OPTIMIZATION AND SIMULATION FOR DESIGNING THE SUPPLY CHAIN OF THE CELLULOSIC BIOFUEL INDUSTRY

A Dissertation

by

HEUNGJO AN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2011

Major Subject: Industrial Engineering

Optimization and Simulation for Designing the Supply Chain of the Cellulosic Biofuel Industry

OPTIMIZATION AND SIMULATION FOR DESIGNING THE SUPPLY CHAIN OF THE CELLULOSIC BIOFUEL INDUSTRY

A Dissertation

by

HEUNGJO AN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Wilbert E. Wilhelm
	Stephen W. Searcy
Committee Members,	Halit Üster
	Lewis Ntaimo
Head of Department,	César O. Malavé

December 2011

Major Subject: Industrial Engineering

ABSTRACT

Optimization and Simulation for Designing the Supply Chain of the Cellulosic Biofuel Industry.

(December 2011)

Heungjo An, B.S.; M.S., Hanyang University

Co-Chairs of Advisory Committee, Dr. Wilbert E. Wilhelm
Dr. Stephen W. Searcy

The purpose of this dissertation is to provide an effective approach to design the supply chain (SC) of the cellulosic biofuel industry in order that it will support and accelerate the successful commercialization of the cellulosic biofuel industry. The methods of approach to this problem are (1) to assess the state-of-the-art biofuel SC studies, (2) to provide a decision support tool based on a mixed integer programming (MIP) model for the cellulosic biofuel supply chain design problem (BSCP), (3) to devise an exact solution method to solve large-scale instances of BSCP, (4) to evaluate a biomass logistics system based on biomass modules, by using new simulation elements for new machines, and (5) to compare several biomass logistics systems based on biomass module, bale, and silage, using simulation models.

The first part of this dissertation broadly reviews the literature on biofuel SCs, analyzing the state-of-the-art biofuel and petroleum-based fuel SC studies as well as relating generic SC models (An et al., 2010a). The second part of this dissertation formulates BSCP as a MIP model, which is a time-staged, multi-commodity flow, network design problem with an objective of maximizing profit (An et al., 2010b), providing a case study to demonstrate managerial use in application to a region in Central Texas.

The third part of this dissertation provides an exact solution method to solve BSCP. An embedded structure can be transformed to a generalized minimum cost flow problem, which is

used as a sub-problem in a CG approach. This study proposes a dynamic programming algorithm to solve the sub-problem in *O(m),* generating improving path-flows. To accelerate branch-and-bound (B&B) search, it develops an inequality, called the *partial objective constraint* (POC), which is based on the portion of the objective function associated with binary variables.

The fourth part of this dissertation evaluates a biomass module system, which is a conceptual logistics system based on large packages of chopped biomass with sufficient size and density to provide maximized legal highway loads and quick load/unload times. The last part of this dissertation evaluates economic benefits of the biomass module system, comparing it to bale and silage systems.

DEDICATION

To my family

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Wilbert E. Wilhelm, for his dedicated teachings and guidance throughout my Ph.D. studies. He introduced me to the field of the agricultural industry as well as Integer Programming. Professor Wilhelm contributed to the model formulation and solution methods described in Chapters III and IV; and, through the editing process, he composed much of the wording in Chapters I-IV. Moreover, his enthusiastic and highly respectable attitude as a scholar impressed and motivated me to have an aspiration to pass down his academic spirit to the next generation.

I also want to show my deep appreciation to my co-advisor, Dr. Searcy, for providing me an opportunity to experience practical, diverse, and valuable tasks as a research assistant for him in a great project for biomass logistics system, the results of which are some parts (i.e., chapters V and VI) of this dissertation. Professor Searcy contributed to the model formulation and data described in Chapters V and VI. His domain knowledge in agriculture and bioenergy areas was a base to define the problem for this dissertation. His broad horizon as well as practical-oriented approach influenced me greatly on having a balanced, academic philosophy.

I am also grateful to Dr. Üster and Dr. Ntaimo for serving as members of my advisory committee and for giving helpful suggestions. I gained lots of knowledge for logistics through Dr. Üster's three courses (601, 603, and 661). One of my major research outputs is based on what I have learned in the 'Network-Based Planning and Scheduling Systems' course. Moreover, when he was an advisor of the INFORMS student chapter, I learned about benefits of academic service and collaboration through an experience as a webmaster. Dr. Ntaimo transferred his highly valuable knowledge for stochastic programming to me. Computational experience for large-scale optimization through the "Large-Scale Stochastic Optimization" course was greatly

helpful to the computational tests for this dissertation.

Lastly, I want to mention my family's love and sacrifice. Without their patience and enormous support, this dissertation would not be possible. In particular, I would like to dedicate this dissertation to my father who went to heaven two years ago.

NOMENCLATURE

B&B	Branch and Bound
BRA	Backward Reaching Algorithm
BSCP	Cellulosic Biofuel Supply Chain Design Problem
CG	Column Generation
CRP	Conservation Reserve Program
DOE	U.S. Department of Energy
ETOH	Ethanol
GFP	Generalized Flow Problem
MIP	Mixed Integer Programming
POC	Partial Objective Constraint
RFS	Renewable Fuel Standard
SC	Supply Chain
SPFH	Self-Propelled Forage Harvester
USDA	U.S. Department of Agriculture

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Biofuel is one of the most important renewable energies. It can be distributed easily in the current technology and can be readily transported. Other renewable energies (e.g. solar, wind, hydropower, ocean, and geothermal) primarily produce electricity, which cannot readily power transportation vehicles using current technologies. During the last decade, countries around the world, especially in the U.S. and Europe, have tried to accelerate commercialization of biofuel to provide environmentally friendly, renewable energy.

The U.S. government passed Renewable Fuel Standard (RFS) 1 as part of the Energy Independence and Security Act in 2007 and amended it in 2010 (i.e., RFS2). RFS2 requires the U.S. to use 36 billion gallons per year (BGY) of domestic renewable fuel, including 16 BGY of cellulosic biofuel, starting in 2022, (see Figure 1) (U.S. Environmental Protection Agency, 2010). Based on RFS2, an obligated party (i.e., refiners and importers of gasoline) must produce or supply biofuels in proportion to the total amount of gasoline each supplies to the U.S. market.

First-generation biofuels have been commonly produced from animal fats and edible crops (e.g., corn). However, they have led to a concern that they could cause food prices to rise, leading to a crisis. This concern has stimulated development of second-generation biofuels (e.g. cellulosic ethanol (ETOH)), which are produced from cellulosic biomass, the residual, non-edible parts of food crops (e.g., stems, leaves and husks) as well as other non-food crops (e.g., forest wood, switch grass, jatropha, cereals that bear little grain, and industrial waste such as wood chips, skins and pulp from fruit pressing). Because it has a high oil content and is highly

---

This dissertation follows the style and format of *IIE Transactions*.

productive on a per-unit-area basis, algae has recently been highlighted as a feedstock for third-generation biofuel and is under development in research labs. However, since it will take some time to develop algae that can produce biofuel on a commercial-scale and RFS2 mandates production of substantial amounts of cellulosic biofuel, second generation biofuels will be the focus of commercial endeavors in the near future.

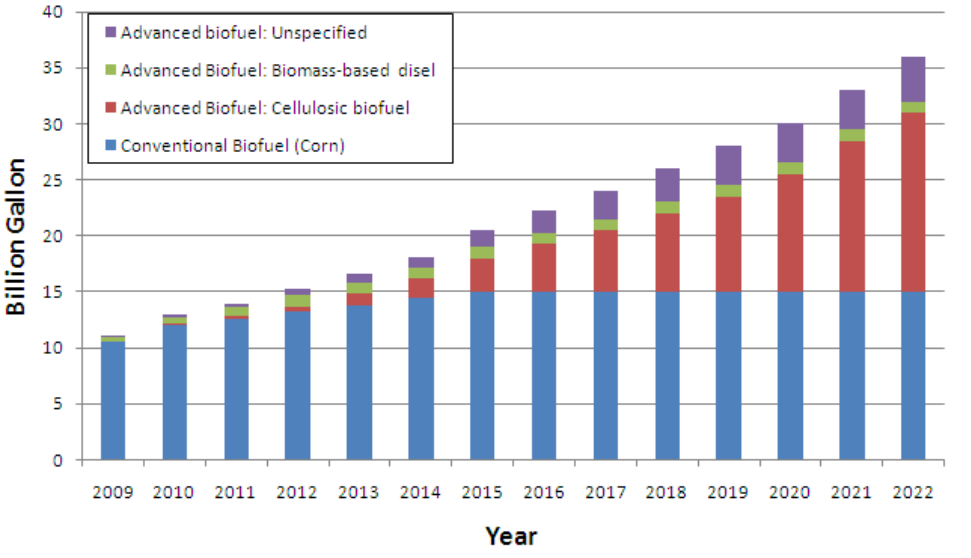

Figure 1 Renewable fuel requirements mandated by RFS2 (2010)

Second-generation biofuels are under development in a number of research labs and a few pilot refineries. However, second-generation feedstock has quite low energy density, is geographically distributed, and is degradable. Moreover, some feedstocks can be harvested only in specific seasons. Designing an economically viable SC to provide biofuel on a year around basis is a challenging problem.

The DOE reports that biomass logistics, which includes harvesting/collecting, storing, and pre-processing, constitutes as much as 20% of the current cost of supplying cellulosic ETOH (U.S.

Department of Agriculture (USDA) and DOE, 2008). Thus, reducing the logistics cost of cellulosic biomass is a key factor in successful commercialization. The DOE recently announced a goal for the biomass multi-year program plan: "The feedstock logistics goal is to reduce the dry herbaceous feedstock logistics cost to $0.39 per gallon (*$0.103 per liter*) of ETOH (equivalent to approximately $35 per dry ton *[$38.59/dry Mg]* in 2007 $) by 2012, with further reduction to $0.33 per gallon (*$0.087 per liter*) of ETOH by 2017" (DOE, 2010).

Moreover, the biofuel industry must cope with uncertainties (e.g. harvest yield due to weather and the market price of fuel). The moisture content of biomass fluctuates with weather conditions and is also uncertain. In addition, biofuel must compete with petroleum-based fuels, the price of which is highly variable due to complex relationships between supply, demand, fuel trading futures, and inventories. However, due to RFS2, the effect of price variation on biofuel production quantity will be relatively low, but its effect on the profit of biofuel suppliers may be significant. To make it possible for second generation biofuels to be competitive with first generation biofuels and petroleum-based fuels, it is essential that the feedstock logistics system as well as the SC be designed optimally. Developing a multi-period deterministic model for biofuel SC design is necessary to identify dynamic features in the biofuel industry and can lead to insights into system operations and interactions. Moreover, it is required to address details of several unique features of cellulosic feedstocks (e.g., high moisture content, dry matter loss in storage facilities, and single destination for feedstock supply) so that a model reflects the actual features that characterize the industry.

