \quad 1, \ldots, I, b \quad 1, \ldots, B \quad (14) \\
\phi, z, y, \alpha, e & \geq 0 \quad (15) \\
x_{il} & \in \{0,1\}; \quad x_{jn}^t \in \{0,1\} \quad (16) \\
\end{align*}

where: \( \xi_{kibt} \quad \theta_{bt} \delta_{kb} \gamma_{kb} L_{kb} \quad \forall k \quad 1, \ldots, K, b \quad 1, \ldots, B, t \quad 1, \ldots, T \quad (17) \)

- \( L_{kb} \): total available land at harvesting site \( k \) for biomass type \( b \) (acres)
- \( \gamma_{kb} \): proportion of land that can be harvested at site \( k \) for producing biofuels from biomass type \( b \) (%)
- \( \delta_{kb} \): yield for biomass type \( b \) at harvesting site \( k \) (tons/acre)
- \( \theta_{bt} \): harvesting-time factor (captures seasonality for biomass type \( b \))
\[ \theta_{bt} \begin{cases} f_{bt} & \text{if } f_{bt} > 0 \text{ when biomass type } b \text{ harvested in period } t \\ 0 & \text{otherwise} \end{cases} \]

\( \xi_{bkt} \) represents the amount of biomass type \( b \) available at a particular harvesting site \( k \) in time period \( t \). For example suppose we have 1,000 acres of land available (\( L_{kb} \ 1000 \)), for producing corn stover, as shown in Figure 3 below. Due to erosion constraints, it is advised that no more than 33\% of stover be removed from land. Thus, out of 1000 acres, we can use only 33\% (\( \gamma_{kb} \ 0.33 \)) for harvesting biomass for biofuels’ production, i.e., biomass harvested from 330 acres can be used for producing biofuels. Similar considerations are taken into account when calculating availability of other biomass types, like in the case of woody biomass, the percentage of woody biomass used for biofuels’ production depends on its use by paper and furniture industries, more the paper and furniture industries use, the less percentage of woody biomass is available for biofuels’ production and vice versa. Also, certain biomass can be harvested during certain months only, for e.g., corn stover can be harvested in only 3 months out of 12, i.e., September, October and November. Therefore, if we assume an equal amount of proportions to be harvested in the harvesting months then we would harvest 33.33\% (\( \theta_{bt} \ 0.33 \)) of total corn stover in each of September, October and November, and for rest of the months, we would not harvest (\( \theta_{bt} \ 0 \)). Thus for the months of September, October and November, amount of land harvested is 110 acres and no land is harvested in the remaining months. Now if yield is 3 tons/acre (\( \delta_{kb} \ 3 \)), then in each of the 3 months we would get 330 tons (\( \xi_{bkt} = 330 \)) of corn stover. Therefore, a maximum of 990 tons of corn stover is available for producing biofuels in a year.
Constraint (2) indicates that the amount of biomass for biofuels in a given period is limited by biomass availability. Constraints (3), (4) and (5) are flow conservation constraints which indicate that no more biomass can be shipped from a site than what is available at that site in a given period. Constraint (6) indicates that the amount of biofuels produced is limited by the amount of biomass processed. Constraint (7) shows that the amount of biofuels shipped from the biorefinery is not bigger than what is available. Constraint (8) and (9) are the capacity constraints on the storage of biomass in a given period at collection facilities and the biorefineries respectively. Constraint (10) indicates that the demand at each blending facility should be satisfied. Constraint (11) is the biofuels production capacity constraint in a biorefinery. Constraint (12) and (13) are the
location constraints that say that only one biorefinery or collection facility can be opened at a given location. Constraint (14) says that the initial inventory for biomass and biofuels is zero. Constraint (15) is non-negativity constraints. Constraint (16) is binary constraints.

**Related Problems**

Our problem is a combination of three well studied combinatorial problems. These problems are: capacitated facility location problem (CFLP), network design problem (NDP) and production planning problem (PPP). The formulation of these problems and their relation to our problem is explained in detail in the following paragraphs.

**Capacitated Facility Location Problem**

CFLP identifies locations for facilities. These facilities have capacities. Our problem formulation incorporates facility location decision variables. Collection facilities and biorefineries have capacities. Thus, CFLP is a sub-problem of our main problem. The general formulation for CFLP is as shown.

\[
\begin{align*}
\text{minimize} & \quad \sum_{i} \sum_{j} c_{ij} x_{ij} + \sum_{i} \sum_{j} g_{ij} x_{ij} \\
\text{subject to:} & \\
\sum_{j} p_{ij} x_{ij} & \leq V_i y_i \quad \forall i, 1, \ldots, I \\
\sum_{i} x_{ij} & = 1 \quad \forall j, 1, \ldots, J
\end{align*}
\]  

(17)
\( x_{ij}, y_i \in \{0, 1\}, i \in I, j \in J \) \tag{20}

\[
x_{ij} = \begin{cases} 
1 & \text{if client } j \text{ is serviced by facility } i \\
0 & \text{otherwise}
\end{cases}
\]

\[
y_i = \begin{cases} 
1 & \text{if facility } i \text{ is opened} \\
0 & \text{otherwise}
\end{cases}
\]

\( I \) is the set of potential facilities to open, \( J \) is the set of customers, \( c_i \) is the cost of opening a facility at location \( i \), \( g_{ij} \) is the unit production and transportation costs, \( p_{ij} \) is the amount produced at facility \( i \) for customer \( j \), and \( V_i \) is the production capacity at facility \( i \). Please refer to [101] for more details about this formulation.

In the model we propose above, constraints (18) are represented in our model by constraints (10) and (11). Investment costs \( c_i \) and transportation costs \( g_{ij} \) are represented by \( \psi_{it} \) and \( c_{im}^3 \) respectively in our model. Production quantity \( p_{ij} \) is represented by \( y_{imt} \) in our model.

CFLP can be considered as a special case of our problem when it satisfies the following properties:

1. One time period (i.e. \( t = 1 \)),
2. One commodity (i.e. \( b = 1 \)),
3. One size of biorefinery (i.e. \( l = 1 \))

Therefore, CFLP is a special case of our problem. If we just consider location of biorefinery based on biofuels demand, and if we assume sufficient supply of biomass and
infinite storage capacity for biomass along with one time period \((t = 1)\), one biomass type \((b = 1)\) and one biorefinery size \((l = 1)\) then constraints (2)-(6), (8), (9), and (12)-(14) are redundant. If \(d_m \leq 1\) and all costs are assumed to be zero except \(c_{im}\) and \(\psi_i\) in the objective function then our model will look like:

\[
\begin{align*}
[P^1] & \quad \min \left[ I \sum_{i=1}^{I} \psi_i x_i + M \sum_{i=1}^{I} \sum_{m=1}^{M} c_{im} y_{im}^3 \right] \\
\text{subject to:} & \quad e_i \leq C x_i, \quad \forall i, 1, \ldots, I \quad (22) \\
& \quad e_i \leq \sum_{m=1}^{M} y_{im}^3, \quad \forall i, 1, \ldots, I \quad (23) \\
& \quad \sum_{i=1}^{I} y_{im}^3 \leq 1, \quad \forall m, 1, \ldots, M \quad (24) \\
& \quad y_{im}^3, e_i \geq 0, \quad \forall i, 1, \ldots, I, \forall m, 1, \ldots, M \quad (25) \\
& \quad x_i \in \{0,1\}; \quad (26)
\end{align*}
\]

The objective function of our model will be similar to the CFLP and constraints (7) and (11) from our model represent constraints (18) for CFPL, constraints (10) and (16) in our model represent constraints (19) and (20) for CFLP respectively. Thus CFLP is a sub-problem to our model. CFLP is a NP-Hard problem. For more details please refer to [102].
Multi-Commodity Network Flow Problem

Our model also develops a distribution network for biomass flow which makes multi-commodity network flow problem (MCNFP) a special case of our problem. The general formulation for MCNFP is as shown:

\[
\text{minimize} \quad \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k
\]
subject to:

\[
x_{ij}^k - x_{ji}^k \leq b(i) \quad \forall i \in N
\]

\[
\sum_{(i,j) \in A} x_{ij}^k \leq u_{ij} \quad \forall (i,j) \in A
\]

where \( \sum_{i \in 1}^n b(i) = 0 \)

Constraint (28) are known as flow conservation constraints, first term in the constraints represents the outflow of the node and the second term represents the inflow for the node. These constraints state that (outflow - inflow) should equal supply/demand of that node [103]. Constraints (29) are the capacity constraints for the arcs in the network. Finally, the term \( \sum_{i \in 1}^n b(i) = 0 \) tells us that total supply should equal total demand.

In our model, we can find similar flow conservation constraints in constraints (3), (4), and (5) for biomass and constraints (7) for biofuels.

MCNFP can be considered as a special case of our model and a sub-set for our model when the following properties are satisfied:

1. One collection facility and biorefinery size \( l = 1 \)

2. Outdoor and Indoor deterioration rate is zero (i.e. \( \alpha_1^1 \alpha_2^2 = 0 \) )
3. No inventories \( (z_{kbi}^1, z_{jbi}^2, z_{ibi}^3, z_{ubi}^4, 0) \)

MCNFP is a special case our main problem. If we assume that we know the locations of the harvesting sites, collection facilities, biorefineries and markets, and if we assume an abundant supply of biomass and infinite storage capacity for biomass along with one size \( (l = 1) \) each for collection facility and biorefinery, and with zero deterioration rates \( (\alpha^1, \alpha^2, 0) \) and no inventories are considered \( (z_{kbi}^1, z_{jbi}^2, z_{ibi}^3, z_{ubi}^4, 0) \) then constraints (2), (8), (9), (11)-(14), and (16) are redundant.

We can have two network flow problems within one problem, one for biomass and one for biofuel. If all costs are assumed to be zero except \( c_{bki}^1, c_{bji}^2, c_{imi}^3 \) in the objective function then our model will look like:

\[
[P^2] \quad \min \left[ \sum_{k=1}^{K} \sum_{j=1}^{J} c_{bki}^1 y_{kjb}^1 + \sum_{j=1}^{J} \sum_{i=1}^{I} c_{bji}^2 y_{jib}^2 + \sum_{i=1}^{I} \sum_{m=1}^{M} c_{imi}^3 y_{im}^3 \right]
\]

subject to:

\[
\phi_{kb} \quad y_{kjb}^1 \quad \forall k \quad 1, \ldots, K, \forall b \quad 1, \ldots, B \quad (31)
\]

\[
y_{kjb}^1 \quad y_{jib}^2 \quad \forall j \quad 1, \ldots, J, \forall b \quad 1, \ldots, B \quad (32)
\]

\[
y_{jib}^2 \quad w_{ib} \quad \forall i \quad 1, \ldots, I, \forall b \quad 1, \ldots, B \quad (33)
\]

\[
e_{i} \quad \beta_{b} w_{ib} \quad \forall i \quad 1, \ldots, I \quad (34)
\]
\begin{equation}
e_i \sum_{m=1}^{M} y_{im}^3 \quad \forall i \ 1,\ldots, I \tag{35}
\end{equation}

\begin{equation}
y_{im} \quad \forall m \ 1,\ldots,M \tag{36}
\end{equation}

\begin{equation}
\phi_{kb}, y_{kib}^1, y_{kib}^2, w_{ib}, e_i, y_{im}^3 \geq 0 \quad \forall k \ 1,\ldots,K, \forall j \ 1,\ldots,J,
\end{equation}

\begin{equation}
\forall i \ 1,\ldots,I, \forall m \ 1,\ldots,M \tag{37}
\end{equation}

For a biomass network, constraints (3)-(5) in our model represent constraints (28) for the MCNFP. If we assume infinite capacity for the arcs in our model then, constraints (15) in our model represent constraints (29) for MCNFP. For a biofuel network, constraints (7) in our model represent constraints (28) for the MCNFP. If we assume infinite capacity for the arcs in our model then, constraints (15) in our model represent constraints (29) for MCNFP. Therefore, MCNFP is a special case of our model.

**Production Planning Problem**

Finally, our model solves a production planning problem (PPP) for producing biofuels and their distribution. The general formulation for PPP is as shown.

\begin{equation}
\text{minimize} \sum_{i=1}^{n} (p_i(x_i) + h_i(I_i)) \tag{38}
\end{equation}

subject to:

\begin{equation}
I_j - X_i - R_i \quad \forall i \ (1,\ldots,n) \tag{39}
\end{equation}

\begin{equation}
0 \quad x_i - c_i \quad \forall i \ (1,\ldots,n) \tag{40}
\end{equation}

\begin{equation}
I_i \geq 0 \quad \forall i \ (1,\ldots,n) \tag{41}
\end{equation}

\begin{equation}
I_n \quad 0 \tag{42}
\end{equation}
where \( b_i \) is the production set-up cost, \( r_i \) is the demand, and \( c_i \) is the production capacity for \( i = 1, 2, ..., n \).

\[ p_i(x) \] \( b_i \) \( p_i(x) \), is the cost of producing \( x_i \) units in period \( i \).