In addition, since reducing the cost of supplying feedstock is of primary importance, a computer model that estimates important measures (e.g., cost, energy used, CO_2 emission, biomass yield) by simulating detailed logistics operations is needed. A simulation model can consider operations in detail and accurately estimate costs that can be input to the proposed

optimization model.

Based on this background, section 1.1 introduces research objectives. Sections 1.2 and 1.3 give a problem statement and an overview of our method of approach for each research objective, respectively. The last section describes the organization of the dissertation.

1.1 Research objectives

The long term goal of this study is to provide an effective approach to design the cellulosic biofuel SC that will support and accelerate the successful commercialization of the biofuel industry. Specific research objectives are:

(1) Assessing the state-of-the-art biofuel SC studies;

(2) Providing a mathematical-based decision support tool to design the economically viable cellulosic biofuel SC, in which the overall system including location, capacity and technology of each facility is structured in the most profitable way and biomass are supplied to meet a year-round biofuel demand;

(3) Evaluating biofuel economics in the test region of Central Texas under several scenarios;

(4) Solving effectively large-scale instances of the modeled biofuel SC within reasonable run time;

(5) Developing new IBSAL simulation elements to model a new biomass logistics system based on large packages of chopped biomass, biomass module system; and

(6) Evaluating the performance of alternative biomass logistics systems, including bale, silage and module systems.

1.2 Problem statement

Figure 2 depicts the five-echelon biofuel SC, including feedstock production, preprocessing, conversion in refineries, distribution, and consumption by customers. The term

upstream is commonly used to refer to the portion of the SC from feedstock production to conversion plants; and *downstream*, from conversion plants, which are included in both upstream and downstream, to customer zones. Even though it is important to improve technologies and efficiencies in each echelon, integrating technologies and coordinating echelons is necessary for the system to be most profitable.

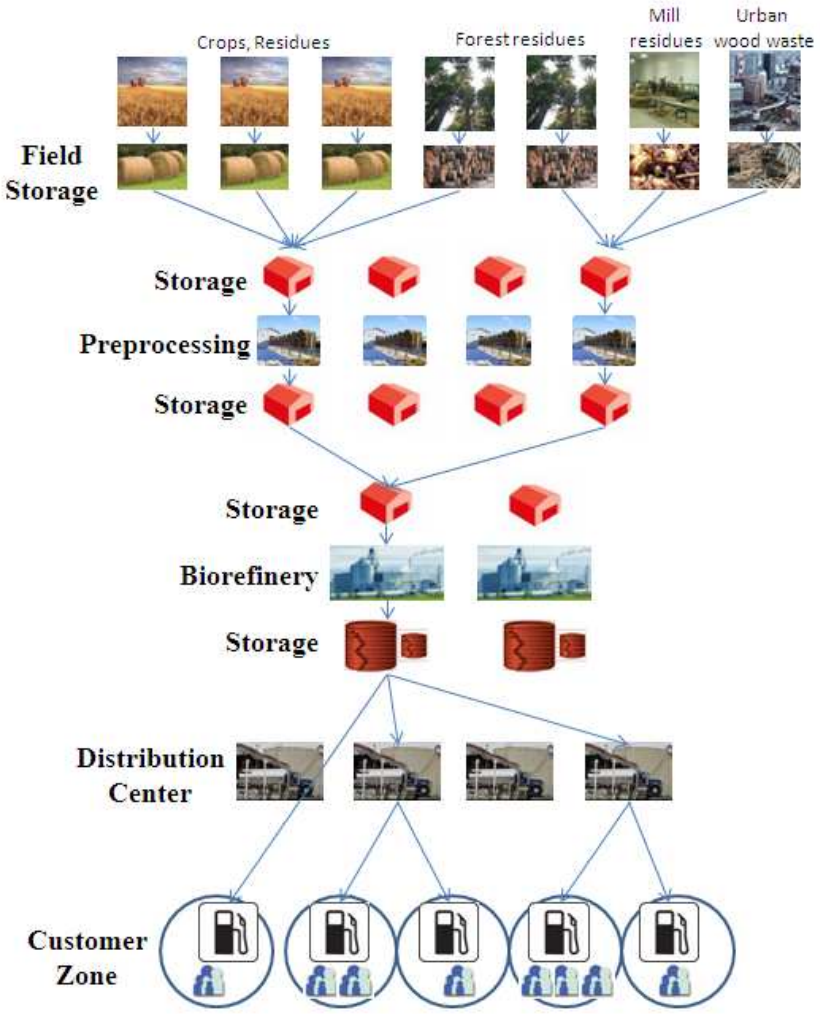

Figure 2 Schematic of facility alternatives of cellulosic biofuel SC system in time period t

The optimization problem is to prescribe strategic level decisions (i.e., locations where

facilities should be opened, and capacities and technology types of opened facilities) as well as a strategic plan for material flow, including single destination transfers in the upstream; and quantities transported, inventoried, and produced, while using a mix of different types of biomass as feedstocks and their changes from feedstocks to preprocessed products, and subsequently to biofuels. The objective is to maximize after tax profit of the total system from feedstock suppliers to customer zones. The model can be used to prescribe specific material flows in the tactical level at which the SC design has been fixed and accurate forecasts of demands, weather conditions, and other features are available to plan specific processing, transporting, and storage quantities over a planning horizon of intermediate duration.

Several assumptions will be invoked to structure the system: available feedstock supplies are given; the demand for biofuel in each customer zone and in each time period is known; preprocessors, refineries, distributors collaborate perfectly throughout the SC to maximize total system profit; preprocessing includes biomass drying and size-reduction operations, the main purposes of which are to reduce feedstock logistics cost (by transporting less water and denser biomass) and to promote conversion process efficiently; and the generated biofuel could be ETOH or drop-in-fuel. We assume that one of the most promising conversion technologies currently being developed is employed in the SC.

For the perspective of the operational level, simulation models evaluate the performance of several biomass logistics systems, including the bale system, which is one most commonly used for herbaceous (i.e., non-woody cellulosic) biomass; the silage system; and the module system, which is currently being developed at Texas A&M University. A biomass logistics system integrates several processing steps before conversion, including harvesting, collection, transportation, storage, and preprocessing. For a given set of input values (e.g., farm land area, biomass yield, weather conditions, moisture content, and harvesting schedule), several measures

(e.g., cost, energy, and CO_2 emissions) for each operation can be estimated based on a discrete event simulation model that reflects actual phenomena (e.g., operation delay due to weather conditions, physical and chemical dry matter loss, moisture content variation, and biomass density change).

1.3 Method of approach

This section describes the approach to achieve research objectives. Each of subsections corresponds with each of objectives that we introduce in section 1.1.

1.3.1 Research objective 1

Since the biofuel industry is just being initiated, a literature review of relevant SC issues has not yet been published. The biofuel industry shares features with the petroleum-based fuel industry in terms of market, but it also has some unique characteristics, especially related to the feedstock supply. More generally, certain prior work on generic SCs may be adapted to the biofuel SC.

Therefore, this literature review will analyze the state-of-the-art fuel SC research and compare biofuel and petroleum-based fuel SC studies. Also, it will relate research on generic SCs to the biofuel SC. Based on the perspective gained from the review, fertile avenues for future research will be recommended.

1.3.2 Research objective 2

The flow of biofuels in the downstream and of several kinds of biomass feedstocks in the upstream can be described as multi-commodity flow with commodity-type changes. Before preprocessing operations in the upstream, each commodity represents a combination of a biomass type and a moisture content; after a preprocessing operation, the commodity type is changed to preprocessed biomass (i.e., size-reduced and dried). In the downstream, commodities are biofuels, which are converted in refineries from preprocessed biomass. The multi-

commodities in the system must be processed, stored and transported, based on prescribed capacities of the processing and storage facilities that are opened, and on the available capacities of selected transportation routes. To deal with these features, two types of decision variables will be incorporated in the model: binary variables prescribe the opening of facilities and the single destination (i.e., facility) to which each facility in the upstream ships; and continuous variables, the capacities of facilities opened and the quantities of material flows (i.e., in processing, storage and transportation).

This paper formulates MIP model which is a time-staged, multi-commodity flow, network design problem with an objective of maximizing profit. The model invokes five types of constraints. First, each facility must employ only one technology type from among several alternatives. Second, the capacity of each opened facility will be restricted to or be less than the capacity limit associated with the corresponding location. Third, in order to facilitate management in the upstream, a constraint will be employed to assure that each field storage facility supplies a single open preprocessing facility, which, in turn, supplies a single open refinery. Fourth, the amount of materials flowing on each arc will be restricted by the capacity of that arc. Finally, flow balance equations will be employed to represent operations in each echelon: moisture content reduction in each preprocessing facility; chemical dry matter loss in each biomass storage facility; conversion from biomass to biofuel in each refinery; and material flow balance at each biofuel storage facility.

1.3.3 Research objective 3

This research will present a case study to demonstrate how managers can use the proposed models. According to the analysis of Milbrandt (2005) (see Figure 3), the northern, central and eastern regions of Texas have some available biomass that can be used to generate biofuel. The central region is selected as a test bed because it represents an example that does not

provide a sufficient amount of crop residue to meet its own biofuel needs so that some energy crops must be planted to meet demand. Moreover, since the central region has no first-generation biofuel refinery yet, a sensitivity analysis to investigate the significance of economic factors will be more focused than in the case that also involves some first generation refineries. For instance, a few first generation biorefineries are already operating in northern panhandle area of Texas.

A sensitivity analysis will be conducted relative to the selected region by evaluating relevant economic factors: feedstock price vs. biofuel price and feedstock supply vs. biofuel demand. Through the sensitivity analysis, we will be able to identify which factors are most economically significant to the biofuel SC system.

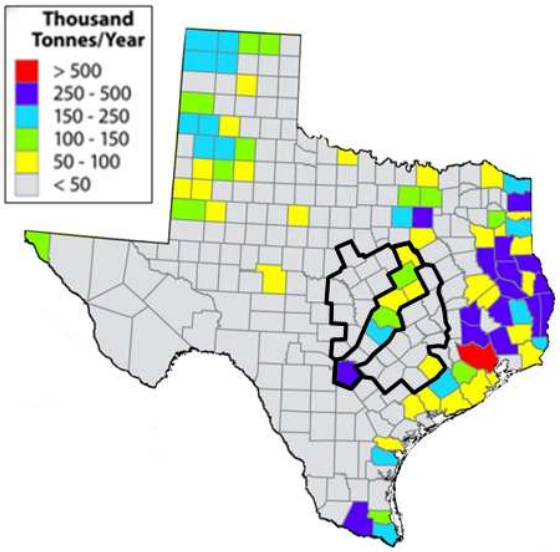

Figure 3 Available biomass distribution in Texas

1.3.4 Research objective 4

Through preliminary computational work, instances have been found to be very large, mainly due to the number of flow variables. This section describes a solution approach to deal efficiently with instances of large size.

(1) Reformulation. In the upstream, a commodity is defined by a combination of biomass type and moisture content. Since moisture content can vary according to biomass source, location and time period, considering a large area or many time periods would generate a huge number of commodity types. Therefore, if commodity type can be redefined by modeling a commodity that does not depend on moisture content, the formulation would be improved. To resolve this issue, this study provides a way to redefine the commodity type.

In addition, to simplify the model further, another technique, a node-arc split with respect to biomass types, has been employed. This technique transforms multi-commodity flow into single commodity flow on each split arc. A single commodity on each split arc can be considered simply as energy flow on it.

(2) Decomposition. Based on the reformulation, the embedded flow problem, which is an embedded generalized flow problem (GFP), can be decomposed into a subproblem. This study deals with the uncapacitated, embedded GFP in BSCP under a CG-based decomposition approach, which solves the linear relaxation of BSCP at the root node. To solve UEGFP effectively, this study develops a new dynamic programming algorithm, which is called backward reaching algorithm (BRA), generating improving path-flows (i.e., columns) effectively in $O(m)$. The master problem prescribes optimal, profitable flow quantities, considering side constraints.

(3) POC. This idea starts from the observation that the optimal solution opens facilities only if the system is profitable; and, accordingly, the fixed cost related to opened facilities could be restricted by available profit. Since the binary variable that prescribes the opening of each facility alternative is associated with its fixed cost, the relationship between fixed costs of opening facilities and profit would induce another relationship between binary variables and profit. And, possibly, the induced relationship could be used as an inequality to restrict binary

variables. This study develops a POC that is based on a portion of the objective function in MIP and may restrict the B&B search tree so that it can accelerate B&B procedure.

1.3.5 Research objective 5

As a precursor to the design and fabrication of the biofuel processes, evaluating the significance of several factors and their interactions in the biomass logistics system will help to optimize the SC design. The IBSAL simulation framework developed by the DOE is an efficient tool to evaluate biomass logistics operations. It calculates several performance measures for each operation (e.g., biomass yield, dry matter loss, costs, energy used, CO_2 emissions, and working hours). In addition, it provides several elements, each of which represents a unit operation, so that composing a new logistics model based on IBSAL is straightforward.

However, since the biomass module system involves some new types of processes that are not yet available commercially, IBSAL does not provide elements to represent them. Thus, this research develops IBSAL elements to simulate the operation and cost of processes that can form and handle biomass modules. Then, a simulation model is used to identify factors that are critical in achieving high performance and low cost of the biomass module system.

1.3.6 Research objective 6

To compare performances of alternative biomass logistics systems, including bale, silage and biomass module, simulation models for those systems are developed. This research investigates advantages/disadvantages of each biomass logistics system under various environments (e.g., crop yield and transportation distance).

1.4 Organization of the dissertation

This dissertation is organized in seven chapters. Chapter II reviews literature associated with the first objective, specifically about biofuel SC studies, comparing those to petroleum-based fuel SC studies. Chapter III formulates a MIP model that is a time-staged, multi-commodity

flow, network design problem and gives a case study in a region of Central Texas, addressing the second and third objectives. Chapter IV describes a solution method devised, which includes reformulation, CG for an uncapacitated, embedded GFP, and POC to deal with large size instances, and provides computational results, addressing the fourth objective. Chapter V addresses the fifth objective by developing new IBSAL simulation elements for the biomass module logistics system. Chapter VI compares performance measures between several biomass logistics systems by using IBSAL simulation models addressing the sixth objective. Finally, Chapter VII presents conclusions and some recommendations for future research.

CHAPTER II

LITERATURE REVIEW

This chapter reviews the literature related to SCs of biomass for liquid fuels and for cofiring in power plants, as well as petroleum-based fuel. The objectives of this review are (1) an analysis of the state-of-the-art fuel SC research, (2) a comparison of biofuel and petroleum-based fuel SC studies, (3) a comparison of the characteristics of the biofuel SC and the generic SC, and (4) a perspective from which we suggest fertile avenues for future research that would expedite biofuel commercialization.

The body of this chapter is organized in five sections. Sections 1 through 5 discuss SC studies related to biofuel and petroleum-based fuel according to the upstream, midstream, and downstream categorization. In particular, section 1 reviews operational-level studies, and section 2 reviews tactical-level studies of supply chain management (SCM) related to fuels. Section 3 reviews models that integrate strategic and tactical decisions in fuel SCM; and section 4 relates models capable of integrating strategic, tactical, and operational decisions in fuel SCM. Section 5 discusses other review papers that have dealt with the management of the generic SC. Section 6 analyzes fuel SCM.

2.1 Operational level studies in fuel SCM

Operational-level SC studies deal with decisions that affect the short term (e.g., hourly, daily, or weekly). Related biofuel SC studies have dealt with costs of operations, harvest scheduling, and upstream transportation. Studies of the petroleum-based fuel SC have dealt with refining operations in midstream.

2.1.1 Biofuel

Early studies of the operational level of biofuel SCs analyzed economic factors, for

example, estimating the cost of each operation from farm to conversion plant. The primary objective of most early studies was to assess the economic feasibility of the biofuel industry by estimating the cost of biomass logistics. Jenkins et al. (1984) estimated costs for alternative processes in biomass logistics, including collection and transportation of several types of biomass (e.g., rice, wheat, and barley straws; corn and sorghum stover; cotton stalks; orchard pruning; and forest slash). Gemtos and Tsiricoglou (1999) and Tatsiopoulos and Tolis (2003) provided a similar analysis, based on cotton stalk, that dealt with an electricity-generating plant in Greece. Petrou and Mihiotis (2007) reported a study for cotton-stalk biomass as a feedstock of biofuel, using another example in Greece.

Hamelinck et al. (2007) addressed international bioenergy logistics. They reported that, in Sweden and the Netherlands, several green-energy producers already import biomass, requiring the supply of long-distance biomass transportation. They analyzed SCs of Europe, including transportation of biomass from Latin American to conversion plants in Europe. They estimated the cost for each possible operation in such an international biomass SC. Gronalt and Rauch (2007) proposed a method to evaluate the total cost of supplying woody biomass from forest to conversion plants for a state within Austria, comparing central and local chipping alternatives. The system cost they calculated includes the costs of transportation from forest to terminals and from terminals to conversion plants and of operating terminals, but does not include harvesting costs.

Simulation models have been based on economic analyses to estimate important measures, including cost, energy consumption, and carbon emission. Mantovani and Gibson (1992) introduced a simulation model to analyze biomass harvesting and transportation costs by using SLAM simulation language. They considered three kinds of biomass: maize, low-quality hay, and wood chips. Gallis (1996) provided a simulation model also based in SLAM to estimate

logistics cost of forest biomass for several scenarios in Greece. De Mol et al. (1997) proposed both simulation and optimization models. Their simulation model analyzes all operations and calculates annual costs, energy consumption, and the flow of biomass from farm to conversion plant within a given network of facilities. They estimated dry matter and moisture losses. They classified the costs associated with biofuel in three categories: feedstock, logistics, and power plant. The main purpose of this simulation model was to assess the economic feasibility of the biofuel industry, but the paper does not provide a numerical example or a detailed mathematical formula for each simulation module.

Sokhansanj et al. (2006a) reported a similar, integrated biomass supply analysis and logistics (IBSAL) model. Their ExtendSim simulation model is similar to that of De Mol et al. (1997), but it calculates the carbon emissions that result from processing and transportation, and includes formulae that give good estimates of physical phenomena and logistics operations, including biomass availability, moisture content, weather factors affecting field operations, equipment performance, dry matter loss, and costs. They analyzed a numerical example that applied their model to corn stover collection and subsequent transportation in bales. In 2007, Kumar and Sokhansanj (2007) employed IBSAL to study switch grass logistics and compared several options for collection and transportation. They noted that the simulation approach is limited because it uses a given network structure, which specifies facilities and their capacities as well as transportation distances, which depend on facility locations.

Ravula et al. (2008) used a simulation model to study transportation in the cotton logistics network as a possible model for the biomass system. Typically, cotton is collected and then compressed in long blocks, known as cotton modules, for transportation. Then, cotton modules built by several farmers are transported to a gin for processing. To consider continuous cotton module supply in a biomass transportation system, they proposed a knapsack optimization

model with binary decision variables that prescribe truck schedules and module pickups. They used their simulation model to study a transportation system, investigating greedy strategies, such as “shortest first” or “longest first, shortest second,” to schedule transportation operations using a limited number of equipments that are specialized for harvesting cotton.

Two other important issues in biomass are maintaining a sustainable supply and assuring reduced environmental impact. Forests are a primary source of cellulosic biomass and related studies have been focused on harvest sustainability since the 1980s. Those studies are applicable but, since they focus on forest products rather than biofuel, this paper simply outlines recent trends in forest harvesting research. Major concerns associated with scheduling forest harvests include producing greatest benefits, achieving consistent and stable harvest yields, and reducing the environmental impact of required treatment operations. Murray (1999) reviewed two basic harvest-scheduling models: the unit restriction model (URM) and the area restriction model (ARM). These two models differ in just one constraint, which involves spatial restrictions. URM does not allow harvesting in an area that is adjacent to another one that has been harvested, while ARM limits the extent of contiguous harvesting areas in each cluster of areas. Most recent studies have focused on different solving methodologies to overcome the computational challenges of the ARM. Martins et al. (2005) posed a column-generation approach that solved sub-problems using heuristics. Gunn and Richards (2005) formulated a mixed integer program (MIP), including strengthening and lifting constraints, to improve the basic ARM formulation. Goycoolea et al. (2005) proposed a MIP based on the node packing problem. They devised an exact-optimizing algorithm that uses strong valid inequalities, which form clique representations of a projected constraint. Constantino et al. (2008) presented a new MIP model, which comprises polynomial numbers of variables and constraints, and used branch and bound to solve it.

Since first-generation biomass (e.g., edible crops, such as corn and sugar cane) has been used to produce ETOH, we also consider SC research in the agri-food area. A few papers have dealt with the sugarcane SC. Higgins (1999, 2002) formulated a MIP to improve yield by scheduling sugarcane harvests and solved it using several heuristics. Higgins and Postma (2004) addressed a scheduling problem to optimize the rostering of harvesting groups into sugarcane rail and road sidings to reduce transportation and harvesting costs. They used a tabu search to solve their MIP, which uses a weighted, multi-objective function, including transportation capacity, the total amount of movement across sidings, and a measure of adherence to a schedule that reflects perfect equity between farmers. Lejars et al. (2008) used the MAGI simulation package to plan sugarcane harvesting. Its main purpose was also to improve yield and, thus, profit.

2.1.2 Petroleum-based fuel

The first studies of the operational level of the petroleum-based fuel SC were reported in the 1960s. Aronofsky and Williams (1962) developed a multi-period, linear programming model to prescribe oil well production. They developed two models: one schedules production rates from either single or multi-well systems; the other, drilling and rig operations. Decision variables in the first model set of production rates for oil wells and those in the second model prescribe the number of wells completely drilled, the number of rigs purchased, and the number of rigs in operation.