\( I_i \), is the inventory level in period \( i \),

\[ X_i \]

\( X_i \), is the cumulative production level up to period \( i \),

\[ R_i \]

\( R_i \), is the cumulative demand level up to period \( i \),

Now when we consider our model, constraints (7) and (10) represent constraint (39) for PPP. Also, constraints (40) of PPP are represented by constraints (11) in our model.

PPP can be considered as a special case of our model and a sub-set for our model when the following properties are satisfied:

1. One biomass type (\( b = 1 \))
2. One Location Facility (i.e. \( i = 1 \))
3. One Size of Facility (i.e. \( l = 1 \))
4. Unified biofuel demand (\( m = 1 \))
5. Investment costs for the facility are zero (\( \psi_{il} = 0 \))

PPP is a special case of our main problem. If we just consider production of biofuel at the biorefinery and its demand, and if we assume an abundant supply of biomass and an infinite storage capacity for biomass along with one biomass type (\( b = 1 \)) and one location (\( i = 1 \)) and size (\( l = 1 \)) of the facility and a unified demand of blending...
facility \((m = 1)\) then constraints (2)-(6), (8), (9), (12)-(13) and (16) are redundant. If all the costs are assumed to be zero except \(h^e\) and \(c^3\) then our model will be as follows:

\[
[P^3] \quad \min \left[ \sum_{t=1}^{T} c^3 y^3_t \quad \sum_{t=1}^{T} h^e z^4_t \right]
\]

subject to:

\[
e_t, z^4_{t-1}, d_t, z^4_t \quad \forall t \quad 1, \ldots, T
\]

\[
e_t, C \quad \forall t \quad 1, \ldots T
\]

\[
z^4_t \leq 0 \quad \forall t \quad 0
\]

\[
z^4_t, y^3_t, e_t \geq 0 \quad \forall t \quad 1, \ldots T
\]

In our model, constraint (10) becomes \(y^3_t \leq d_t, \forall t \quad 1, \ldots, T\), which becomes a redundant constraint and therefore we replaced \(y^3_t\) by \(d_t\) in constraint (7) above so that our model resembles the PPP model.

If we assume zero production set-up costs \((b_i)\) then the first part of our objective function \(\left( \sum_{t=1}^{T} c^3 y^3_t \right)\) represents total production costs and the second part \(\left( \sum_{t=1}^{T} h^e z^4_t \right)\) represents total storage costs respectively for the general production planning problem.

Constraints (7) and (10) of our model represent constraints (39) for PPP, constraints (11) and (15) of our model represent constraints (40) for PPP, constraints (15) of our model also represent constraints (41) for PPP and finally constraints (42) for PPP are taken care of by constraints (7) and (10) in our model. Thus, PPP is a sub-problem to our model.
It can be summarized that our problem is a special case of three problems (CFLP, MCNFP and PPP). All three problems have been proven to be difficult problems to be solved by exact methods. For more details about the computational complexity of the above mentioned problems please refer to [102-105]. Thus, heuristics procedures have been employed to solve these problems to near optimality. Our problem, being a special case of all three problems, is no exception and therefore, heuristic procedures are explored for solving our problem to near optimality. The computational complexity of our problem and the solution procedures are discussed in details in further chapters.
CHAPTER IV
SOLUTION PROCEDURES

In this chapter, we will discuss solution procedures used for solving our problem. As described in the previous chapter, our problem is a difficult problem to solve by using exact solution methods and thus heuristic methods are suggested. In this chapter we discuss the computational complexity of our problem, present a review of the literature for the existing heuristic approaches for solving similar problems, and finally we present a Lagrangean Decomposition approach for solving our problem.

Computational Complexity of the Problem
As described in the previous chapter, the CFLP, MCNFP and PPP are special cases of our problem. CFLP and PPP are proven NP-Hard problems [102, 105]. Therefore, the computational complexity of our problem is NP-Hard. As the problem is difficult to solve computationally, exact methods cannot be used effectively and therefore we need to look for heuristic approaches to solve the problem in an efficient way.

Review of Solution Procedures
We reviewed articles which developed exact as well as heuristic procedures to solve deterministic models similar to our problem. Our literature review of solution
procedures is divided into three main categories. Category 1 includes articles which used exact mathematical programming techniques, category 2 includes articles which designed heuristic algorithms and category 3 includes articles which used decomposition methods for solving the problem. In each category we review articles related to all the three different kinds of problems (CFLP, MCNFP and PPP) on which our model is based.

Exact Mathematical Programming Techniques

In this section, we reviewed methods that use exact techniques to get to the solution of the problem. The first paragraph deals with articles that used exact methods for CFLP. Consecutive paragraphs review articles that use exact methods for MCNFP and PPP. One of the many exact methods is the Branch-&-Bound method, which is used to solve mixed integer programming problems like facility location problems. Branch-&-Bound method is an enumeration method where enumeration (solution space) is restricted by upper and lower bounding the quantity to be optimized and therefore reducing the number of enumerations needed. At each iteration an integer variable is bounded, forcing it to yield an integer solution. This is repeated until all the integer variables yield an integer solution which is within the upper and lower bounds for the optimal solution. Soland [106] proposed a Branch-&-Bound algorithm for locating facilities with concave costs, choosing from a pool of potential facilities in order to satisfy customer demands at minimum costs. Laporte et al. [107] reported a branch and cut method for solving a capacitated facility location problem with stochastic demands. Tcha and Lee [108] developed a Branch-&-Bound procedure for an uncapacitated facility location problem which opened facilities at multiple levels. They used a dual ascent procedure (DAP) and
primal descent procedure (PDP) along with a node simplification procedure (NSP) to significantly reduce the branching in their Branch-&-Bound algorithm. Klose and Görtz [109] applied a column generation algorithm within the branch-and-price algorithm to solve the capacitated facility location problem. Few articles discuss a Branch-&-Bound method based on Lagrangean heuristics for solving different types of facility location problems. These methods are discussed in the later sections where we talk about the decomposition techniques. Most of the exact methods for solving facility location problems involved some modification of the Branch-&-Bound algorithm.

Next, we discuss articles which solve MCNFP with the use of exact methods. Tomlin [110] formulated the minimum cost MCNFP using two different formulations, Node-Arc formulation and Arc-Chain formulation. Tomlin showed that due to their special structures, the former can be tackled by the Dantzig-Wolfe Decomposition and the latter can be solved using A Shortest Chain Algorithm. The author showed that although both the formulations were differently formulated, essentially both were equivalent. The latter having an advantage of being capable of dealing with both directed as well as undirected arcs. Rabinowitz and Mehrez [111] developed a nonlinear cost minimization formulation for Dead Sea Works Ltd. The model dealt with a multi-echelon multi-commodity logistic system and was specially formulated for the company. The model was solved using Excel and GAMS and sensitivity analysis was performed. Ford and Fulkerson in [112] proposed a simplex computation for the maximal MCNFP. The problem was formulated as an arc-chain formulation. This method reduced the size of the bases matrix as compared to a simplex method. The method worked fine with the small size problems but was not tested for the larger size problems and so its practicability was
not proved. Foulds [113] formulated a multi-commodity flow problem as a network
design problem. He used Branch-&-Bound technique to obtain the solutions. In [114] the
authors suggest two easy improvements for obtaining better and efficient lower bounds to
the Branch-&-Bound algorithm used by Warszawski [115] for the multi-commodity
location problem. Gregoriadis and White [116] formulated the routing of freight cars as a
MCNFP and developed a partitioning algorithm based on the primal partitioning
algorithm. In this method, the sub-problems of dual are solved and feasible solutions are
checked for optimality, if the solution is not optimal then a base change is made. This is
done for a finite number of steps. They compared their results to the linear programming
solutions (obtained using MPS/360 software) and found their algorithm to be efficient in
both computation time and number of iterations required for solving the problem. Lin and
Lin [117] used a well-known projected Jacobi method to solve the nonlinear MCNFP.
The method combined a new dual projected pseudo-quasi-Newton (DPPQN) method
with the projected Jacobi method to solve the quadratic sub-problems induced from the
latter method. The DPPQN method unlike conventional Lagrangean Newton method,
finds a constant sparse approximate Hessian matrix which makes the algorithm more
efficient than the latter method. McBride [118] compared four different methods used for
solving MCNFP. These methods are, (i) Decomposition techniques, (ii) Interior-Point
methods, (iii) Simplex with Advance Basis, and (iv) Base-Partition method (EMNET).
The techniques were compared on the PDS (Patient-Distribution System) and KEN
(Kennington Test Problems) problems from NETLIB [119]. It was concluded that the
Base-Partitioning methods had an advantage over the other three techniques. Awerbuch
et al. [120] present approximate solution algorithms for two types of multi-commodity
flow problems: Maximum concurrent flow and Maximum-benefit flow. Their algorithms are based on natural approximate steepest descent framework and are efficient in both the distributed and parallel environment.

We reviewed articles developing exact methods for PPP. Wagner and Whitin [121] proposed a dynamic programming algorithm to solve a dynamic lot sizing problem. In the dynamic lot sizing problem, an order quantity is to be determined such that it minimizes the sum of set-up and inventory costs when demand for each time period is known. Wagner and Whitin proposed an algorithm for the case when we have different demand amounts for different periods and inventory and production costs vary with time. The Wagner-Whitin [121] algorithm starts from the end of time horizon and works its way up to the first period. In each period, $t$, the algorithm decides whether to produce in period $t$ or not, and how much to produce. The algorithm runs in $O(n^2)$, where $n$ is the number of periods. Florian and Klien [122] studied a multi-period single commodity production planning problem similar to Wagner and Whitin. Unlike Wagner and Whitin, Florian and Klien introduced production capacities to their problem. They assume production and storage costs to be concave. They used a dynamic programming approach similar to Wagner and Whitin and showed that their approach solves the problem to optimality if production capacities are equal for all periods. Ahuja and Hochbaum [123] studied the capacitated dynamic lot sizing problem with linear production costs, i.e., zero setup costs. They showed how a successive shortest path algorithm for minimum cost flow problems can be used to solve these problems in $O(n^2)$ time. Further, they showed how these problems can be solved in $O(n \log n)$ time with the use of dynamic trees. $n$ is the number of periods.
Heuristic Algorithms

In this section we review articles related to the heuristic procedures developed for solving CFLP, MCNFP and PPP. We start our review with articles related to CFLP. Korupolu et al. [124] studied a local search heuristic for CFLP along with uncapacitated facility location problem (UFLP) and k-median problem and proved that constant factor approximation bounds can be obtained in polynomial time for uncapacitated/capacitated k-median problems and uncapacitated/capacitated facility location problems. The local heuristics starts with a feasible solution, i.e., a set of opened facilities and then uses ADD, DROP and SWAP algorithms in the local search, to add a facility to the set of open facilities, to drop a facility from a set of open facilities and to swap a facility of the same size from a set of open facilities respectively, reducing the costs significantly. Arya et al. [125] worked with the same local search heuristic as Korupolu et al. [124], but they added a new operation where they were able to drop more than one facility at a time. To keep the procedure within the polynomial time bounds, they used a procedure called T-hunt for this operation. T-hunt is a procedure where in a series of knapsack problems is solved to add or remove a facility from a sub-set to determine multiple sub-sets of facilities opened. This procedure reduced the locality gap to between 3 and 4 as compared to 5 obtained by Korupolu et al. [124]. Jain et al. [126] designed a greedy algorithm for finding good approximation bounds for UFLP and use that algorithm to find better approximation algorithms for the CFLP. They used a LP dual-fitting technique which is similar to a primal-dual algorithm, but in which inequalities are just relaxed in the dual giving a feasible solution for the primal and an infeasible solution for the dual problem.