Most operational-level studies of the petroleum-based fuel SC deal with the midstream (i.e., operations at refineries). Pinto and Moro (2000) developed a modeling framework for planning and scheduling refinery operations. Considering the market demand for oil derivatives typically supplied by a refinery in Brazil (e.g., metropolitan diesel, characterized by low sulfur levels for environment; regular diesel, used in areas with no special concern for atmospheric

pollution; and low-valued maritime diesel, featured with high flashing point for safety), their non-linear optimization model prescribes the amount of each oil derivative to be produced, while satisfying all process constraints. Neiro and Pinto (2004) extended this earlier work, providing a general modeling framework for the operational planning of sub-facilities within a refinery (e.g., processing units that modify the material physically or chemically, tanks for mixture and storage of the different feed streams, and pipelines for transportation of crude oil and products). Decision variables prescribe material flow through each process step. They incorporated detailed operational constraints to represent practical features associated with the refining process. Because they defined each outlet flow as the product of the associated feed flow rate and certain properties of the input, the resulting multi-period model is a large-scale mixed integer non-linear program (MINLP). Even though they proposed a decomposition scheme to solve the MINLP, they gave no further details or computational results.

Research has also addressed operational-level planning for downstream operations. Ronen (1995) addressed a scheduling problem associated with the distribution of petroleum products, considering the two basic types of plants: refineries and lube plants. Refineries produce light products (e.g. gasoline, kerosene, diesel oil, aviation fuel) as well as heavy products (e.g. base stock for lubes, and residual oil). Lube plants manufacture lube oils, greases, and waxes. Ronen explained four different types of operational environments encountered in practice: light products transported in bulk from refineries to tank terminals and industrial customers; light products, from tank terminals to retail outlets; bulk lubes, from lube plants to industrial customers; and packaged lubes, from lube plants to retail outlets and industrial customers. He proposed two scheduling formulations: set partitioning for minimizing cost and set packing for maximizing profit.

2.2 Tactical level studies in fuel SCM

Tactical level studies deal with planning decisions that focus on a somewhat longer time period (e.g., monthly) than addressed by operational level studies. Generally, those decisions prescribe inventory policy and material flow.

2.2.1 Biofuel

Apparently, no study has dealt with the tactical level of the biomass SC. We surmise this is because generic SC studies that predated early biomass SC research contributed strategic, tactical, and operational models as well as models that integrate the various levels. Some representative generic SC studies are introduced in section 6. In addition, the biofuel industry has been in a state of flux as it investigates several types of biomass and several possible processes to convert biomass to biofuel (e.g. chemical or thermochemical processes). Tactical-level studies will doubtlessly be undertaken once the conversion process is fully developed.

2.2.2 Petroleum-based fuel

In contrast, research in the petroleum-based fuel industry has studied tactical planning of inventory management and product flow. Catchpole (1962) proposed a linear program to plan the flow of crude oil from wells to refineries, and the flow of oil products from refineries to distribution centers. In addition, he analyzed the convexity and non-linearity of his model. For example, the use of lead to improve octane for motor gasoline has a decreasing effect as concentration increases, so that the objective function and constraints are non-linear. Capital and operating costs are concave non-linear functions. In addition, he noted that costs and demand must be considered as stochastic factors. He solved the problem using Dantzig-Wolfe decomposition. Klingman et al. (1987) formulated a production and inventory planning model for refineries. Their model, named the SDM (Supply Distribution Marketing) system, has been used by the Citgo Petroleum Corporation. Decision variables prescribe production amounts, spot

purchase amounts, oil product exchange with other companies, inventory, and distribution. This multi-period network flow planning model requires the network structure to be input.

Sear (1993) addressed the problem of distribution in the downstream segment of the oil SC. He dealt with three product classes: motor spirits (petrol), which have a low flashpoint; middle distillates (diesel fuel, oil for heating, etc.), which have high flashpoints; and fuel oils (black oils), which often require heated storage. His model prescribes flows of the various oil products, given a network structure. He also derived an equation to estimate delivery cost based on the assumption that delivery involves vehicle routing.

Escudero et al. (1999) formulated a model, named CORO, to prescribe crude oil supply, transformation in refineries, and distribution to customers for a given facility network structure, while considering demand and spot price uncertainties. They formulated a scenario-based stochastic programming model for which decision variables prescribe the supply volume of crude oil, the volume transformed at the refinery, the transported volume, the spot volumes supplied by other companies, the spot volumes sold to other companies at the depot, and the volumes of excess and deficit at the depot. Dempster et al. (2000) applied the CORO model and extended it to the multi-stage case. Through some numerical examples, they concluded that the multistage model is more robust in the face of uncertainty than the two-stage model; it requires significantly less memory, and provides more realistic cost estimates. In one of their examples, the deterministic equivalent to the multistage model comprised about 50% fewer constraints than the two-stage model.

Chenga and Duran (2004) developed a simulation model for planning worldwide crude oil transportation and a Markov decision model for the stochastic optimal control of the inventory/transportation system. They solved their model using an approximation algorithm based on problem decomposition and function approximation to deal with the large state-space

involved. Lababidi et al. (2004) formulated a production planning model for a petrochemical company. For a given facility network, they dealt with raw material procurement and production costs, as well as lost demand, backlogging, transportation, and storage penalties. Their model includes operational level constraints related to refinery processing and uses a two-stage stochastic program to deal with uncertain demands, market prices, raw material costs, and production yields. MirHassani (2008) also formulated a scenario-based stochastic program to plan the distribution of petroleum-based oil products in the face of uncertain demand. Decision variables prescribe the volumes transported by each transportation mode and the volumes of shortage and excess of each product.

A study by Levary and Dean (1980) proposed a SC model for another form of energy (natural gas) and it may be applicable to biofuel. Their multi-period, multi-objective model deals with demand uncertainty caused by weather variability. The duration of each time period varied according to the variance of demand. Their model incorporates lead time from supply to destination, both through constraints and decision variables, which prescribe the amount of gas flow. Within one period (at least one month in duration), some equations are based on minor (daily) time periods to incorporate the daily rate of flow and the maximum possible rate of flow. This device allowed the authors to integrate time-period mismatches to model different operations.

2.3 Models for integrating strategic and tactical decisions in fuel SCM

Models that integrate strategic and tactical decisions typically deal with both types of decisions and relate to a monthly or yearly time frame. The production and distribution problem is one example of such an integrated model.

2.3.1 Biofuel

Most biofuel SC models deal simultaneously with the strategic and tactical levels of

upstream operations (i.e., from farms to the refinery). Several studies have built upon a classical production/distribution MIP model. De Mol et al. (1997) introduced a MIP that prescribes plant openings as well as annual flows of biomass. They compared advantages and disadvantages of simulation and optimization models. In 2004, Gunnarsson et al. (2004) extended the classical production/distribution model to the forest-fuel industry. They considered preprocessing alternatives and limited the used of less attractive biomass in order to produce forest fuels of desired quality. A multi-period MIP for forestry production and logistics formulated by Troncosoa and Garrido (2005) prescribes discrete capacity decisions for forest processing facilities **(**e.g., sawmills, board plants, and remanufacture plants) as well as material flow decisions, considering harvesting and forest-area limitations.

Huang et al. (2010) developed a multistage model for strategic biofuel SCM from biowaste feestock fields to end users to determine locations and sizes of new refineries, additional capacities, and quantity of material flows on a yearly basis. Whereas, several studies have dealt with a one-year planning horizon, considering loss of dry mass over time. Ekşioğlu et al. (2009) proposed a SC design model that uses corn and corn stover biomass to produce ETOH. Their multi-period MIP prescribes the network design as well as material flows from the upstream to the downstream. In 2010, Ekşioğlu et al. (2010) extended their study by considering modes of transportation. Zhu et al. (2010) formulated a MIP to prescribe locations of biomass storage and conversion facilities, modes of transportation from farms to refineries, and flows of biomass in the upstream.

Several studies have dealt with uncertainty in the biofuel SC. Cundiff et al. (1997) developed a two-stage stochastic program to manage production uncertainty caused by weather over multiple time periods. Their model prescribes optimal storage capacities, including installation of new, or expansion of existing, facilities. Even though they considered some

strategic design aspects, the main decisions, which deal with uncertainty, are associated with the tactical level.

Gigler et al. (2002) applied an agri-product SC study to the biofuel SC, considering biomass as a specific type of agri-product. Their dynamic programming model for the agri-product SC deals with the characteristic change of quality and appearance over time. Moisture content is used as a measure of quality and particle size represents appearance. Before arriving at a refinery, biomass should be in an appropriate state as measured by moisture content and particle size. They gave a numerical example that illustrates the delivery of willow biomass to energy plants. They modeled processing steps from harvesting to conversion plant, calculating quality, appearance, and costs to prescribe an optimal processing path.

Dunnett et al. (2007) addressed a modeling framework for the upstream SC of biomass used for combustion plant, discussing system components that must be represented appropriately (e.g., SC structure, storage strategy, and task schedule). They formulated a MIP, which is based on a state-task network, to prescribe the selection of structural components, the number of units installed on components, and assignment of task to units. Through a numerical example of a 20 megawatt (MW) peak output heat plant with Miscanthus as a feedstock, they showed that optimization solutions result in 5-25% cost reduction compared to a simple heuristic strategy.

2.3.2 Petroleum-based fuel

The first model that integrated strategic and tactical planning for the petroleum-based fuel SC appeared in the early 2000s. Brimberg et al. (2003) formulated a model to design oil pipelines running from onshore wells to a port. Decision variables prescribe, for a single time period, strategic decisions (which pipeline segment to open and the capacity of each segment) and tactical decisions (the amount of flow in each segment). They solved their model based on a decomposition scheme, using an interactive B&B procedure, which incorporates several

techniques at each branching node, including applying augmented valid inequalities and either tabu search or variable neighborhood search. They applied their approach to an example based on a south Gabon oil field.

Most studies of downstream operations have dealt with designing the network (strategic) and prescribing material flows (tactical). Al-Qahtani and Elkamel (2008) formulated a model to design processes and integrate production capacity expansions in a multiple refinery complex using different feedstocks, each with a unique chemical composition. Each refinery consists of different production units that can operate in different operating modes (i.e., processes), which depend on targeted production quantities. Their single-period MIP deals with strategic capacity expansion, transshipments between refineries, and tactical flows of crude oils and intermediate materials between refineries (e.g., heavy naphtha is transported to a catalytic reformer; and light naphtha, to a light naphtha pool or an isomerization unit), considering practical features associated with refining processes. Their numerical example deals with three complex refineries in one industrial area, a common configuration, and showed that the integrated planning of refineries in an area is economically attractive in comparison with decentralized management. Khor et al. (2008) formulated a petroleum refinery planning model that addresses uncertainties associated with demand, price and yield. They formulated four models to hedge against uncertainty: (1) Markowitz's mean-variance (MV) model, which minimizes variance; (2) two-stage stochastic program to deal with randomness of constraint coefficients; (3) mean–risk model, which incorporates the MV model within the framework developed in (2); and (4) a reformulation of the mean–risk model in (3) that adopts the mean-absolute deviation (MAD) as the measure of operational risk imposed by recourse costs. Decision variables prescribe production capacities, which can vary over time; material flows; and inventories for each processing operation. They used three scenarios in a numerical example and analyzed solutions

prescribed by the four models. Kim et al. (2008) formulated a model to integrate strategic supply-network design and tactical production planning for several products (e.g., gasoline and diesel) at the downstream level. Their distribution network model combines a network design model and a production planning model for multi-site refineries. They coupled their MIP with the non-linear production planning model of Li et al. (2005), which deals with operations at a crude distillation unit and a blender. Importantly, through several examples, they showed that a model that integrates strategic and tactical decisions can improve profit in comparison with using separate models at individual refineries.