57
Heuristic procedures for MCNFP are discussed in this section. Khuller et al. [127] developed an approximate algorithm to solve multi-commodity network design problems. They concentrate on the case when the given network is a tree. They develop a divide-and-conquer algorithm based on Leighton and Rao’s algorithm [128], which is a balanced-separator algorithm to divide the graph and obtain an approximation factor of $O(\log n)$. Awerbuch and Leighton [129] proposed a local-control algorithm for transporting multiple commodities from their source to sink nodes in a dynamically changing distributed network where the capacities vary with time. Their algorithm was based on the “edge-balancing” technique in which a commodity is sent across the edge, $e = (u, v)$, if there are more commodities at $u$ than there are at $v$. The algorithm moves in rounds, where each round has 4 phases, i.e., adding new flow to source, pushing flow across edges, removing flow from sink, and rebalancing nodes. Schneur and Orlin [130] developed a penalty-function for the minimum cost multi-commodity flow problem by relaxing the capacity bundle constraints and adding them to the objective function. They developed a scaling algorithm for solving this problem. The scaling algorithm solves a sequence of penalty problems by sending flow around negative costs cycles and makes use of a penalty parameter ($\rho$) and units of commodity ($\delta$) for determining the solution quality. The scaling algorithm utilizes the problem’s special network structure and finds approximately optimal solutions. Agarwal [131] modeled a telecommunication network design problem as a multi-commodity network design problem. The algorithm is similar to the neighborhood search; it starts with an initial feasible solution and at each step tries to improve the solution. At each iteration, a subset of links called sub-network is selected and sub-problems are created, keeping the rest of the network unchanged. These sub-
problems are formulated as multiple choice knapsack problems and are solved using a
dynamic programming algorithm. The algorithm is run until there is no improvement to
the solution. Poh et al. [132] modeled the multi-period multi-commodity transportation
problem as a MIP problem. They showed that remodeling the problem into two separate
models by backward decomposition and then solving them iteratively improved the
solution quality and runtime for the problem. For solving large size problems they
developed a heuristic method which is based on heuristics used for bin-packing problems
and other search heuristics, because of the similarity of the first sub-problem to the bin-
packing problem. Chauhan et al. [133] model the forest supply chain as a MCNFP. The
problem is modeled as an MIP and is solved using commercial software CPLEX. A
scenario improvement heuristic and a branch-and-price algorithm are proposed specific to
the model. Scenario improvement heuristic starts with a given scenario where a product-
mix is determined for each block and then improves it to reduce overall costs. Once a
scenario is fixed, then the problem becomes a linear program which is solved using a
transportation simplex algorithm. Results from all three methods were compared and
showed that the scenario improvement heuristics performed well for the small size
problems and the branch-and-price algorithm did well with the large scale problems.

Heuristic procedures developed for PPP are reviewed in this section. Baykasoglu
and Gocken [134] developed a tabu search algorithm for solving a fuzzy logic goal
programming model for aggregate production planning. Hung et al. [135] formulated a
production planning problem with setup decisions (both cost and time) as a mixed integer
programming problem. The problem is a hybrid of a multi-item capacitated lot-sizing
problem and an aggregate production planning problem. They developed three different
versions of genetic algorithms to solve the problem. Torabi et al. [136] developed a fuzzy logic approach for solving the hierarchical production planning problem. The model involves two levels of decision making, the first level involves planning for family of products and the second level involves planning for individual product items within each family of products. At the first level, an aggregate production planning model is solved using a fuzzy linear program and at the second level, it is disaggregated using another fuzzy linear programming model.

**Decomposition Methods**

In this section we reviewed relaxation and decomposition techniques developed for solving CFLP, MCNFP and PPP. We start our review with decomposition methods used for CFLP. Beasley [137] developed a Lagrangean relaxation heuristic framework for different types of facility location problems including capacitated warehouse location and capacitated warehouse location with single source constraint problems. Beasley relaxes the demand constraints and the warehouse capacity constraints and reduces the problem to a 0-1 program which determines whether to open a warehouse or not. Beasley developed a special procedure to solve this problem which made use of Lagrangean multipliers to determine whether to open a warehouse or not. Klincewicz and Luss [138] developed a Lagrangean relaxation heuristic for the capacitated facility location problem with single-source constraints. They relaxed the capacity constraints to get UFLP as a sub-problem which they solved using a dual ascent algorithm. In addition to this, they also developed an *add* heuristic, which adds a facility to the set of open facilities, one at a time, to obtain an initial feasible solution to the problem and a *final adjustment* heuristic,
which improves the customer assignments obtained by Lagrangean relaxation to improve the bounds of the Lagrangean relaxation. Mazzola and Neebe [139] developed a Lagrangean relaxation technique for a multi-product capacitated facility location problem to generate lower bounds for the problem. They decomposed the main problem into a number of UFLP and a 0-1 Knapsack problem. The UFLPs are solved using the dual-based procedure as described by Erlenkotter [140] and 0-1 Knapsack problem can be solved using dynamic programming. They also developed a Branch-&-Bound procedure using the bounds generated by the Lagrangean relaxation to obtain the upper bounds for the problem. Nauss [141] developed a Lagrangean relaxation procedure for CFLP by relaxing the demand constraints. Nauss showed that by the addition of two improvements, i.e., (i) selecting better Lagrangean multipliers and (ii) adding a constraint that makes sure that enough facilities are opened so as to satisfy cumulative customer demands, tighter bounds on Lagrangean relaxation can be obtained. The sub-problems obtained in the Lagrangean relaxation are the continuous knapsack problem and the 0-1 knapsack problem. Nauss also developed a Branch-&-Bound procedure based on the bounds obtained from the Lagrangean relaxation of the capacitated facility location problem. Holmberg et al. [142] developed a Lagrangean relaxation based Branch-&-Bound procedure for a single-sourcing capacitated facility location problem. They relaxed the single-sourcing constraints to obtain a relaxation of the problem. The sub-problems obtained were knapsack problems, one for each facility. The feasible solutions for the problem (upper bounds) are obtained by converting the problem into a matching problem and solving the problem using a repeated matching heuristic framework. The Lagrangean relaxation along with the repeated matching heuristic is then embedded in the
Branch-&-Bound framework to determine both the upper and lower bounds.

Tragantalerngsak et al. [143] proposed a Lagrangean relaxation based Branch-&-Bound technique for a capacitated facility location problem with single-source constraints where facilities are required to be opened at two-echelons.

Decomposition techniques for MCNFP are discussed in this section. The most common decomposition technique is to divide the original problem into as many smaller sub-problems as commodities. These sub-problems are identical problems and according to Ralphs and Galati [144], decomposition methods are effectively applied to such problems. Karkazis and Boffey [145] considered multi-commodity facilities location problem introduced by Warszawski and Peer in [115]. They developed two solution approaches, (i) the Dual-based approach, in which they start with the dual of the problem and then iteratively solve the sub-gradient optimization algorithm to minimize the function and (ii) Lagrangean Dual based approach with Hill Climbing (HC), in which Lagrangean multipliers are obtained using HC algorithm. HC is a technique to minimize (or maximize) a function similar to the gradient ascent but the only difference is that HC is done over a discrete space whereas the latter is done over continuous space. It was found that both the methods were computationally effective but Hill Climbing was the better of the two. Balakrishnan and Graves [146] considered minimum cost MCNFP where the total cost for each arc is a piecewise linear, concave function of the total flow on that arc. They developed a Lagrangean relaxation technique with sub-gradient optimization and dual ascent to generate lower bounds. Sub-gradient optimization is used to determine the optimal Lagrangean multiplier values for each step of the Lagrangean relaxation. Upper bounds are then generated heuristically making use of solutions from
the Lagrangean relaxation. Ibaraki and Fukushima [147] present a primal-dual proximal point algorithm for the convex MCNFP. In each iteration, the algorithm finds an approximate saddle point of the augmented Lagrangean of the problem and checks that the solutions of the sub-problems always satisfy the flow conservation constraints for all commodities. Pirkul and Jayaraman [5] develop a MIP model for a multi-commodity plant and warehouse location problem where the objective is to minimize the total transportation and distribution costs along with investment costs for operating plants and warehouses. They employ the Lagrangean relaxation technique with a sub-gradient method to generate lower bounds and a heuristic method to generate upper bounds for the problem. Holmberg [148] discusses the use of Lagrangean relaxation heuristics for two variants of the MCNFP, one with single origin and single destination for each commodity, and another with multiple origins and multiple destinations for each commodity. He concluded that Lagrangean heuristics were capable of yielding near-optimal solutions. Pirkul and Jayaraman [6] formulate a MIP for the multi-commodity, multi-plant, capacitated facility location problem. A Lagrangean relaxation and a heuristic method are developed for the problem to calculate the lower and upper bounds respectively. The heuristic method utilizes the plant and warehouse locations obtained by solving the Lagrangean relaxation sub-problems as inputs, and does the assignment of customers to the warehouse based on the ratio of non-assignment penalty cost to demand requirements. Numerical experiments are done to show the efficiency of the algorithms. Babonneau and Vial [149] proposed an implementation of analytic center cutting plane method (ACCPM) to solve nonlinear MCNFP with Lagrangean relaxation. Babonneau et al. [150] developed an efficient heuristic to solve large scale linear MCNFP. The
heuristic uses partial Lagrangean relaxation on the set of arcs which is determined by an active set strategy. The Lagrangean dual is then solved using ACCPM. Shen [7] proposed a nonlinear integer program for designing a multi-commodity supply chain and developed a Lagrangean relaxation solution algorithm. The problem is relaxed to get sub-problems separable by commodity-facility pairs. A Branch-&-Bound procedure is used to generate upper bounds for the problem. The Branch-&-Bound procedure takes Lagrangean relaxation solution as an input to find an initial feasible solution. Eksioglu et al. [151] studied an integrated transportation and production planning problem in a two stage supply chain. The problem was formulated as MCNFP with fixed charge costs and Lagrangean decomposition technique was used to calculate the lower and upper bounds for the problem. The efficiency of these bounds were tested on the randomly generated problems. Wu and Golbasi [152] developed a Lagrangean decomposition for a multi-item, multi-facility supply chain planning problem. The main problem is decomposed into a resource sub-problem and a number of product-level sub-problems. The sub-problems are solved and Lagrangean multipliers are updated using a sub-gradient algorithm. The method yielded high quality solutions.

Finally, we reviewed decomposition models for PPP. Graves [153] developed a Lagrangean relaxation model for hierarchical production planning problems. The author decomposed the problem into two sub-problems, one an aggregate planning problem and another is a disaggregation problem. Lagrangean relaxation yields lower bounds for the problem and upper bounds are calculated using a procedure to find the feasible solution of the problem at each iteration. Gupta and Maranas [154] proposed a hierarchical Lagrangean relaxation for midterm production planning problems. The main problem is
decomposed into smaller problems by relaxing constraints in three different stages. Tighter lower bounds are obtained with this procedure. Certain structure of relaxed problem is retained and embedded into an upper bound finding procedure which then finds a feasible solution for the problem.

**A Heuristic-Based Solution Approach**

MCNFP arise in a wide variety of applications like transportation, telecommunication, *etc.* and many large-scale models are formulated as MCNFP [118]. MCNFP with integral flow belongs to NP-complete class of problems [155]. Thus any exact methods and algorithms cannot be used to solve large size problems optimally. The problem has a large number of variables and constraints even for a small size problem. For example if we have $m$ commodities and $n$ arcs in a network, the corresponding formulation will have at least $(mn)$ variables and $(m + n)$ constraints.

Several approaches have been developed by the researchers in solving the multi-commodity flow problems. These approaches in general can be classified in three basic categories:

1. Price-directive decomposition

2. Resource-directive decomposition

3. Partitioning methods

Price-directive decompositions include methods like Lagrangian Relaxation, Dantzig-Wolfe decomposition, *etc.* These methods remove the bundle constraints (capacity constraints) from the constraint matrix and put them in the objective function
by applying a penalty price to them. The multi-commodity flow problem is converted into several small single-commodity network flow problem and are solved separately. The method is initiated by solving a set of minimum cost flow problems and then updating the multipliers using some algorithmic procedures. This process is done iteratively until a set of stopping conditions is fulfilled.

In a Resource-directive method, unlike using prices to decompose the problem, it allocates the joint bundle capacity of each arc to the individual commodities. We would allocate the capacity to the commodities and solve single-commodity flow problems as a set of independent single-commodity flow problems. Thus, resource directive methods can be seen as a capacity allocation problem. The method solves the problem iteratively, it initially solves a resource allocation problem and then finds a sub-gradient to determine the direction and step length which would take it to convergence.

Finally, Partitioning methods work on the basis that MCNFP are specially structured linear programs embedded with network flow problems. These methods work on the principle of spanning tree interpretation of a linear programming basis and the fact that the linear programming basis for multi-commodity flow problems contains a basis for each commodity. For details the reader is directed to Ahuja et al. [103].

Based on our literature review we found out that not many people have used partitioning methods. There are few complexities which arise when working with partitioning methods which are described by McBride [118]. Different types of heuristics have been developed for MCNFP but all those heuristics exploit a specific structure of the problem being considered and cannot be generalized for all kinds of MCNFP. Many researchers have used decomposition techniques to solve the MCNFP most of the time.
decomposing it into SCNFP. A very common decomposition method used by researchers was found to be the Lagrangean relaxation technique. Thus we have employed a Lagrangean relaxation technique for our problem.