2.4 Models for integrating strategic, tactical, and operational decisions in fuel SCM

To our knowledge, no available models integrate biofuel and petroleum-based fuel SCs in all decision levels. However, Fiedler et al. (2002) provided a conceptual study. Using a strategic model based on a geographic information system (GIS) to design cost-efficient supply logistics for the industrialized use of biomass, they analyzed options available to managers, including choice of sources of biomass (e.g. owned or contracted supplier) and types of biomass, assessment of necessary logistics processes (e.g., pre-processing, storage, and transportation mode), and design of the transportation network.

2.5 Trends in generic SCM

A rather extensive literature has addressed SC management over last two decades. Since the biofuel SC is a specific type of SC, much of the work on generic SCs may apply. Here, we discuss several recent review papers to identify research trends related to SCM.

Over the last two decades, information technology (IT) has been improved and has become a major driving force for SC innovations. Min and Zhou (2002) reviewed SC models, using four major categories: deterministic, stochastic, hybrid, and IT-driven models. IT-driven models aim to integrate and coordinate various phases of SC planning on a real-time basis,

including models of warehouse management systems, transportation management systems, collaborative planning and forecasting replenishment, material requirement planning, enterprise resource planning, and GIS. Bilgen and Ozkarahan (2004) reviewed strategic, tactical, and operational models for production and distribution. They classified papers in terms of the solution methodology used: optimization-based, metaheuristic, IT-driven, and hybrid models. They suggested that future research should address multiple conflicting objectives, stochastic factors, and other issues, including bill of materials, fuzzy constraints, logical constraints, operational decisions, and reverse SCs. Many IT-driven models (e.g., real time planning and management based on information systems) are applicable to the biofuel SC. In particular, GIS is very useful in dealing with dispersed biomass supply locations. Graham et al. (2000) employed GIS to calculate exact transportation distances and costs for supplying switchgrass in eleven US states.

Another important trend relates to sustainable SCM. Sustainable products provide environmental, social, and economic benefits, while protecting public health and welfare, and the environment over their life cycles. Seuring and Muller (2008) offered a review of the literature on sustainable SCM from 1994 to 2007. They described distinctive features of sustainable SCM, which must consider the SC in a long-term perspective; risk management; and a wider set of performance objectives, including economical, environmental and social benefits. Sustainable SCM entails a need for companies to collaborate. They proposed that future research should identify shortcomings of existing models since sustainable development is primarily one-dimensional, focusing on environmental issues. A perspective that integrates social issues (e.g., partnership issues related to human rights and working conditions between a company and its suppliers) is required. Most papers report empirical research, so that theoretical developments are needed. Biofuel is a potentially sustainable product, so generic sustainable SCM studies

could be specialized to deal with the sustainability and environmental friendliness of biofuel SCM. The impact of using land for energy production (rather than food) is affecting government policy toward biofuels, and studies of sustainable SCM may be able to address related concerns.

Due to issues related to public health, agri-food SCs have been the focus of a significant amount of recent research (Wilson, 1996; Ahumada and Villalobos, 2009). Ahumada and Villalobos (2009) reviewed the literature related to agri-food supply chains (ASCs). They compared SC studies that deal with both perishable and non-perishable agri-products, considering their objectives, planning scopes, decision variables, and modeling approaches. They concluded that the state-of-the-art planning models for ASCs lags that in other industries, including electronics and automotive. They recommended that future research should address the perishability of products and formulated operational models that integrate production and distribution decisions. Perishability is common to agri-products in general and to biomass in particular, and it is very important aspect of SCM. Problems that involve perishable inventories, one important part of SCM, have been studied for several decades. Goyal and Giri (2001) reviewed deteriorating inventory models proposed since the early 1990s. They classified perishable inventory models depending upon the life time of the product: fixed, random (exponential decay), and decay corresponding to the proportion of inventory decrease per unit time. Thus, some aspects of agri-product SCM may be applicable to biofuel SCM. In addition, other generic agricultural studies, such as farm planning (1987) and agri-product production planning (1998), also may be directly related to biomass productivity.

Facility location is a classical problem that is a component in the strategic planning of SCs. Klose and Drexl (2005) reviewed facility location models. They provided basic formulations for representative models, including a continuous location model, a network location model, uncapacitated/capacitated single/multi-layer models, a dynamic model, and a

probabilistic model. Melo et al. (2009) provided a comprehensive review of the facility location problem. They concluded that capacity, inventory, and production decisions are important in addition to location-allocation decisions. Even though procurement, transportation-mode selection, and routing are important, not many papers have considered them. They also observed the recent trend to incorporate risk management in the design of SCs. Relatively few studies have been able to address uncertainty (Birge and Louveaux, 1997) related to generic (Santoso et al., 2005), petroleum-based fuel (Escudero et al., 1999), or biofuel supply chains (Cundiff et al., 1997). Melo et al. (2009) concluded that there is still a lack of research that addresses multi-period, multi-layer, stochastic models, noting that most papers deal with a single-period and assume a deterministic environment. Since biomass has very low energy density and is geographically dispersed, the structure of the facility network must be optimized so that the biomass SC can contribute to the economic viability of the industry. It may be possible to ameliorate the computational challenges posed by the multi-level structure of the biofuel SC through applying results obtained by generic facility location studies.

With a trend towards globalization, international issues have been major topics of concern. Goetschalckx et al. (2002) reviewed the literature on modeling and designing global logistics systems, focusing on the savings potential generated by integrating the strategic design of global SC networks with tactical production-distribution decisions. They analyzed the international features represented in strategic logistics models (e.g. stochastic features, taxation, transfer prices, and trade barriers) and formulated a model to prescribe transfer prices in global logistics systems. Meixell and Gargeya (2005) provided a comprehensive review of global SC design since 1980. They analyzed the literature based on four dimensions: decision variables, performance measures, SC integration, and globalization considerations. They concluded that several industries (i.e., food and medication) have not been explored in the context of global SC

models, while others (e.g. electronics manufacturing, apparel, fiber and textile, and automotive) have been studied extensively. In addition, they suggested that future research should focus on multi-tier SCs with both internal production sites and external suppliers. While most countries are requiring more extensive use of biofuel, not all countries around the world have land and environment capable of producing necessary quantities of feedstock for biofuel, so that an international market is likely to develop for biofuel. Moreover, even though the energy density of biomass is too low to offset the energy needed to transport it over a long-distance, some countries in Europe (e.g. Netherlands and Sweden) import biomass (e.g., wood pellets) from South America because the European Union has rules requiring member countries to generate 20% of their electricity from renewable sources by 2020. Their demands have led some companies in the U.S. to export pellets to Europe (e.g., Green Circle Bio Energy Inc., Dixie Pellet LLC, Phoenix Renewable Energy LLC) (The Wall Street Journal, 2009b).

Due to numerous efforts in SCM research, several sub-areas have evolved in SCM. Kouvelis et al. (2006) provided a comprehensive review of all topics addressed in SCM studies, including SC design, uncertainty, the bullwhip effect, contracts and SC coordination, capacity, sourcing decisions, applications, and the practice and teaching of SCM. If the biofuel industry matures and needed interactions between entities (e.g., supplier, manufacturer, buyer, and transporter) evolve, these topics will have to be investigated to deal with these complicated phenomena as they relate to biofuel.

2.6 Discussions

We categorize each paper examined according to two dimensions (i.e., SC level and decision level) as shown by Figure 4 for biofuel and Figure 5 for petroleum-based fuel. Figures 4 and 5 show the number of papers we review in each category and Tables 1 and 2 give taxonomies of these papers.

It is noteworthy that biofuel studies predominantly relate to upstream processes (i.e, from biomass suppliers to conversion plants) especially at operational and integrated levels. The reason is that prior research has focused on evaluating various types of feedstock and improving the efficiency of feedstock logistics. A second observation is that no study optimizes conversion operations in midstream because several technologies are still being developed and tested. Virtually, no study has dealt with the downstream and recently only a few studies have dealt with all echelons from the upstream to the downstream, likely because downstream processes for biofuel may be similar to those for petroleum fuel. In practice, some biofuels (i.e., so called drop-in fuels) can be distributed within the existing infrastructure for petroleum-based fuels, while other biofuels (e.g., ETOH) must be handled separately from petroleum-based fuels.

In contrast, studies of petroleum-based fuel have addressed upstream, midstream, downstream, and mid/downstream at operational, tactical, and integrated levels. Studies that have focused on planning at the operational level have addressed upstream, midstream and downstream levels of the SC. The most frequently studied combination has dealt with the mid/downstream (i.e., refining and distribution) and the tactical level of production planning. Integrated planning models have been formulated for upstream (pipeline layout design and material flow, refinery facility location or capacity expansion), midstream (intermediate product flow decision between sub-facilities within a refinery, sub-facility network design, and refinery production planning) or mid/downstream (refining and distribution decisions). More aspects of the petroleum-based fuel SC have been addressed than of the biofuel SC, apparently because the later is a developing industry that has only recently been promoted by government policies.

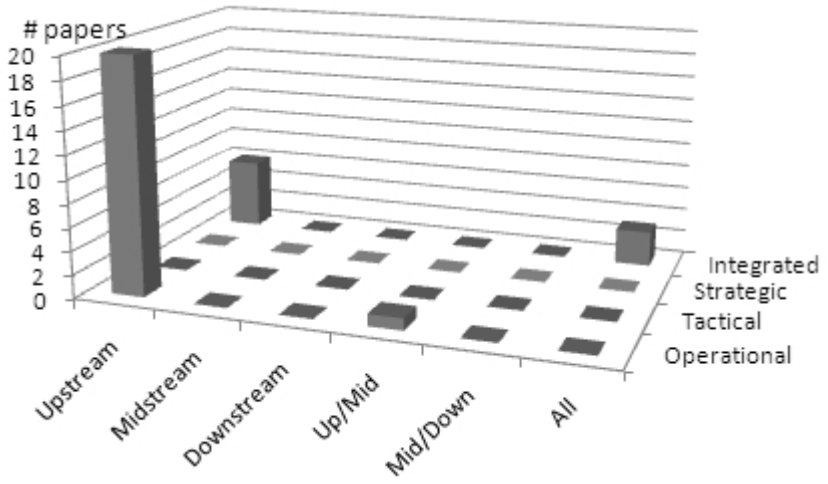

Figure 4 Categorization of papers related to biofuel SCM papers

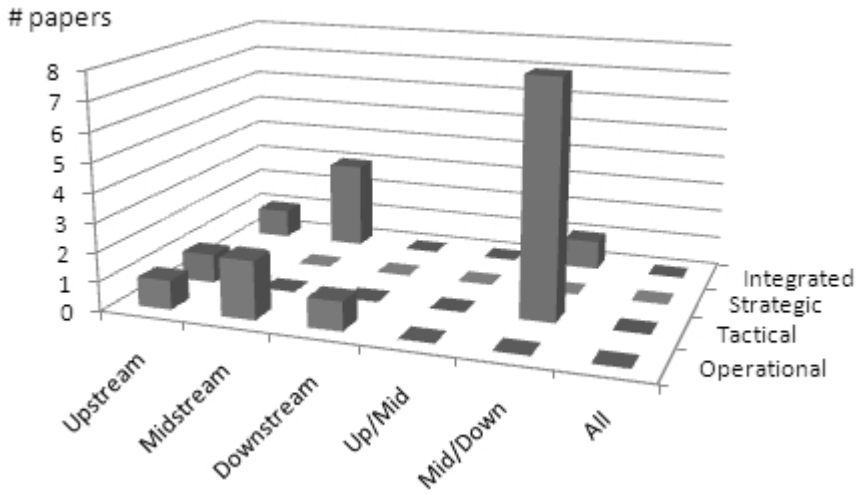

Figure 5 Categorization of papers related to petroleum-based SCM papers

Table 1 Taxonomy of biofuel SCM papers

SCM planning level	Layer in SC	Year	Researchers
Operational	Upstream	1984	Jenkins et al.
		1992	Mantovani and Gibson
		1996	Gallis
		1997	De Mol et al.
		1999	Gemtos and Tsiricoglo
		1999	Murray
		2002	Higgins
		2003	Tatsiopoulos and Tolis
		2004	Higgins and Postma
		2005	Goycoolea et al.
		2005	Gunn and Richards
		2005	Hamelinck et al.
		2005	Martins et al.
		2006	Sokhansanj et al.
		2007	Gronalt and Rauch Peter
		2007	Kumar and Sokhansanj
		2007	Petrou and Mihiotis
		2008	Constantino et al.
		2008	Lejars et al.
		2008	Ravula et al.
	Up/Midstream	1999	Higgins
Integrated	Upstream	1997	Cundiff et al.
		2002	Gigler et al.
		2004	Gunnarsson et al.
		2005	Troncosoa and Garrido
		2007	Dunnett et al.
		2010	Zhu et al.
	All	2009	Ekşioğlu et al.
		2010	Ekşioğlu et al.
		2010	Huang et al.

Table 2 Taxonomy of petroleum-based fuel SCM papers

SCM planning level	Layer in SC	Year	Researchers
Operational	Upstream	1962	Aronofsky and Williams
	Midstream	2000	Pinto and Moro
		2004	Neiro and Pinto
	Downstream	1995	Ronen
Tactical	Upstream	2004	Chenga and Duran
	Mid/Downstream	1962	Catchpole
		1980	Levary and Dean
		1987	Klingman et al.
		1993	Sear
		1999	Escudero et al.
		2000	Dempster
		2004	Lababidi et al.
		2008	MirHassani
Integrated	Upstream	2003	Brimberg et al.
	Midstream	2005	Li et al.
		2008	Al-Qahtani and Elkamel
		2008	Khor et al.
	Mid/Downstream	2008	Kim et al.

CHAPTER III

MIP MODEL AND A CASE STUDY ON A REGION IN CENTRAL TEXAS

This chapter formulates a model to maximize the profit of a cellulosic biofuel SC ranging from feedstock suppliers to biofuel customers. The model deals with a time-staged, multi-commodity flow, production/distribution system, prescribing facility locations and capacities, technologies, and material flows. A case study based on a region in Central Texas demonstrates application of the proposed model to design the most profitable biofuel SC under each of several scenarios. A sensitivity analysis identifies that ETOH price is the most significant factor in the economic viability of a cellulosic biofuel SC.

This study holds two primary research objectives. The first is to formulate a mathematical model to prescribe an optimal biofuel SC that allows use of various types of cellulosic biomass and deals with upstream and downstream material flows. The second is to apply the model in a case study to demonstrate its use in providing decision support for industry managers and government officials.

The body of this chapter comprises three sections. Section 1 describes the system and section 2 presents our mathematical model. Section 3 provides a case study based on a region in Central Texas. Section 4 analyzes impacts of several economic factors based on computational results and gives recommendation of future research. Finally, Section 5 gives conclusions.

3.1 System description

The biofuel SC system considered comprises five echelons: feedstock production, preprocessing, production in conversion plants, distribution, and consumption by customers, and including possible storage locations. Each facility can use one of several technology alternatives. For example, biomass can be stored using outdoor-uncovered, outdoor-covered, indoor-aerobic,

or indoor-anaerobic technologies. Preprocessing technology could include size reduction, drying moisture content, or both. Moreover, conversion technology may involve a biochemical, a thermochemical, or a bio-thermochemical process. Even though improving technologies and efficiencies in each echelon is important, integrating technologies and coordinating echelons is necessary for the system to be most successful economically.

Materials flowing in the SC must be stored before being processed either at preprocessing or conversion facilities, and again stored as they wait to be transported after processing. While it is being stored at upstream locations, biomass degrades over time, losing some portion of its mass due to chemical reactions (e.g., fermentation and breakdown of carbohydrates) (Sokhansanj et al., 2006b). The rate of dry matter loss in storage depends on the type of biomass, moisture content, and storage conditions.

Some feedstocks contain high moisture content and must be dried on the field immediately after harvesting and/or in a preprocessing facility to reduce the cost of transporting it and to meet requirements of the conversion technology selected. Since cellulosic biomass typically has low energy density, it is important to reduce moisture content so that energy is not expended in transporting it. In particular, transportation routes must be carefully prescribed so that the system is able to achieve a net production of energy while managing green house gas emissions. Preprocessing facilities may also involve a size reduction operation to reduce the transportation cost by increasing density and to facilitate the conversion process.

In the upstream, the various types of feedstocks must share the capacities of transportation vehicles and processing facilities. In the downstream, some biofuels (e.g., ETOH) must be transported and/or stored separately from petroleum-based fuels, while other biofuels (i.e. so called drop-in fuels) are compatible with petroleum-based fuels and can be handled easily within the existing infrastructure.

The variability of several important factors over time (e.g., the seasonality of biomass availability, biomass moisture content, and the demand and price of fuel) could significantly affect the SC design so that a time-staged model is required. A strategic model, which deals with a long-range forecast of demand can capture such dynamics using a time period of one quarter (i.e., 3 months) duration.

We invoke several assumptions to structure the system: available feedstock supplies are known; the demand for biofuel in each customer zone in each period is known; preprocessors, refineries, and distributors collaborate perfectly in all operations from field storage to customer zones to maximize total profit; preprocessing includes drying and size-reduction operations. We model the types of preprocessing technologies that appear to be the most attractive among the ones currently under development. Since the energy density of biomass is relatively meager, it cannot be transported over long distances if the SC is to result in a net production of energy. Further, transport over longer distances increases green house gas emissions. We invoke the single destination assumption to reflect the needs to result in net energy production and manage green house gas emissions as well as to promote management efficiency.

3.2 Mathematical model

We use a multi-commodity flow model to represent several kinds of biomass feedstocks in the upstream and of biofuels in the downstream. In the upstream, each commodity represents a combination of a feedstock type and a range of moisture content. Commodities must be processed, stored and transported, based on prescribed capacities of processing plants and storage facilities, and on the available capacities of transportation routes. The capacity of a storage facility is the maximum amount of biomass or biofuel it can store; that of a preprocessing (conversion) facility is the maximum amount of biomass (biofuel) that it can process in a year.

Our model prescribes two types of decisions variables: binary variables select facilities

to open and arcs representing routes; and continuous variables prescribe the capacities of open facilities and the quantity of material flow on each arc. Table 3 shows notations used in the formulation.

Table 3 Notation

Sets

A : Directed arcs : $= A^{+}_{frt} \cup A^{inv}_{frt}$

A^{+}_{frt} (A^{-}_{frt}): Directed arcs in period *t* that start or end at node *frt*

A^{inv}_{frt} : Arc that represents inventory held at facility *f* of type *r* from period *t* to period *t+1*

F_l : Candidate locations for facilities in echelon l or feedstock supply site (F_{F1}, F_{F2}) or customer zone (F_{CZ}), $l \in L$

F_{UP} : Upstream facilities : $= F_{F1} \cup F_{F2} \cup F_{P1} \cup F_{P2} \cup F_{P3} \cup F_{R1}$

F_{DOWN} : Downstream facilities := $F_{R2} \cup F_{R3} \cup F_{DC} \cup F_{CZ}$

F : All facilities : $= F_{UP} \cup F_{DOWN}$

F_{WH0} : Warehouses where biomass is held before preprocessing : $= F_{F2} \cup F_{P1}$

F_{WH1} : Warehouses where biomass is held : $= F_{WH0} \cup F_{P3} \cup F_{R1}$

F_{WH2} : Warehouses where biofuel is held : $= F_{R3} \cup F_{DC}$

F_{WH}: Warehouses : $= F_{WH1} \cup F_{WH2}$

F_{PR} : Process facilities (preprocessing, refinery) : $= F_{P2} \cup F_{R2}$

F_{OP} : Operating facilities : $= F \backslash (F_{F1} \cup F_{CZ})$

K_1: Feedstock commodities := $\{(f, t, p)\}, f \in F, t \in T, p \in P_f$

K_2: Biofuel commodities := $\{e\}, e \in E$

K: Commodities := $K_1 \cup K_2$

L : Echelons, $\{F1, F2, P1, P2, P3, R1, R2, R3, DC, CZ\}$ (*F1*: farm, *F2*: field storage, *P1-3*: preprocessing facilities, *R1-3*: conversion facilities, *DC*: distribution center, and *CZ*: customer zone)

P_f : Feedstock(biomass) types raised or gathered at $f \in F_{F1}$

R_f : Types of technologies at facility *f*, $f \in F_l$

T : Time Periods (monthly)

Indices

a : arc, $a \in A$

e : biofuel type, $e \in E$

f : facility, $f \in F$

k : commodity type, $k \in K$

l : layer, $l \in L$

p : biomass type, $p \in P$

Table 3 Continued

r : technology type, $r \in R$

t : time, $t \in T$

Parameters

C_a^T : Fixed cost associated with arc *a* (fixed transportation cost on a transportation arc, fixed holding cost on an inventory arc) (dollar)