**A Lagrangean Decomposition-Based Solution Approach**

Lagrangean relaxation (LR) is a technique in which the hard constraints are moved to the objective function in order to determine the penalty of not satisfying the constraints. Lagrangean decomposition (LD) is a Lagrangean relaxation technique in which for a given MIP problem with two matrix constraints, the relaxation problem decomposes into two sub-problems, each having one of the two matrices of the original problem as constraints. Guignard and Kim [156] defined LD and proved that the optimal values obtained by using LD techniques are better than those obtained by using LR techniques.
Figure 4 shows various decomposition approaches. LD can be applied in three different ways, i.e. structural, functional and temporal. In structural decomposition the problem is decomposed by structure. For example, the multi-commodity network flow problem can be decomposed into smaller networks by commodities. In functional decomposition the problem is decomposed based on functions. For example, a capacitated facility location problem can be decomposed into two sub-problems based on functions, facility locations can be determined in the first sub-problem and customer-facility assignments can be done in the second sub-problem. In temporal decomposition the problem is decomposed by time. For example, a multi-period scheduling problem can be decomposed by time periods. For our problem we cannot apply temporal decomposition because of the inventory arcs which carried inventory from one time.
period to another. Also functional decomposition is difficult because it would only divide
the problem into two sub-problems which would not provide good bounds, but structural
decomposition can be applied because we can obtain as many smaller sub-problems as
the number of commodities.

We apply LD heuristic to our main problem \((P)\) in such a way as to reduce the
computational burden of the problem by separating our multi-commodity problem into
several single-commodity problems and solving these problems and combining their
solutions to get the lower bound on our main problem \((P)\). We start by introducing a
variable \(a_{ibt}\), which is the amount of biomass type \(b\) processed at biorefinery, \(i\), in time
period \(t\). \(a_{ibt}\) is equal to \(w_{ibt}\), and therefore we introduce constraint (49) to our main
problem \((P)\). Note that constraint (49) is a “less than equal to” constraint although
variables \(a_{ibt}\) and \(w_{ibt}\) are equal. We do this to get better quality lower bounds on our LD
problem. For details regarding this solution approach please refer to [157, 158]. Also, the
variable \(a_{ibt}\) is bounded by the amount of biomass type \(b\) available at harvesting sites in
period \(t\). Therefore, we add an additional constraint (50) to our existing model \((P)\). The
modified problem \((P')\) model is as follows:

\[
\begin{align*}
\text{[P']} & \quad \min \left[ \sum_{k \in K} \sum_{b \in B} \sum_{t \in T} p_{bkt} \phi_{ibt} \right. \\
& \quad \quad \left. + \sum_{b \in B} \sum_{t \in T} h_b \left\{ \sum_{k \in K} z_{kbt}^1 \sum_{j \in J} z_{jbt}^2 \sum_{i \in I} z_{ibt}^3 \right\} \right. \\
& \quad \quad \left. + \sum_{i \in I} \sum_{t \in T} \sum_{k \in K} \sum_{j \in J} \sum_{b \in B} \sum_{t \in T} h^e \z_{it}^4 \left\{ \sum_{k \in K} \sum_{j \in J} \sum_{b \in B} \sum_{t \in T} \sum_{i \in I} c_{bktj} \psi_{ibtj} \psi_{ibtj} \right\} \right. \\
& \quad \quad \left. + \sum_{i \in I} \sum_{m \in M} \sum_{t \in T} \sum_{j \in J} \sum_{n \in N} c_{imn} \z_{imn} \right. \\
& \quad \quad \left. + \sum_{i \in I} \sum_{b \in B} \sum_{t \in T} \sum_{l \in L} \sum_{a_{ibt} \in A_{ibt}} \omega_{a_{ibt}} \right. \\
& \quad \quad \left. + \sum_{j \in J} \sum_{n \in N} \sum_{x_{ijn} \in X_{ijn}} \psi_{ijn} \right. \\
& \quad \quad \left. + \sum_{j \in J} \sum_{n \in N} \sum_{x_{ijn} \in X_{ijn}} \psi_{ijn} \right. \\
& \quad \quad \left. + \sum_{j \in J} \sum_{n \in N} \psi_{ijn} \right. \\
& \quad \quad \left. \right] 
\end{align*}
\]
subject to:

Constraints (2)-(16) and

\[
\begin{align*}
\alpha_{ibt} & \quad \omega_{ibt} \quad \forall i \quad 1, ..., I, b \quad 1, ..., B, t \quad 1, ..., T & (49) \\
\sum_{i} \alpha_{ibt} & \quad \phi_{ibt} \quad \forall b \quad 1, ..., B, t \quad 1, ..., T & (50)
\end{align*}
\]

Our Lagrangean decomposition model (LD) is as follows:

\[
[LD] \quad \min \begin{bmatrix}
K & B & T & K & B & T & J & B & T \\
I & B & T & I & T & K & J & B & T \\
I & L & B & T & I & M & T & I & B & T \\
I & L & I & B & T & I & M & I & T & I & B & T \\
I & L & I & L & T & J & N & T & I & L & T & I & L & T \\
\end{bmatrix}
\begin{bmatrix}
p_{b} \phi_{ibt} \\
h_{b} z_{ibt}^{1} \\
(h_{b} \lambda_{jbt}^{1}) z_{ibt}^{2} \\
(h_{b} \lambda_{ibt}^{2}) z_{ibt}^{3} \\
(h_{b} \lambda_{ibt}^{2}) z_{ibt}^{4} \\
(h_{b} \lambda_{ibt}^{2}) c_{bjt}^{1} y_{ibt}^{1} \\
(h_{b} \lambda_{ibt}^{2}) c_{bjt}^{2} y_{ibt}^{2} \\
(h_{b} \lambda_{ibt}^{2}) c_{bjt}^{3} y_{ibt}^{3} \\
(h_{b} \lambda_{ibt}^{2}) \omega_{b} w_{ibt} \\
(\psi_{il} - S_{ibr}^{2}) x_{il}^{i} \\
(\psi_{jn} - S_{nCF}^{1}) x_{jn}^{i} \\
\end{bmatrix}
\]

subject to:

Constraints (2)-(7), (10)-(16), (49), (50) and

\[
\lambda_{jbt}^{1} \geq 0 \quad \forall j \quad 1, ..., J, t \quad 1, ..., T & (52) \\
\lambda_{ibt}^{2} \geq 0 \quad \forall i \quad 1, ..., I, t \quad 1, ..., T & (53)
\]
LD can be separated into following two sub-problems.

\[
\text{[SPI]} \quad \min \left[ \sum_{k \in K, b \in B, t \in T} p_k \phi_{kbt} + \sum_{k \in K, b \in B, t \in T} h_b z_{kbt}^1 + \sum_{j \in J, b \in B, t \in T} \lambda_{jbt}^1 z_{jbt}^2 \right]
\]

subject to:

Constraints (2)-(5), (14), (15), (49), (52), and (53)

and

\[
\text{[SP2]} \quad \min \left[ \sum_{i \in I, t \in T} h^i z_{it}^4 + \sum_{i \in I, m \in M, t \in T} c_{im}^3 y_{imt}^3 + \sum_{i \in I, b \in B, t \in T} \omega_b a_{ibt} \right]
\]

subject to:

Constraints (6), (7), (10)-(16), (50), (52) and (53)

The Lagrangean dual is: \[ \max_{\lambda} LD(\lambda) \]

SPI determines the flow of biomass from harvesting sites to collection facilities and to biorefineries. SP2 determines the biofuel production and distribution from biorefineries to blending facilities. Sub-problem SPI can further be decomposed by commodity into b sub-problems (SPIb). Each SPIb is a transportation model and is a linear program which can be solved using CPLEX. SP2 is a combination of a capacitated facility location and a production planning problem which can be solved using available heuristic procedures.
The upper bounding procedure consists of making use of the binary variable solutions obtained from solving the Lagrangean sub-problems and adding them to the original problem \((P)\) as constraints making it an easy linear program which is solved using CPLEX 9.0

**Sub-Gradient Optimization**

Lagrangean multipliers \((\lambda)\) are calculated using a sub-gradient optimization method. For details regarding the sub-gradient optimization method, please refer to [159, 160]. At each iteration \(r\), \(\lambda\) s are calculated using the following equations.

\[
\begin{align*}
\lambda_{jt}^{1r} & = \lambda_{jt}^{1r-1} + \zeta^r \left( \frac{B}{b} \sum_{j=1}^{J} z_{jbt}^2 \right) \quad \forall j \in J, t \in T \\
\lambda_{il}^{2r} & = \lambda_{il}^{2r-1} + \zeta^r \left( \frac{L}{l} \sum_{i=1}^{I} z_{ibt}^3 \right) \quad \forall i \in I, t \in T
\end{align*}
\]  

(56)  

(57)

where \(\zeta^r\) is the step size at iteration \(r\) and is calculated using following equation.

\[
\zeta^r = \frac{\varepsilon^r \left( Z^* - LD(\lambda^r) \right)}{\sum_{j=1}^{J} z_{jbt}^2 \sum_{i=1}^{I} z_{ibt}^3}
\]

\[
\sum_{b=1}^{B} \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{l=1}^{L}
\]

(58)

\(Z^*\) is the upper bound on the optimal solution of problem LD, \(\varepsilon^r\) is a positive scalar, \(\varepsilon^r = 2\). Initially, at iteration \(r = 0\), \(\varepsilon^r\) is set equal to 2. The value of \(\varepsilon^r\) is halved when the solution to LD has not changed for a given number of iterations. In our algorithm, we set the limit to 20 iterations. The algorithm is terminated if one of the three conditions is satisfied, \((i)\) the percentage gap calculated between the upper bound \((UB)\) and the LD solution is less than 0.1\%, \((ii)\) the solution value for LD does not change for
30 iterations, or \((iii)\) the algorithm is run for specified number of iterations (50 in our case). The number of iterations is decided by trial and error technique, where scenarios were run for a different number of iterations, before deciding upon the specified number.
CHAPTER V
A DECISION SUPPORT SYSTEM

In this chapter we develop an Excel-based decision support system (DSS) for the biomass-to-biorefinery supply chain problems. More often the mathematical models developed by academicians are more complex and require special softwares to solve them. By simplifying the models and developing tools that use existing softwares we are trying to increase the practicality of the models developed by academicians so that they can be readily used by practitioners in the real world.

We propose an interactive software-based system intended to support decision makers for the design and management of the biomass-to-biorefinery supply chain. Visual basic for applications (VBA) in Excel is used to model the algorithms that support the findings of this DSS. The model presented coordinates the long-term type decisions of designing a supply chain and the medium to short term decisions of logistics management. This system has the ability to (a) identify locations and capacities for biorefineries given the availability of biomass, and costs; (b) estimate the minimum cost of delivering biofuels. These costs include transportation, inventory, investment and processing costs; (c) perform sensitivity analyses with respect to a number of parameters.
Algorithms for Facility Location Problems

The algorithms designed in this paper are essentially based on facility location problems (FLP). FLP have been studied for many years and therefore, a lot of literature is available in the area of facility location problems. Generally facility location problems are NP-Hard problems and therefore use of exact solution methods is limited by the size of the problem. Researchers have studied a variety of facility location problems and have developed different heuristic procedures for solving each of these problems. The following paragraph gives some insights into a few of the heuristic methods developed for different types of facility location problems.

We reviewed some of the methods used for solving facility location problems. LP-Rounding method is a method in which a linear program (LP) is solved, and its solution is modified by replacing continuous variables by integer variables ([161], [162]). In local search method, the solution usually starts with an empty set and locations are added to the solution, one at a time, until the required numbers of locations are added. The solution obtained is then analyzed further by evaluating cost implications of adding or dropping a facility ([125], [163]). Greedy heuristics are another approach for solving facility location problems. It is similar to a local search heuristic but works on some greedy rule which defines the method (for example, minimum distance between facilities) ([126], [164]). Erlenkotter [140] developed a dual-based method for solving uncapacitated facility location problems. The author showed that the method provided upper bounds if not integral solutions to the problem and later developed a Branch-&-Bound algorithm to provide optimal integral solutions. Ross and Soland [165] showed how the facility location problems can be modeled as generalized assignment problems.
and can be solved using algorithms designed for solving generalized assignment problems. Nauss [141] developed a Lagrangean relaxation technique and a Branch-&-Bound algorithm for solving capacitated facility location problems. Jain and Vazirani [166] developed a primal-dual algorithm for the facility location problem.

Most of the algorithms stated here are designed either for the uncapacitated version or the k-median version of the facility location problem. Some algorithms consider the capacitated version of the facility location problem but assume soft capacities meaning multiple copies of a facility can be opened at a location. A sample list of algorithms and problems solved by these algorithms are summarized in the following Table 3.
Table 3  Different Algorithms and Facility Location Problems

<table>
<thead>
<tr>
<th>Author</th>
<th>UFLP</th>
<th>CFLP</th>
<th>K-median</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuehn &amp; Hamburger [163]</td>
<td>x</td>
<td></td>
<td>x</td>
<td>Local Search</td>
</tr>
<tr>
<td>Arya <em>et al.</em> [125]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Local Search</td>
</tr>
<tr>
<td>Lin &amp; Vitter [161]</td>
<td>x</td>
<td></td>
<td>x</td>
<td>LP Rounding</td>
</tr>
<tr>
<td>Shmoys <em>et al.</em> [162]</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>LP Rounding</td>
</tr>
<tr>
<td>Erlenkotter [140]</td>
<td>x</td>
<td></td>
<td></td>
<td>Dual –based Ascent</td>
</tr>
<tr>
<td>Nauss [141]</td>
<td></td>
<td>x</td>
<td></td>
<td>Lagrangean Relaxation</td>
</tr>
<tr>
<td>Mahadian <em>et al.</em> [164]</td>
<td>x</td>
<td>x</td>
<td></td>
<td>Greedy Algorithm</td>
</tr>
<tr>
<td>Jain <em>et al.</em> [126]</td>
<td></td>
<td>x</td>
<td>x</td>
<td>Greedy Algorithm</td>
</tr>
<tr>
<td>Jain &amp; Vazirani [166]</td>
<td>x</td>
<td></td>
<td>x</td>
<td>Primal-Dual, Lagrangean Relaxation</td>
</tr>
<tr>
<td>Hochbaum [167]</td>
<td>x</td>
<td></td>
<td></td>
<td>Greedy Algorithm</td>
</tr>
<tr>
<td>Charikar &amp; Guha [168]</td>
<td></td>
<td>x</td>
<td></td>
<td>LP Rounding, Primal-Dual, Greedy Augmentation</td>
</tr>
<tr>
<td>Charikar <em>et al.</em> [169]</td>
<td></td>
<td>x</td>
<td></td>
<td>LP Rounding</td>
</tr>
</tbody>
</table>

Our problem consists of locating a facility of a given capacity at a single location selected from a set of potential locations. The problems consider finding the minimum distance between the facilities to determine the optimal facility location. Therefore, we consider a simple local search and greedy heuristic approach based on the weighted gravity method described in the book by Nahmias [170], to find the distances between facilities and choose the facility with the minimum distance. The assignment between the facilities is done based on the same principle, i.e. the facility with the minimum overall
distance will be assigned first and so on. The algorithm is discussed in details in the following section.

**The Modified Problem**

In chapter 3 we modeled the biomass supply chain problem as a mixed integer programming problem (MIP). The original problem \((P)\) is a computationally difficult problem to solve as CFLP is a special case of the model. CFLP is a known NP-Hard problem [102], therefore the model itself is a difficult one to solve and also requires the use of special LP/MIP commercial software like CPLEX. The new problem \((MP)\) solved here is a modification of the original problem \((P)\) and its formulation is shown below.

**Problem Parameters**

\[ p_{kb} \] unit price of planting biomass type \( b \) at the harvesting site \( k \) ($/acre)

\[ c^1_b \] transportation cost for transporting a unit biomass of type \( b \) ($/tons/mile)

\[ c^2_e \] transportation cost for transporting unit gallon of biofuel ($/gallons/mile)

\[ \psi^1_{jn} \] amortized fixed investment cost for a collection facility of size \( n \) at a location \( j \) ($/year)

\[ \psi^2_{il} \] amortized fixed investment cost for a biorefinery of size \( l \) at a location \( i \) ($/year)

\[ \gamma_{kb} \] proportion of land that can be harvested at a harvesting site \( k \) for producing ethanol from biomass type \( b \) (%)

\[ L_{kb} \] total land available at harvesting site \( k \) for biomass type \( b \) (acres)

\[ \delta_{kb} \] yield for biomass type \( b \) at harvesting site \( k \) (tons/acre)
\[ \beta_b \] conversion rate for converting biomass type \( b \) to ethanol (gallons/tons)

\[ S_{nCF} \] storage capacity of a collection facility \( j \) of size \( n \) (tons)

\[ d_m \] demand for ethanol at market \( m \) (gallons)

\[ C_l \] capacity of a biorefinery of size \( l \) (gallons/year)

**Decision Variables**

\[ \phi_{kb} \] amount of biomass type \( b \) available at harvesting site \( k \) (tons)

\[ y^1_{kjb} \] amount of biomass type \( b \) shipped from harvesting site \( k \) to collection facility \( j \) (tons)

\[ y^2_{jib} \] amount of biomass type \( b \) shipped from collection facility \( j \) to biorefinery \( i \) (tons)

\[ y^3_{im} \] amount of ethanol shipped from biorefinery \( i \) to market \( m \) (gallons)

\[ x^1_{jm} \] binary variable equals 1 if a collection facility of size \( n \) is opened at location \( j \), 0 otherwise

\[ x^2_{il} \] binary variable equals 1 if a biorefinery of size \( l \) is opened at a location \( i \), 0 otherwise

**Modified MIP Model**

\[
[MP] \quad \min \left\{ \sum_{k \in B} \sum_{b \in B} \sum_{k \in K} \sum_{b \in B} \phi_{kb} \right\} + \sum_{k \in B} \sum_{b \in B} \sum_{j \in J} c^1_{b,y^1_{kjb}} y^1_{kjb} + \sum_{j \in J} \sum_{b \in B} \sum_{i \in I} c^2_{b,y^2_{jib}} y^2_{jib} + \sum_{i \in I} \sum_{m \in M} c^3_{e,y^3_{im}} y^3_{im}
\]

\[
+ \sum_{i \in I} \sum_{j \in J} \sum_{b \in B} \sum_{j \in J} \sum_{n \in N} \sum_{b \in B} \sum_{j \in J} \omega_{b,y^2_{jib}} x^1_{jm} x^1_{jm} + \sum_{i \in I} \sum_{l \in L} \sum_{i \in I} \sum_{l \in L} \psi^1_{im} x^1_{jm} + \sum_{i \in I} \sum_{l \in L} \sum_{i \in I} \sum_{l \in L} \psi^2_{im} x^2_{il} \right\} \right]
\]
subject to:

\[ \phi_{kb} - \sum_{j=1}^{J} y_{kjb}^1 \geq 0 \quad \forall k = 1,...,K, b = 1,...,B \]  
(60)

\[ \sum_{k=1}^{K} y_{kjb}^1 - \sum_{i=1}^{I} y_{jib}^2 = 0 \quad \forall j = 1,...,I, b = 1,...,B \]  
(61)

\[ \sum_{b=1}^{B} \beta_b y_{jib}^2 - \sum_{m=1}^{M} y_{im}^3 = 0 \quad \forall i = 1,...,I \]  
(62)

\[ \sum_{b=1}^{B} y_{kjb}^1 - \sum_{n=1}^{N} S_{nCF} x_{jn}^1 = 0 \quad \forall j = 1,...,J \]  
(63)

\[ \sum_{i=1}^{I} y_{im}^3 = d_m \quad \forall m = 1,...,M \]  
(64)

\[ \sum_{m=1}^{M} y_{im}^3 - \sum_{l=1}^{L} C_{il} x_{jl}^2 = 0 \quad \forall i = 1,...,I \]  
(65)

\[ \sum_{l=1}^{L} x_{jl}^2 = 1 \quad \forall i = 1,...,I \]  
(66)

\[ \sum_{n=1}^{N} x_{jn}^1 = 1 \quad \forall j = 1,...,J \]  
(67)

\[ y, \omega \geq 0 \]  
(68)

\[ 0 \leq \phi_{kb} \leq \xi_{kb} \]  
(69)

\[ x \in \{0,1\} \]  
(70)

where: \( \xi_{kb}, \delta_{kb} y_{k} L_{kb} \)  
\( \forall k = 1,...,K, b = 1,...,B \)

**Different Supply Chain Problems**

From our interactions with the experts and practitioners from the biomass and biofuel areas, we figured out that the two most important parameters in designing and
managing the biomass-to-biofuel supply chains are: (a) location of the facilities and (b) sizes of the facilities. Based on the information from practitioners, we divided the existing biomass-to-biofuel supply chain problem into four smaller problems and designed algorithms to solve these problems. Figure 5 shows different problems solved by the DSS.

![Figure 5 Different Types of Biomass-To-Biorefinery Supply Chain Problems](image)

The problems are based on two basic problems, the first is the simple transportation problem, which is described in the book by Winston [171] and the second is the CFLP described in previous sections. Problem S1 corresponds to a simple transportation problem where we know the locations for the facilities and we also know the sizes for the facilities. Problem S2 is a capacity allocation problem where we know the location for the facility but don’t know the size of the facility. Problem 3 is a facility location problem where we know the size of the facility but we don’t know the location
of the facility. Finally problem 4 is a capacitated facility location problem, where we need to decide size as well as location for our facility.

**DSS Algorithm**

General flow of the DSS is as shown in Figure 6. The DSS starts by getting the inputs from the user. The data entered by the user is processed and stored in excel. The next step is defining the problem. Once the user determines the problem to solve, a particular algorithm is applied to solve the problem. After the problem is solved, the user is presented with the model outputs. The user then has an option either to perform sensitivity analysis based on certain parameters or look at the detailed reports.

![Flow Chart for DSS](image)

*Figure 6  Flow Chart for DSS*
Our main algorithm is *Basic Module*. It contains other smaller algorithms as listed above and will be explained in details in the following paragraphs. We call Basic Module for all the problem types as described above. Only difference is in the calling of the other algorithms with in *Basic Module*.

**BASIC MODULE:**

1. **PRE-PROCESSING**
2. **FIND – \(N_{HS}\)**
3. **FIND – \(N_{CF}\)**
4. **LOCATE – \(N_{CF}\)**
5. **FIND – \(N_{BR}\)**
6. **LOCATE – \(N_{BR}\)**
7. **ASSIGN – \(HS – CF\)**
8. **ASSIGN – \(M – BR\)**
9. **DET – \(BF – PROD\)**
10. **DET – \(BM – PROC\)**
11. **ASSIGN – \(BR – CF\)**

**PRE-PROCESSING:** This algorithm is run to determine the feasibility of the problem. It calculates the maximum amount of biomass (\(BM\)) available, calculates the maximum amount of biofuel that can be produced. Thus making sure that supply is always greater than or equal to demand, ensuring the feasibility of the problem at hand.
**FIND**: This algorithm is used to determine the sizes and number of facilities to be opened. A greedy approach is taken in determining facility sizes. Sizes that minimize the difference between total quantities demanded are chosen. For example, if we have 2 potential sizes, i.e. 15 and 20 dry tons per year, for collection facilities, and the biomass demand is 10 dry tons then a facility size handling 15 dry tons per year will be opened.

The second part after the hyphen in the notation describes the facility for which the algorithm is called for, for example, FIND – NCF determines the sizes and numbers of collection facilities (CF) required to store the biomass.

**LOCATE**: This algorithm is used to determine the locations of the facilities to be opened. A greedy approach is taken and locations which minimize the distance are chosen. For example, LOCATE – NCF determines the locations for the CFs. The algorithm is run until the required number of CFs is opened as determined by FIND – NCF.

**DET – BF – PROD**: This algorithm is used to determine the amount of biofuel produced at a biorefinery (BR). This algorithm makes use of the assignment information obtained by running ASSIGN – M – BR to determine the biofuel production values.

**DET – BM – PROC**: This algorithm is used to determine the amount and type of biomass (BM) processed. This algorithm employs a greedy procedure, which determines the ratio of biomass costs ($/dry tons) to conversion rates (gallons/dry tons), and chooses the one which minimizes this ratio. It uses the information from DET – BF – PROD to determine the amount and type of BM processed.
**ASSIGN**: This algorithm is used to do the assignment between two facilities. For example, **ASSIGN – HS – CF** is used to assign biomass (**BM**) from the harvesting site (**HS**) to **CF**.

The costs parameters like, investment costs, transportation costs, *etc.* are calculated within the **BASIC MODULE** and is represented to the user in terms of unit delivery cost ($/gallon) of biofuel. Other reports are generated to give users in-depth analysis of the system. Users are also given an option for doing sensitivity analysis of the system based on changes in different parameters like investment costs, harvesting costs, inventory holding costs, *etc.*
CHAPTER VI

EXPERIMENTAL RESULTS

Case Study

The state of Mississippi is chosen as the case study. The data for the model is collected from United States Department of Agriculture’s (USDA) affiliated websites like National Agricultural Statistics Service (NASS), Agricultural Marketing Service (AMS) and Economic Research Service (ERS). At present, for the case study our model deals with just two types of biomass: corn, stover, and forest residues, but it can be used for any type of biomass and any number of biomass. Forest residues are further divided as sawtimber and pulpwood. We use CPLEX 9.0, a commercial LP/MIP solver, to solve the problem for the case study. One of the disadvantages of solving MILP using CPLEX is that for large scale of problems, it goes out of memory. Our problem is no exception and so for larger instances of our problems CPLEX cannot be used.