C_{fr}^O : Fixed cost of opening facility *f* of type *r* (dollar)

C_k : Cost of commodity *k*, *k* => feedstock *p* from farm *f* in period *t* (dollar)

D_{kft} : Demand of end product *k* in customer zone *f* in period *t*, *k* := *e* (liter)

P_{kft} : Price of biofuel type k in customer zone *f* in period *t* (dollar)

V_a^T : Variable cost for a unit of flow on arc *a* (variable transportation cost on transportation arc, variable holding cost on inventory arc) (dollar)

V_{fr} : Variable cost per unit of capacity of opening facility *f* of type *r* (dollar)

Q_a^T : Flow capacity of arc *a* (Mg/single period)

Q_f^F : Capacity of facility *f* (biomass storage (Mg), preprocessing (Mg), refinery (liter), or biofuel storage (liter))

Q_k^F : Supply capacity of commodity *k* at farm *f* during period *t* for feedstock type *p* (Mg)

δ_k : Moisture content of commodity *k* (decimal fraction)

$\alpha_{k_1k_2fr}$: Amount of biofuel k_2 produced from one unit of pre-processed feedstock k_1 at conversion plant *f* using technology *r* (decimal fraction)

γ_{kfr} : Chemical dry mass loss rate (fraction) of feedstock *k* held at warehouse *f* of storage type *r*

Decision Variables

q_{fr} : Capacity of facility *f* of type *r*, $f \in F_{OP}, r \in R_f$

x_{fr} : 1 if facility *f* of type *r* is open, 0 otherwise, $f \in F_{OP}, r \in R_f$

y_a: 1 if arc *a* is used, 0 otherwise, $a \in A$

z_{ka} : Flow amount of commodity *k* on arc *a*, $k \in K, a \in A$

Model P describes the formulation:

$$\mathbf{P}: Z^* = \text{Max} \sum_{k\in K_2}\sum_{f\in F_{CZ}}\sum_{t\in T}\sum_{a\in A^-_{frt,}r=0} P_{kft}z_{ka} - \sum_{f\in F_{OP}}\sum_{r\in R_f}\left(C_{fr}^O x_{fr} + V_{fr}q_{fr}\right) - \sum_{a\in A} C_a^T y_a - \sum_{k\in K}\sum_{a\in A} V_a^T z_{ka} - \sum_{k\in K}\sum_{f\in F_{F1}}\sum_{t\in T}\sum_{a\in A^+_{frt},r=0} C_k z_{ka} \quad (3.1)$$

s.t.

$$\sum_{r\in R_f} x_{fr} \le 1 \qquad f \in F_{OP} \quad (3.2)$$

$$q_{fr} - Q_f^F x_{fr} \le 0 \qquad f \in F_{OP}, r \in R_f \quad (3.3)$$

$$\sum_{a\in A^{+}_{frt}} y_a \le 1 \qquad f \in F_{P1} \cup F_{R1}, r \in R_f, t \in T \qquad (3.4)$$

$$\sum_{k\in K} z_{ka} - Q^{T}_{a} y_a \le 0 \qquad f \in F, r \in R_f, t \in T, a \in A^{+}_{frt} \qquad (3.5a)$$

$$\sum_{k\in K} z_{ka, a\in A^{inv}_{frt}} - q_{fr} \le 0 \qquad f \in F_{WH}, r \in R_f, t \in T \qquad (3.5b)$$

$$\sum_{k\in K} \sum_{a\in \mathrm{A^{+}_{frt}}} z_{ka} - q_{fr} \le 0 \qquad f \in F_{PR}, r \in R_f, t \in T \qquad (3.5c)$$

$$\sum_{a\in \mathrm{A^{+}_{frt}}} z_{ka} \le Q^{F}_{k} \qquad k \in K, f \in F_{F1}, r \in R_f, t \in T \qquad (3.6a)$$

$$\sum_{a\in A^{+}_{frt}} z_{ka} - \sum_{a\in A^{-}_{frt}} (1-\delta_{\mathrm{k}}) z_{ka} = 0 \qquad k \in K_1, f \in F_{P2}, r \in R_f, t \in T \qquad (3.6b)$$

$$\sum_{a\in A^{+}_{frt}} z_{ka} + z_{ka, a\in A^{Inv}_{frt}} - \sum_{a\in A^{-}_{frt}} z_{ka} - \left(1-\gamma_{\mathrm{p_1 fr}}\right) z_{ka, a\in A^{Inv}_{frt-1}} = 0 \qquad k \in K_1, f \in F_{WH1}, r \in R_f, t \in T \qquad (3.6c)$$

$$\sum_{a\in A^{+}_{frt}} z_{k_2 a} - \sum_{k\in K_1} \sum_{a\in A^{-}_{frt}} \alpha_{\mathrm{k_1 k_2 fr}} z_{k_1 a} = 0 \qquad k_2 \in K_2, f \in F_{R2}, r \in R_f, t \in T \qquad (3.6d)$$

$$\sum_{a\in A^{+}_{frt}} z_{ka} + z_{ka, a\in A^{Inv}_{frt}} - \sum_{a\in A^{-}_{frt}} z_{ka} - z_{ka, a\in A^{Inv}_{frt-1}} = 0 \qquad k \in K_2, f \in F_{WH2}, r \in R_f, t \in T \qquad (3.6e)$$

$$\sum_{a\in A^{-}_{frt}} z_{ka} \le D_{kft} \qquad k \in K_2, f \in F_{CZ}, r \in R_f, t \in T \qquad (3.6f)$$

$$x_{fr} \in \{0,1\} \qquad \mathrm{f} \in F_{OP}, r \in R_f \qquad (3.7a)$$

$$y_a \in \{0,1\}, \qquad a \in A \qquad (3.7b)$$

$$q_{fr} \ge 0 \qquad f \in F_{OP}, r \in R_f \qquad (3.7c)$$

$$z_{ka} \ge 0 \qquad k \in K, a \in A \qquad (3.7d)$$

Objective (3.1): Maximize the present worth of total system profit, defined as the discounted revenue earned from selling biofuels in the customer echelon minus all discounted costs, including the fixed cost of capital for each facility opened and variable costs associated with operating facilities, purchasing feedstocks, carrying inventory, and transporting biomass and biofuel

Constraint (3.2): At most, one technology type can be selected for each facility

Constraint (3.3): If facility *f* is opened, the amount of flow out of it is restricted by its capacity; otherwise, the facility can sustain no flow

Constraint (3.4): Each field storage (preprocessing) facility must use a transport link to a single destination preprocessing (refinery) facility

Constraint (3.5.a): If arc a is selected, the flow amount is restricted by the arc capacity; otherwise, the flow amount on the arc must be zero

Constraint (3.5b): The capacity of storage facility $f \in F_{WH}$ restricts the amount of inventory that it can hold

Constraint (3.5c): The capacity of processing facility $f \in F_{PR}$ restricts the amount that it produces

Constraint (3.6a): The capacity at farm $f \in F_{F1}$ (i.e., supply limit) restricts the amount of biomass it can supply

Constraint (3.6b): Flow balance at preprocessing facility $f \in F_{P2}$

Constraint (3.6c): Flow balance at biomass storage facility $f \in F_{WH1}$, including dry mass loss

Constraint (3.6d): Flow balance at conversion facility $f \in F_{R2}$

Constraint (3.6e): Flow balance at biofuel storage facility $f \in F_{WH2}$

Constraint (3.6f): Flow balance at customer zone $f \in F_{CZ}$ (the inflow of biofuel must be less than or equal to demand)

Constraint (3.7a): Binary restrictions on decision variables x_{fr}.

Constraint (3.7b): Binary on decision variables y_a.

Constraint (3.7c): Non-negativity restrictions on decision variables q_{fr}.

Constraint (3.7d): Non-negativity restrictions on decision variables z_{ka}.

3.3 Case study

We now present a case study to demonstrate the types of analysis our model will support. We select nine counties in Central Texas as a test bed because this region has relatively high biomass availability (Milbrandt, 2005) and it is representative of regions that cannot provide a

sufficient amount of crop residues to meet its own biofuel needs, so they must be supplemented with energy crops to meet demand. Even though this region has relatively high biomass availability, our analysis found that most crop residues must be left in the fields after harvest to maintain soil fertility (see section 3.3.3). Thus, instead of crop residues, we employ switchgrass as a feedstock to meet demand under the assumption that it will be grown on some of the land currently used to grow food crops as well as some of the land set aside by the Conservation Reserve Program (CRP), which is a voluntary conservation program administrated by USDA to assist agricultural producers in enhancing environmentally sensitive lands.

We assume that a single biofuel (i.e., ETOH) is produced from several types of cellulosic feedstocks (i.e., switchgrass, mill residues, and urban wood wastes). We use a one-year planning horizon in which each period represents a quarter (i.e., three months).

We study the effect of four factors on SC design (switchgrass cost, ETOH price, switchgrass yield, and ETOH demand) under 18 scenarios that evaluate switchgrass cost vs. ETOH price and three other scenarios that assess supply vs. demand (see Table 4). We analyze several measures of the prescribed SC performance (e.g., profit, revenue, cost, material flow pattern, ratio of supply to demand, and land area used) for each scenario. We base ETOH price on the trend forecasted by the U.S. Energy Information Administration for the price of regular-grade gasoline, including taxes. Even though counties are in the same region, each may offer different soil, terrain, and weather conditions, affecting the price of biomass. Furthermore, contracts with growers may differ. While our model can accommodate county-dependant parameter values, our case study assumes the same values for all counties due to the lack of more specific data.

The following sub-sections describe the data we gathered to formulate our case study. Most data are available from papers, reports and data services provided by the U.S. government.

Table 4 Scenarios of sensitivity analysis for a region in Central Texas

Scenario Group	Scenario No	Switchgrass Cost	ETOH Price		Switchgrass Yield		ETOH Demand
		$/Mg	$/gal	$/L	Mg/acre	Mg/ha	
A. Cost vs. Price	1	50	2.50	0.66	10.00	24.71	E10 for local area
	2	50	2.60	0.69	10.00	24.71	E10 for local area
	3	50	2.70	0.71	10.00	24.71	E10 for local area
	4	50	2.80	0.74	10.00	24.71	E10 for local area
	5	50	2.90	0.77	10.00	24.71	E10 for local area
	6	50	3.00	0.79	10.00	24.71	E10 for local area
	7	60	2.50	0.66	10.00	24.71	E10 for local area
	8	60	2.60	0.69	10.00	24.71	E10 for local area
	9	60	2.70	0.71	10.00	24.71	E10 for local area
	10	60	2.80	0.74	10.00	24.71	E10 for local area
	11	60	2.90	0.77	10.00	24.71	E10 for local area
	12	60	3.00	0.79	10.00	24.71	E10 for local area
	13	70	2.50	0.66	10.00	24.71	E10 for local area
	14	70	2.60	0.69	10.00	24.71	E10 for local area
	15	70	2.70	0.71	10.00	24.71	E10 for local area
	16	70	2.80	0.74	10.00	24.71	E10 for local area
	17	70	2.90	0.77	10.00	24.71	E10 for local area
	18	70	3.00	0.79	10.00	24.71	E10 for local area
B. Supply vs. Demand	19	60	2.90	0.77	7.00	17.30	E10 for local area
	20	60	2.90	0.77	13.00	32.12	E10 for local area
	21	60	2.90	0.77	10.00	24.71	E20 for local area

3.3.1 Cost estimates and technical factors

Table 5 presents a list of cost estimates and technical factors used in the case study. We have consulted a number of publically available sources to gather data with the goal of making the case study as realistic as possible.