Input Data

Data collection was done for running and validating the model. We assumed that biofuel to be produced is cellulosic ethanol (C-ethanol). We used mainly two types of biomass for this purpose, i.e., corn stover and forest residues (sawtimber and pulpwood).
We collected the data for each of the three entities. Apart from that, data associated with the facilities was also collected.

Corn Stover

The harvesting period for corn stover is from September till November [172]. The amount of corn stover produced in a county is equal to the amount of corn harvested, assuming the dry-weight ratio of stover to corn grain is 1:1 [96]. Thus to get data for corn stover, we collected data for corn from the National Agricultural Statistics Services’ (NASS) website [173]. Every year NASS publishes reports covering every aspect of US Agriculture. Data we needed from NASS’s website for our model was: total planted acres and yield of corn per acre. NASS publishes this data at the county level and so the smallest unit we consider in our model is a county. We assume that corn stover is collected and then baled using a large rectangular (square) baler. The bales are staged at the field edge and then loaded onto flatbeds pulled by semi-trucks to storage facilities. Semi-trucks are also used to transport bales from storage to the biorefineries. The cost per ton per loaded mile is estimated $0.195 (under the assumption that the distance traveled is less than 25 miles), is estimated $0.143 (under the assumption that the distance traveled is between 26 and 100 miles) and $0.078 (under the assumption that the distance traveled is more than 100 miles). The transportation costs were calculated on the basis that the trucks do the round trip. The cargo weight for a load is 44,736 lbs or 952 bushels (1 bu = 56 lbs) [96]. In order to calculate the distance traveled, we identified the coordinates of each supply and demand points and then calculated the geographical distance between
points. The feedstock deterioration rate (α) is estimated to be 0.5% and 0.1% per month for outdoor and indoor storage, respectively [89].

Forest Residues

The data regarding forest residues in Mississippi came from a report published by Mississippi Institute for Forest Inventory (MIFI) [174]. The report published the amount of sawtimber and pulpwood volumes available for each county. The report published so far just considers counties in southeast and southwest regions of Mississippi. The forest residues are collected year-round except in the winter months of December, January and February. The estimated total transportation costs for forest residues are $0.125 per ton per mile.

C-ethanol

The American Coalition for Ethanol Handbook reported ethanol demands for all 50 states in the US [175]. According to that report the demand for ethanol in Mississippi in 2005 was 168 million gallons per year (MGY). The transportation cost for ethanol is estimated as $0.001 per gallon per mile.

Facilities

MIFI published a report presenting estimates about the investment costs to build a biorefinery in Wiggins, Mississippi [176]. The report states that the investment costs to build a biorefinery in Wiggins, MS which produces 58 MGY of ethanol is $310,102,000.
Wallace et al. stated in his study that doubling the size of a plant (from 50 to 100 MGY) would increase the investment costs by a factor of 1.6 [177]. We used these estimates to calculate estimates for the investment costs for biorefineries of different sizes. We used a projected life of 20 years and an interest rate of 15% to calculate the annuity ($\psi_d$). After discussions with the experts, we assumed that the storage capacities of the biorefineries will be 10% of their annual production capacities. The biorefineries were assumed to run for 330 days a year. Fifteen potential biorefinery locations were selected based on the higher quantity of biomass availability. Seven discrete biorefinery sizes were selected, i.e. biorefineries producing 10, 20, 30, 40, 60, 100 and 150 MGY.

**Computational Results**

Using the data collected in the Data Collection section, a base scenario was developed. An additional 41 scenarios were created by changing one factor at a time for the data from the base scenario and are shown in Table 4.

**MIP Model**

The model formulated above in chapter 3 is fed into CPLEX 9.0, a commercial LP/MIP solver and coded in C++. 
Table 4  Different Scenarios

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>80%</td>
<td>80%</td>
<td>80%</td>
<td>110%</td>
<td>80%</td>
<td>50%</td>
</tr>
<tr>
<td>90%</td>
<td>90%</td>
<td>90%</td>
<td>90%</td>
<td>120%</td>
<td>90%</td>
<td>75%</td>
</tr>
<tr>
<td>110%</td>
<td>110%</td>
<td>110%</td>
<td>110%</td>
<td>130%</td>
<td>110%</td>
<td></td>
</tr>
<tr>
<td>120%</td>
<td>120%</td>
<td>120%</td>
<td>120%</td>
<td>140%</td>
<td>120%</td>
<td></td>
</tr>
<tr>
<td>150%</td>
<td>130%</td>
<td>130%</td>
<td>130%</td>
<td>150%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>175%</td>
<td>140%</td>
<td>140%</td>
<td>140%</td>
<td>200%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200%</td>
<td>150%</td>
<td>150%</td>
<td>150%</td>
<td>250%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The model was solved a total of 42 times and the results obtained are presented in Table 5. For 29 out of 42 scenarios two biorefineries were opened, one at Attala County with a production capacity of 150 MGY and another at Simpson County with a production capacity of 100 MGY. In 6 other scenarios two biorefineries were opened, one at Leflore County with a production capacity of 150 MGY and another at Smith County with a production capacity of 100 MGY. Out of the remaining 7 scenarios (all Biomass Supply Scenarios), in 4 scenarios, two biorefineries were opened, one at Attala County with a production capacity of 150 MGY and another at Smith County with a production capacity of 20 MGY; in 2 scenarios, two biorefineries were opened, one at Smith County with production capacity of 150 MGY and another at Attala County with production capacity of 20 MGY; and finally in remaining one scenario, two biorefineries were opened, one at Attala County with a production capacity of 150 MGY and another
at Simpson County with a production capacity of 20 MGY. The location of the biorefinery is influenced by the locations of harvesting sites and local markets. Biorefinery size and location are also influenced by the locations and the capacities of the collection facilities if the biomass is routed to the biorefineries via collection facilities. Thus in our case, even though economies of scale would suggest biorefinery sizes of 150 and 20 MGY, the model selects 150 and 100 MGY respectively for most of the scenarios, except in Biomass Supply scenarios. This is due to the fact that storing biomass is more costly than storing C-ethanol and therefore, biomass was converted into C-ethanol which was stored at biorefineries thus increasing the biorefinery capacity. In Biomass Supply Scenarios, the biomass availability increases and thus there is no need of storing biomass or C-ethanol and therefore the model selects biorefinery sizes of 150 and 20 MGY respectively. Table 5 displays cost distribution of different cost parameters for unit delivery costs for C-ethanol in $/gallon for all the scenarios.
Table 5  Cost Distribution for Different Scenarios

<table>
<thead>
<tr>
<th></th>
<th>Investment Cost ($/gallon)</th>
<th>Harvesting Cost ($/gallon)</th>
<th>Processing Cost ($/gallon)</th>
<th>Total BM Trans Cost ($/gallon)</th>
<th>Total BM Inv Cost ($/gallon)</th>
<th>Total Eth Trans Cost ($/gallon)</th>
<th>Total Eth Inv Cost ($/gallon)</th>
<th>Total Cost ($/gallon)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base Case</strong></td>
<td>100%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>$0.82</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>$0.92</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>$1.13</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>$1.23</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>$1.54</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.40</td>
</tr>
<tr>
<td></td>
<td>175%</td>
<td>$1.79</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.40</td>
</tr>
<tr>
<td></td>
<td>200%</td>
<td>$2.05</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td><strong>Investment Cost</strong></td>
<td>80%</td>
<td>$1.02</td>
<td>$0.17</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>$1.02</td>
<td>$0.19</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>$1.02</td>
<td>$0.23</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>$1.02</td>
<td>$0.25</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>$1.02</td>
<td>$0.27</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>140%</td>
<td>$1.02</td>
<td>$0.29</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>$1.02</td>
<td>$0.31</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td><strong>Biomass Harvesting Cost</strong></td>
<td>80%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.39</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.44</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.54</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.59</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.64</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td>Processing Cost</td>
<td>140%</td>
<td>150%</td>
<td>140%</td>
<td>150%</td>
<td>140%</td>
<td>150%</td>
<td>140%</td>
<td>150%</td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Biomass Transport Cost</td>
<td>80%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.29</td>
<td>$0.01</td>
<td>$0.39</td>
<td>$2.43</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.40</td>
<td>$0.01</td>
<td>$0.22</td>
<td>$2.54</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.48</td>
<td>$0.01</td>
<td>$0.22</td>
<td>$2.62</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.55</td>
<td>$0.01</td>
<td>$0.22</td>
<td>$2.70</td>
</tr>
<tr>
<td>Biomass Supply (Yield)</td>
<td>110%</td>
<td>$0.66</td>
<td>$0.23</td>
<td>$0.50</td>
<td>$0.34</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.00</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>$0.66</td>
<td>$0.22</td>
<td>$0.50</td>
<td>$0.31</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.00</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>$0.66</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.32</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.00</td>
</tr>
<tr>
<td></td>
<td>250%</td>
<td>$0.66</td>
<td>$0.08</td>
<td>$0.44</td>
<td>$0.40</td>
<td>$0.00</td>
<td>$0.02</td>
<td>$0.00</td>
</tr>
<tr>
<td>Inventory Holding Cost</td>
<td>80%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.00</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>$1.02</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>$1.28</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>$1.28</td>
<td>$0.21</td>
<td>$0.49</td>
<td>$0.37</td>
<td>$0.01</td>
<td>$0.02</td>
<td>$0.39</td>
</tr>
</tbody>
</table>

Table 5 (continued)
Figure 7 shows the changes in the unit $/gallon delivery costs for C-ethanol with respect to changes in different cost parameters like investment costs, processing costs, harvesting costs and transportation costs. It should be noted that for each scenario in Figure 7, we change only one parameter keeping others constant. For example, for an Investment Cost scenario, changes are done in investment costs only keeping other costs (processing, harvesting, etc.) constant. It can be seen from the figure that the impact of changes in investment costs is much more as compared with the other three cost parameters.

![Unit Delivery Costs for Different Scenarios](image)

**Figure 7  Changes in Unit Delivery C-ethanol Cost**

Table 6 shows the amount of biomass processed for the scenarios where the biomass supply increases from 110% to 250%. It can be seen from the table that as the
biomass supply increases the amount of corn stover processed decreases while the amount of sawtimber processed increases.

Table 6  Biomass Processed with Biomass Supply

<table>
<thead>
<tr>
<th>Biomass Supply (Yield)</th>
<th>Corn Stover</th>
<th>Sawtimber</th>
<th>Pulpwood</th>
</tr>
</thead>
<tbody>
<tr>
<td>110%</td>
<td>1,273,344</td>
<td>1,747,463</td>
<td>48,855</td>
</tr>
<tr>
<td>120%</td>
<td>1,287,658</td>
<td>1,726,504</td>
<td>39,198</td>
</tr>
<tr>
<td>130%</td>
<td>1,244,373</td>
<td>1,818,676</td>
<td>39,606</td>
</tr>
<tr>
<td>140%</td>
<td>1,211,254</td>
<td>1,890,986</td>
<td>38,132</td>
</tr>
<tr>
<td>150%</td>
<td>1,161,391</td>
<td>1,997,207</td>
<td>38,560</td>
</tr>
<tr>
<td>200%</td>
<td>516,992</td>
<td>3,335,108</td>
<td>78,912</td>
</tr>
<tr>
<td>250%</td>
<td>418,445</td>
<td>3,550,754</td>
<td>74,041</td>
</tr>
</tbody>
</table>

Corn stover has a higher conversion rate than sawtimber and pulpwood, but also has higher harvesting, processing and inventory holding costs. Thus, as more and more biomass is available to reduce these costs, sawtimber is used which has lower conversion rates and higher transportation costs, as it is bulkier and heavier to transport. Figure 8 shows an increase in the amount of sawtimber processed and a decrease in the amount of corn stover. Pulpwood availability for conversion to biofuel is very limited and therefore, a very small amount of pulpwood is processed.
Figure 8  Changes in Amount of Biomass Types Processed

Figure 9 shows the impact of biomass supply on different cost parameters, like harvesting costs, processing costs and transportation costs. It can be seen from the figure below that as the biomass supply increases, processing and harvesting costs decrease while transportation costs increase. This is due to an increased amount of processing of sawtimber which has higher transportation costs and lower harvesting and processing costs as compared to corn stover.
The average cost of producing cellulosic ethanol as estimated by Department of Energy (DoE) is $2.20 per gallon [178] and USDA estimated the same to $2.65 [179]. Both of these costs are just production costs and do not include any distribution costs. Our model estimates a base cost of $2.51 per gallon and an average cost of $2.47 per gallon for producing and distributing cellulosic ethanol with a maximum of $3.53 and a minimum of $1.60 based on different scenarios. One of the reasons our model has lower delivery costs is that with our model we remove the restrictions on distance traveled to obtain biomass which is usually a 50 miles radius for other models.

Further we studied the effect of biomass availability on the biomass and C-ethanol inventory holding costs. It can be seen from Table 5 that as the biomass supply increases, biomass inventory costs decrease. This is due to the fact that as the biomass supply increases, more sawtimber is used which is available 9-months a year as compared to 3-months a year for corn stover, thus reducing the need for its storage. Also, it can be seen that...
from Table 5 that C-ethanol inventory costs are reduced to zero. This is because for all other scenarios it is cheaper to store C-ethanol than biomass, as most of the biomass processed is corn stover which is available only 3-months a year and is more costly to store, thus requiring C-ethanol storage. For Biomass Supply scenarios, as more biomass becomes available, sawtimber is processed more and as it is available 9-months a year there is no need for storing either the biomass or C-ethanol.

In another experiment, we studied the impact of changes in conversion rates with changes in other cost parameters, like investment costs, processing costs, harvesting costs and transportation costs. In all of 140 scenarios generated, two biorefineries were opened, one with a production capacity of 150 MGY and another with a production capacity of 20 MGY. This is because, as the conversion rates increase, more C-ethanol can be produced using less biomass and thus there is no need of storing either the biomass or C-ethanol. Table 7 represents the unit $/gallon C-ethanol delivery costs for different scenarios with changes in conversion rates.
Table 7  Unit Delivery Cost of C-ethanol with Conversion Rates

<table>
<thead>
<tr>
<th>Conversion Rate</th>
<th>110%</th>
<th>120%</th>
<th>130%</th>
<th>140%</th>
<th>150%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>$1.52</td>
<td>$1.43</td>
<td>$1.35</td>
<td>$1.48</td>
<td>$1.23</td>
</tr>
<tr>
<td>90%</td>
<td>$1.59</td>
<td>$1.49</td>
<td>$1.41</td>
<td>$1.52</td>
<td>$1.29</td>
</tr>
<tr>
<td>110%</td>
<td>$1.72</td>
<td>$1.62</td>
<td>$1.55</td>
<td>$1.55</td>
<td>$1.42</td>
</tr>
<tr>
<td>120%</td>
<td>$1.79</td>
<td>$1.69</td>
<td>$1.61</td>
<td>$1.59</td>
<td>$1.49</td>
</tr>
<tr>
<td>150%</td>
<td>$1.98</td>
<td>$1.89</td>
<td>$1.81</td>
<td>$1.38</td>
<td>$1.69</td>
</tr>
<tr>
<td>175%</td>
<td>$2.15</td>
<td>$2.05</td>
<td>$1.97</td>
<td>$1.40</td>
<td>$1.85</td>
</tr>
<tr>
<td>200%</td>
<td>$2.31</td>
<td>$2.21</td>
<td>$2.14</td>
<td>$1.43</td>
<td>$2.02</td>
</tr>
<tr>
<td>Processing Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>$1.56</td>
<td>$1.47</td>
<td>$1.40</td>
<td>$1.44</td>
<td>$1.29</td>
</tr>
<tr>
<td>90%</td>
<td>$1.61</td>
<td>$1.52</td>
<td>$1.44</td>
<td>$1.46</td>
<td>$1.33</td>
</tr>
<tr>
<td>110%</td>
<td>$1.70</td>
<td>$1.60</td>
<td>$1.52</td>
<td>$1.47</td>
<td>$1.39</td>
</tr>
<tr>
<td>120%</td>
<td>$1.74</td>
<td>$1.64</td>
<td>$1.56</td>
<td>$1.48</td>
<td>$1.42</td>
</tr>
<tr>
<td>130%</td>
<td>$1.79</td>
<td>$1.68</td>
<td>$1.59</td>
<td>$1.36</td>
<td>$1.46</td>
</tr>
<tr>
<td>140%</td>
<td>$1.84</td>
<td>$1.72</td>
<td>$1.63</td>
<td>$1.39</td>
<td>$1.49</td>
</tr>
<tr>
<td>150%</td>
<td>$1.88</td>
<td>$1.76</td>
<td>$1.67</td>
<td>$1.44</td>
<td>$1.52</td>
</tr>
<tr>
<td>Biomass Harvesting Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>$1.61</td>
<td>$1.52</td>
<td>$1.44</td>
<td>$1.46</td>
<td>$1.33</td>
</tr>
<tr>
<td>90%</td>
<td>$1.63</td>
<td>$1.54</td>
<td>$1.46</td>
<td>$1.48</td>
<td>$1.34</td>
</tr>
<tr>
<td>110%</td>
<td>$1.67</td>
<td>$1.58</td>
<td>$1.50</td>
<td>$1.51</td>
<td>$1.37</td>
</tr>
<tr>
<td>120%</td>
<td>$1.70</td>
<td>$1.60</td>
<td>$1.51</td>
<td>$1.53</td>
<td>$1.38</td>
</tr>
<tr>
<td>130%</td>
<td>$1.72</td>
<td>$1.61</td>
<td>$1.28</td>
<td>$1.46</td>
<td>$1.39</td>
</tr>
<tr>
<td>140%</td>
<td>$1.74</td>
<td>$1.63</td>
<td>$1.35</td>
<td>$1.47</td>
<td>$1.40</td>
</tr>
<tr>
<td>150%</td>
<td>$1.76</td>
<td>$1.65</td>
<td>$1.48</td>
<td>$1.48</td>
<td>$1.41</td>
</tr>
<tr>
<td>Biomass Transportation Cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>$1.59</td>
<td>$1.50</td>
<td>$1.55</td>
<td>$1.36</td>
<td>$1.30</td>
</tr>
<tr>
<td>90%</td>
<td>$1.62</td>
<td>$1.53</td>
<td>$1.74</td>
<td>$1.39</td>
<td>$1.33</td>
</tr>
<tr>
<td>110%</td>
<td>$1.69</td>
<td>$1.59</td>
<td>$1.91</td>
<td>$1.44</td>
<td>$1.38</td>
</tr>
<tr>
<td>120%</td>
<td>$1.72</td>
<td>$1.61</td>
<td>$2.07</td>
<td>$1.46</td>
<td>$1.40</td>
</tr>
<tr>
<td>130%</td>
<td>$1.75</td>
<td>$1.64</td>
<td>$1.34</td>
<td>$1.48</td>
<td>$1.42</td>
</tr>
<tr>
<td>140%</td>
<td>$1.78</td>
<td>$1.67</td>
<td>$1.38</td>
<td>$1.51</td>
<td>$1.45</td>
</tr>
<tr>
<td>150%</td>
<td>$1.82</td>
<td>$1.69</td>
<td>$1.45</td>
<td>$1.53</td>
<td>$1.47</td>
</tr>
</tbody>
</table>

Table 8 shows CPU times for each of the 140 scenarios. It can be seen that as the conversion rates increase, less time is required to solve the problem.
Table 8  CPU Times vs. Conversion Rates

<table>
<thead>
<tr>
<th>Conversion Rate</th>
<th>110%</th>
<th>120%</th>
<th>130%</th>
<th>140%</th>
<th>150%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass Harvesting Cost</td>
<td>80%</td>
<td>3,323</td>
<td>3,915</td>
<td>2,109</td>
<td>2,377</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>2,979</td>
<td>3,601</td>
<td>2,054</td>
<td>2,009</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>3,283</td>
<td>2,591</td>
<td>1,953</td>
<td>1,951</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>3,740</td>
<td>2,709</td>
<td>1,878</td>
<td>2,457</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>3,093</td>
<td>2,604</td>
<td>2,292</td>
<td>2,237</td>
</tr>
<tr>
<td></td>
<td>175%</td>
<td>3,178</td>
<td>2,152</td>
<td>1,974</td>
<td>1,914</td>
</tr>
<tr>
<td></td>
<td>200%</td>
<td>3,165</td>
<td>2,269</td>
<td>2,063</td>
<td>2,240</td>
</tr>
<tr>
<td>Processing Cost</td>
<td>80%</td>
<td>3,870</td>
<td>2,368</td>
<td>2,111</td>
<td>1,832</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>4,118</td>
<td>2,151</td>
<td>2,040</td>
<td>2,009</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>3,970</td>
<td>2,518</td>
<td>2,321</td>
<td>2,342</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>4,547</td>
<td>2,574</td>
<td>2,602</td>
<td>1,817</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>3,596</td>
<td>2,721</td>
<td>2,248</td>
<td>2,052</td>
</tr>
<tr>
<td></td>
<td>140%</td>
<td>4,247</td>
<td>2,512</td>
<td>2,247</td>
<td>2,532</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>4,778</td>
<td>2,757</td>
<td>2,545</td>
<td>1,811</td>
</tr>
<tr>
<td>Biomass Transportation Cost</td>
<td>80%</td>
<td>3,666</td>
<td>2,286</td>
<td>2,156</td>
<td>2,024</td>
</tr>
<tr>
<td></td>
<td>90%</td>
<td>3,242</td>
<td>2,320</td>
<td>2,250</td>
<td>1,702</td>
</tr>
<tr>
<td></td>
<td>110%</td>
<td>2,888</td>
<td>2,293</td>
<td>2,079</td>
<td>1,782</td>
</tr>
<tr>
<td></td>
<td>120%</td>
<td>3,892</td>
<td>2,356</td>
<td>2,508</td>
<td>1,544</td>
</tr>
<tr>
<td></td>
<td>130%</td>
<td>4,017</td>
<td>2,429</td>
<td>2,490</td>
<td>2,008</td>
</tr>
<tr>
<td></td>
<td>140%</td>
<td>4,964</td>
<td>3,307</td>
<td>2,610</td>
<td>2,333</td>
</tr>
<tr>
<td></td>
<td>150%</td>
<td>4,293</td>
<td>3,066</td>
<td>2,084</td>
<td>1,816</td>
</tr>
</tbody>
</table>

In summary, the biorefinery locations are influenced by location of harvesting sites and markets. The size and location of biorefineries also depend on the size and location of the collection facilities, that is, if biomass is routed through collection...
facilities to biorefineries. Biomass supply does have a negative effect on the biomass inventory holding costs and has a positive effect on the biomass transportation costs. Biomass supply also influences the size of a biorefinery due to the fact that a greater time period of availability of biomass reduces biomass storage requirements reducing the size of a biorefinery. An increase in the biomass supply has a negative effect on biomass harvesting, processing and inventory holding costs and although it increases biomass transportation costs for some scenarios, its overall effect reduces the total costs for the supply chain.

**Lagrangean Decomposition Heuristics**

A Lagrangean decomposition (LD) heuristic is run for all of the 42 cases of the problem generated as described in the above section. Our LD heuristic generated lower bounds \( LB \) and upper bounds \( UB \) for all the problem scenarios. LD was run for 50 iterations; solution values and running times of LD are recorded for each problem. To have an equal comparison, CPLEX is run for the same amount of time as LD for each problem and \( LBs \) and \( UBs \) from CPLEX are recorded. Out of 42 problems solved, CPLEX failed to find the bounds on 16 occasions. Table 9 shows the performance comparison between LD and CPLEX for problems to which CPLEX was able to find the bounds.
Table 9  Computational Results for CPLEX and LD Heuristics

<table>
<thead>
<tr>
<th>Problem Scenarios</th>
<th>Performance GAP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UB</td>
</tr>
<tr>
<td>Investment Cost</td>
<td></td>
</tr>
<tr>
<td>110%</td>
<td>52.79%</td>
</tr>
<tr>
<td>120%</td>
<td>53.91%</td>
</tr>
<tr>
<td>175%</td>
<td>58.09%</td>
</tr>
<tr>
<td>200%</td>
<td>59.31%</td>
</tr>
<tr>
<td>Processing Cost</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>51.85%</td>
</tr>
<tr>
<td>130%</td>
<td>50.41%</td>
</tr>
<tr>
<td>140%</td>
<td>50.06%</td>
</tr>
<tr>
<td>150%</td>
<td>49.72%</td>
</tr>
<tr>
<td>Harvesting Cost</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>51.64%</td>
</tr>
<tr>
<td>110%</td>
<td>51.33%</td>
</tr>
<tr>
<td>130%</td>
<td>51.02%</td>
</tr>
<tr>
<td>140%</td>
<td>50.87%</td>
</tr>
<tr>
<td>Transportation Cost</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>51.82%</td>
</tr>
<tr>
<td>130%</td>
<td>50.51%</td>
</tr>
<tr>
<td>150%</td>
<td>49.88%</td>
</tr>
<tr>
<td>Biomass Supply</td>
<td></td>
</tr>
<tr>
<td>110%</td>
<td>-26.98%</td>
</tr>
<tr>
<td>120%</td>
<td>68.33%</td>
</tr>
<tr>
<td>130%</td>
<td>-30.56%</td>
</tr>
<tr>
<td>140%</td>
<td>-18.31%</td>
</tr>
<tr>
<td>150%</td>
<td>-20.57%</td>
</tr>
<tr>
<td>200%</td>
<td>-9.41%</td>
</tr>
<tr>
<td>250%</td>
<td>-10.79%</td>
</tr>
<tr>
<td>Inventory Holding Cost</td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>51.48%</td>
</tr>
<tr>
<td>90%</td>
<td>51.48%</td>
</tr>
<tr>
<td>110%</td>
<td>51.48%</td>
</tr>
<tr>
<td>Projected Life</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>54.39%</td>
</tr>
</tbody>
</table>

The values under UB and LB columns in Table 9 are the performance gaps between the two methods (LD and CPLEX). The number in those columns is an indicator of how good one method is over the other. For example, for the first problem, i.e. with investment costs at 110%, the LD performs 52.79% better in finding UB but performs 20.87% worse in finding LB as compared to CPLEX. The negative values indicate poor
performance and a 100% indicates that LD was not able to find the bounds. The performance gap is calculated by using the following equations.

\[
GAP(UB) = \frac{\text{CPLEX}_{UB} - \text{LD}_{UB}}{\text{CPLEX}_{UB}} \times 100 \% \quad (71)
\]

\[
GAP(LB) = \frac{\text{LD}_{LB} - \text{CPLEX}_{LB}}{\text{CPLEX}_{LB}} \times 100 \% \quad (72)
\]

The maximum and minimum gaps of the 36 problems for which LD was able to find bounds is 72.71% and 37.54% with an average gap of 44.71%. The average time to solve the problem was 237 seconds of CPU time. The problems were run on a Dell system with Intel Pentium 2.80 GHz processor and 1 GB RAM.

In summary, we proposed a Lagrangean based LD heuristics to compute \( UB \) and \( LB \) for the biomass-to-biorefinery supply chain design and management problem. The LD heuristic was run for different problem scenarios and the computational results are presented in Table 9. The LD heuristic performed better in finding \( UBs \) while CPLEX performed better in finding \( LBs \).

DSS Model

25 counties are selected as potential biorefinery locations from varied areas of the state of Mississippi and 7 sizes are selected as potential biorefinery sizes. For problem S1, where locations and sizes are known, 175 instances are created (25 locations and 7
sizes). For problem S2, where size is known but locations are not, 25 instances are created, one for each location with the chosen size. For problem S3, where location is known but size is not, 7 instances are created, one for each size with the chosen location. Finally for problem S4, 5 instances are created, one with original biofuel demand and other four by changing the demand by ±10% and ±20%. All the instances are solved using both the Excel-based DSS and CPLEX 9.0 commercial MIP solver. The results for the problems are presented in the following tables. Table 10 presents the average unit delivery costs for Ethanol ($/gallons) for both DSS and CPLEX models, Table 11 shows the minimum, average, and maximum time required to solve each type of problem by both methods, and Table 12 shows the minimum, average and maximum percentage gap between the solutions obtained from DSS and CPLEX. For experimental purposes, the biomass considered in both methods is Corn and therefore, the biofuel produced is assumed to be Ethanol.

Table 10  Average Unit Delivery Cost for DSS and CPLEX

<table>
<thead>
<tr>
<th>Case</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSS</td>
<td>1.61</td>
<td>1.54</td>
<td>1.40</td>
<td>1.44</td>
</tr>
<tr>
<td>CPLEX</td>
<td>1.51</td>
<td>1.48</td>
<td>1.36</td>
<td>1.35</td>
</tr>
</tbody>
</table>
Table 11 Minimum, Average and Maximum Solution Times for DSS and CPLEX

<table>
<thead>
<tr>
<th>Problem</th>
<th>Time (seconds)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSS</td>
<td></td>
<td></td>
<td>CPLEX</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>Average</td>
<td>Max</td>
<td>Min</td>
<td>Average</td>
<td>Max</td>
</tr>
<tr>
<td>S1</td>
<td>0.30</td>
<td>0.72</td>
<td>2.92</td>
<td>0.19</td>
<td>13.01</td>
<td>1944.50</td>
</tr>
<tr>
<td>S2</td>
<td>0.64</td>
<td>2.95</td>
<td>6.00</td>
<td>2.00</td>
<td>90.00</td>
<td>1959.00</td>
</tr>
<tr>
<td>S3</td>
<td>2.17</td>
<td>11.04</td>
<td>32.80</td>
<td>12.88</td>
<td>366.79</td>
<td>2080.52</td>
</tr>
<tr>
<td>S4</td>
<td>1.42</td>
<td>2.02</td>
<td>2.95</td>
<td>92.00</td>
<td>270.00</td>
<td>677.00</td>
</tr>
</tbody>
</table>

Table 12: Minimum, Average and Maximum GAP between DSS and CPLEX Solutions

<table>
<thead>
<tr>
<th>Problem</th>
<th>GAP (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S1</td>
<td>S2</td>
<td>S3</td>
<td>S4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSS-CPLEX</td>
<td>Min</td>
<td>1.13</td>
<td>0.13</td>
<td>0.03</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>5.60</td>
<td>3.91</td>
<td>2.99</td>
<td>6.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>14.99</td>
<td>7.39</td>
<td>8.74</td>
<td>9.81</td>
<td></td>
</tr>
</tbody>
</table>

The average gap is calculated using the following formula:

\[
GAP \left( \frac{DSS_{soln} - CPLEX_{soln}}{DSS_{soln}} \right) \times 100 \%
\] \hspace{1cm} (73)

Figure 10 shows the average gap obtained between the DSS and CPLEX solution values for problem S1 solved for each biorefinery size. It can be seen from the figure that as the capacity of the biorefinery increases the average gap decreases until a size 40 MM gallon/year and then increases again as size increases. This shows that the problem with size 40 is the easiest problem to solve. This is due to various factors such as sizes of collection facilities, investment costs, etc.
Figure 10  Average GAP – DSS and CPLEX Solutions for Problem S1

Figure 11 shows the average time taken by both the methods. CPLEX times are higher than DSS times and therefore, for comparison purposes, the natural logarithm of both the times is taken and is plotted in the figure below. It can be seen from the figure that the average time taken to solve the problem increases as the capacity size increases. CPLEX times increase exponentially as compared with DSS times. Also, for one problem with size 150 MM gallons/year, CPLEX goes out of memory and is unable to solve the problem whereas DSS solves the problem in 2.80 seconds.
Figure 11  Average Time – DSS and CPLEX for Problem S1

Figure 12 shows the solution values obtained by both the methods for problem S1. It can be seen from the figure that even though there is a significant gap between the two solution values, the pattern for both is similar, i.e. solutions for both methods decrease with the biorefinery size until it reaches 40 MM gallons/year where the solution is minimal and then increases as the size increases. Both of the methods suggest 40 MM gallons/year to be the optimal facility size to be opened. This is because as the biorefinery size increases, due to economy of scales the unit cost decreases, but this decrease is also influenced by the availability of biomass and unit biomass transportation costs. Thus as the size increases more biomass is required and so the average distance traveled away from biorefinery increases which again increases the unit cost.
For problem S3, in 4 out of 7 problems, the locations selected by the DSS and CPLEX are same, and for the other 3 problems, the locations selected by DSS are within 75 miles of the locations selected by CPLEX.

Table 13  Results for Problem S4 from DSS and CPLEX

<table>
<thead>
<tr>
<th>Ethanol Demands</th>
<th>Solution Value</th>
<th>Time (seconds)</th>
<th>GAP (%)</th>
<th># BR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSS</td>
<td>CPLEX</td>
<td>DSS</td>
<td>CPLEX</td>
</tr>
<tr>
<td>134,400,000</td>
<td>1.39</td>
<td>1.34</td>
<td>1.42</td>
<td>92</td>
</tr>
<tr>
<td>151,200,000</td>
<td>1.42</td>
<td>1.34</td>
<td>1.66</td>
<td>120</td>
</tr>
<tr>
<td>168,000,000</td>
<td>1.44</td>
<td>1.35</td>
<td>1.83</td>
<td>160</td>
</tr>
<tr>
<td>184,800,000</td>
<td>1.44</td>
<td>1.35</td>
<td>2.23</td>
<td>299</td>
</tr>
<tr>
<td>201,600,000</td>
<td>1.50</td>
<td>1.35</td>
<td>2.95</td>
<td>677</td>
</tr>
</tbody>
</table>

Table 13 shows the comparison between DSS and CPLEX results for problem S4. The minimum gap obtained is 3.32% and the maximum gap obtained is 9.81%. The gap
obtained is a function of the biorefinery and collection facility sizes selected as well as the demand for ethanol. The biorefinery locations selected and the sizes opened by DSS and CPLEX are shown in Table 14. The locations selected by DSS are more or less similar to that of CPLEX. For example, in scenario 2 (demand = 151,200,000 gallons) locations selected by CPLEX are Sunflower and Yazoo counties, and locations selected by DSS are Sunflower and Holmes counties. With Yazoo and Holmes counties being adjacent counties the distance between them is minimal.

Table 14  Biorefinery Locations and Sizes for Problem S4 from DSS and CPLEX

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Demand</th>
<th>DSS</th>
<th>CPLEX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BR</td>
<td>BR</td>
</tr>
<tr>
<td>1</td>
<td>134,400,000</td>
<td>Sunflower 150</td>
<td>Sunflower 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yazoo 40</td>
</tr>
<tr>
<td>2</td>
<td>151,200,000</td>
<td>Sunflower 150</td>
<td>Sunflower 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Holmes 10</td>
<td>Yazoo 60</td>
</tr>
<tr>
<td>3</td>
<td>168,000,000</td>
<td>Holmes 150</td>
<td>Hinds 30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sunflower 100</td>
</tr>
<tr>
<td>4</td>
<td>184,800,000</td>
<td>Sunflower 30</td>
<td>Hinds 40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Montgomery 10</td>
<td>Sunflower 150</td>
</tr>
<tr>
<td>5</td>
<td>201,600,000</td>
<td>Montgomery 150</td>
<td>Lafayette 60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Holmes 60</td>
<td>Sunflower 60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yazoo 100</td>
</tr>
</tbody>
</table>

To summarize, we present a modified version (MP) of the original problem (P). The original biomass-to-biorefinery supply chain problem is divided into four simple supply chain problems. Algorithms are developed using VBA in Excel. Results obtained are compared with CPLEX. The gap between DSS and CPLEX solutions ranges
anywhere from 2.5-6.5%. Also, DSS is faster compared to CPLEX as it solves the
problems within a few seconds.

CPLEX is a commercial MIP solver which is more costly to buy and also requires
some basic knowledge of programming languages, whereas DSS has a simple easy to use
interface, where the decision makers can directly input data and click on few buttons to
generate reports for any specific problem.
CHAPTER VII
CONCLUSIONS

In this dissertation we study optimization models that integrate location, production, inventory and transportation decisions for industrial products and apply that information to develop supply chains for agricultural products (biomass). We formulate the biomass-to-biorefinery supply chain problem as a mixed integer linear programming model. The model is applied to a case study. 42 different scenarios based on changes in the problem parameters’ values are constructed. Numerical experiments are done to measure the performance of the model.

We develop Lagrangean decomposition heuristic. In our heuristic, we divide the problem into two sub-problems, sub-problem 1 is a transportation problem and sub-problem 2 is a combination of a capacitated facility location and production planning problem. Sub-problem 2 is further divided by commodities. Our heuristic provides both the bounds (upper and lower) for a given problem scenario. The algorithm is tested for a number of different problem scenarios and its performance is compared with CPLEX solutions for each of the problem scenarios.

We provide a modified version of the original biomass-to-biorefinery supply chain problem and develop a decision support system (DSS). In DSS, the main problem is divided into four easy-to-solve supply chain problems. These problems were
determined based on our knowledge of supply chain and discussions with the experts from the biomass and biofuels’ sector. The DSS is coded using visual basic applications (VBA) for Excel and has a simple user interface which assists the user in running different types of supply chain problems and provides results in the form of reports which are easy to understand. A numerical analysis is done and the solutions obtained by DSS are compared with the CPLEX solutions for the given problem type.
REFERENCES


http://www.12manage.com/description_supply_chain_design.html


[149] Babonneau, F., and J.P. Vial," ACCPM with a nonlinear constraint and an active set strategy to solve nonlinear multicommodity flow problems", *Mathematical Programming*.


[176] "Central Mississippi Forest Inventory": Mississippi Institute of Forest Inventory (MIFI). http://www.mifi.ms.gov