We employ a one-year planning horizon and amortize the cost of capital for opening a facility over a 20-year lifetime at a 10% discount rate. However, since the biofuel industry is in an early stage of development, not much recent data is available. Therefore, a case study

Table 5 Cost estimates and technical factors

Parameters			Value	Reference	Remark
Refinery	Fixed opening cost		$4,511,168	Nguyen and Prince (1996)	
	Variable opening cost		$0.0071/L	Nguyen and Prince (1996)	
	Variable operating cost		$0.032/L	Aden et al. (2002)	
Preprocessing facility	Fixed opening cost		$60,000	Sokhansanj et al., (2006a)	Using a Tub Grinder and a biomass dryer
	Variable opening cost		$1.00/Mg	Sokhansanj et al., (2006a)	
	Variable operating cost		$18.14/Mg	Sokhansanj et al., (2006a)	
Biomass storage facility	Indoor Anaerobic	Fixed cost	$15,153	Anderson and Noyes (2010)	
		Variable cost	$151.46/m^2	Anderson and Noyes (2010)	
	Outdoor uncovered	Fixed cost	$1,515	Assumption	10% of Outdoor uncovered
		Variable cost	$15.15/m^2	Assumption	
Biofuel storage facility	Fixed cost		$1,000	Assumption	
	Variable cost		$0.063/L	Assumption	Referred to eBay and Northern Tool
Transportation cost	Biomass		$6.81/Mg + $0.08/Mg-km	Glassner et al. (1998)	Based on bale system. Same for all biomass
	ETOH		$1.00/Mg + $0.08/Mg-km	Assumption	
Single destination	Management cost		$100/month	Assumption	
Biomass purchase cost	Switchgrass		$50, $60 and $70/Mg	Perrin et al., (2008)	
	Mill residues		$4/Mg	Fehrs (1999)	
	Urban wastes		$12/Mg	Fehrs (1999)	
Conversion factor	From biomass to ETOH		70% of the theoretical estimate	Assumption	Hamelinck et al. (2005)
Dry matter loss	Outdoor uncovered	Switchgrass	2%/month	Shinners and Binversie (2004)	
		Wood wastes	2%/month	Kofman (2006)	
	Indoor anaerobic	Switchgrass	0.3%/month	Shinners and Binversie (2004)	
		Wood wastes	0.5%/month	Kofman (2006)	
Moisture content	Switchgrass		Uniform(20%, 50%)	Kumar and Sokhansanj (2007)	
	Wood wastes		Uniform(10%, 20%)	Fehrs (1999)	

solution might over- or under-estimate the actual value a bit because data that describe the change of cost parameters over time is not yet available.

It was not possible to find sources for some parameters, so we estimate them based on values associated with similar processes. We assume that cost for an outdoor-uncovered storage facility is 10% of the cost of an indoor anaerobic storage facility, and that of transporting ETOH involves a fixed cost of $1.00/Mg plus a variable cost of $0.08/Mg km. We estimate the cost of biofuel storage based on the cost of oil storage tanks posted on some commercial online shopping malls [eBay.com], augmented with the fixed cost of capital for land ($1,000) and the variable operating cost ($0.01/liter). We assume that the cost for the single-destination limitation relates to administration and is about $100/month. A wide range of efficiencies of characterizing the technologies that convert cellulosic biomass to ETOH has been reported (e.g., 35 to 68% (Hamelinck et al., 2005)). We assume a 70% of conversion efficiency for each type of biomass, based on their theoretical estimates. The theoretical ETOH yield of biomass estimated based on Theoretical Ethanol Yield Calculator by DOE are 397.6 liter/Mg for switchgrass, 381.6 liter/Mg for secondary mill residues, and 439.1 liter/Mg for urban wood waste.

3.3.2 Demand

This sub-section describes how we estimate biofuel demand in the Central Texas region. We assume that the demand for ETOH is 10% of the demand for gasoline, because E10 (a mixture of 10% ETOH and 90% gasoline) can be distributed easily in the current infrastructure for petroleum-based fuel.

Table 6 gives the population of each county in the region, the average annual consumption of gasoline from 1998 to 2007 in Texas and an estimate of the demand for ETOH (10% of gasoline consumption based on the population of each county). The overall average demand for ETOH is 1,425 liters/year/person (National Priorities Project Database).

3.3.3 Biomass supply

We now describe our analysis procedure and estimate biomass availability in the selected region. We estimate the amount of available crop residue based on farmed land area and crop yield in each county as provided by USDA-National Agricultural Statistics Service. Wilhelm et al. (2004) reported that some residue associated with certain crops should be left in the field to maintain soil quality: more than 6.0 Mg/hectare/year for corn residue; and more than 3.0 Mg /hectare/year for wheat residue. Based on their estimates, most crop residues available in the Central Texas region should be left in the field. Thus, we have not considered crop residues as possible feedstocks.

To supply an amount of biomass sufficient to meet demand, we assume that switchgrass will be grown on some farm lands instead of the current food crops and on some CRP land areas as well. Table 6 gives an estimated amount of switchgrass that could be made available in each county. In addition to using switchgrass as a feedstock, we include other cellulosic biomass (i.e., mill residues and urban wood wastes) as feedstocks based on the data provided by Milbrandt (2005) so that we consider three types of biomass as a feedstock. We assume that switchgrass is harvested only in Summer and Fall and other biomass (i.e., mill residues and wood wastes) is supplied uniformly in all seasons.

3.3.4 Transportation cost

We estimate transportation distance based on the length of the straight line between the center points of each pair of counties. We invoke two assumptions to estimate transportation distances within each county: preprocessing, biorefinery, and distribution center facilities are at the same location so that the cost of transportation between each relevant pair of echelons is very small; farms and customer zones are uniformly distributed within a county so that the transportation distance between each pair of field storage and preprocessing facilities, and

Table 6 Estimated biofuel demand and switchgrass availability

	County		Hill	McLennan	Falls	Bell	Williamson	Travis	Hays	Comal	Bexar
Demand	**Population (2008)[(a)]**		35,637	230,213	16,900	285,084	394,193	998,543	149,476	109,635	1,622,899
	Gas Consumption[(b)] (KL/year)		50,769	327,962	24,076	406,131	561,568	1,422,526	212,944	156,186	2,311,985
	Demand for ETOH(KL/year)		5,077	32,796	2,408	40,613	56,157	142,253	21,294	15,619	231,198
Supply	**Farm Land Area(2009) (ha)**		72,843	60,662	56,535	53,580	57,749	9,105	1,012	6,111	11,250
	CRP land Area (2009) (ha)		1,912	69	140	349	379	0	0	0	0
	Total Area (ha)		74,755	60,732	56,675	53,929	58,128	9,105	1,012	6,111	11,250
	Switchgrass Production Amount (Mg/year)	**Yield: 7 Mg/ha/year**	523,285	425,121	396,725	377,503	406,893	63,738	7,082	42,775	78,752
		Yield: 10 Mg/ha/year	747,550	607,316	566,750	539,289	581,276	91,054	10,117	61,108	112,503
		Yield: 13 Mg/ha/year	971,815	789,511	736,775	701,076	755,659	118,371	13,152	79,440	146,253

(a) The U.S. Census Bureau, 2008

(b) 1,425 L/year/person (National Priorities Project Database)

between each pair comprising a distribution center and a customer zone is about 19 km, because the size of most counties is approximately 50 by 50 km and the average distance from any location in a county to the center point is about 19 km. The transportation mode considered in this case study is a truck. We assume that the bale system is used to transport biomass.

3.4 Results

This section analyzes the results of our computational experiments for 21 scenarios (See Table 1). The number of binary and continuous variables are 2,034 and 96,030, respectively. IBM ILOG CPLEX 12.1 solved each instance within one hour in a personal computer with Core(TM)2 Duo CPU 3.16GHz and 4G RAM, prescribing an optimal solution for each scenario.

3.4.1 Basic results

This section describes results prescribed by the optimization model for scenario 11, which we consider to be a basic (i.e., bench marking) scenario because it has the smallest ETOH price (i.e., $0.77/Liter) while meeting all customer-zone demands for biofuel using the median cost of switchgrass (i.e., $60/Mg).

3.4.1.1 Facility location, technology, and capacity

Table 7 describes results about the strategic level decisions. While preprocessing facilities are opened at three counties (i.e., 2, 5, and 9), refineries are only opened at counties 2 and 5. Note that we consider only single technology for preprocessing facility and refinery in this case study. Alternative technology is considered for the selection of storage types in preprocessing facility (i.e., outdoor-uncovered and indoor-anaerobic). However, only field storage facilities, the type of which is an outdoor-uncovered, are opened mainly due to high fixed cost of other storages in echelons of preprocessing facility and refinery. Relatively large field storages are prescribed at counties 1, 3, 5, and 6.

Table 7 Facility location, capacity, and technology type for scenario 11

Echelon	Location		Capacity	Technology Type
	No	County Name		
Field Storage	1	Hill	575,722 Mg	outdoor-uncovered
	2	McLennan	4,950 Mg	outdoor-uncovered
	3	Falls	403,626 Mg	outdoor-uncovered
	4	Bell	9,950 Mg	outdoor-uncovered
	5	Williamson	583,141 Mg	outdoor-uncovered
	6	Travis	114,813 Mg	outdoor-uncovered
	7	Hays	8,759 Mg	outdoor-uncovered
	8	Comal	63,341 Mg	outdoor-uncovered
	9	Bexar	16,001 Mg	outdoor-uncovered
Preprocessing	2	McLennan	1,299,048 Mg/year	grinding & drying
	5	Williamson	938,998 Mg/year	grinding & drying
	9	Bexar	64,192 Mg/year	grinding & drying
Biorefinery	2	Bell	205,920 KL/year	70% conversion efficiency
	5	Williamson	222,569 KL/year	70% conversion efficiency
DC layer	-	-	-	-

3.4.1.2 Material flow pattern

We analyze the material flow pattern in each time period under the scenario 11. Figure 6 depicts the material flow pattern in each season. Each straight arc represents the aggregated flow of all types of biomass in the upstream and ETOH in the downstream. Rectangular icons in the field storage echelon represent inventory carryovers from one period to the next. Using facilities opened (See Table 7), materials flow through three preprocessing facilities (in counties 2, 5, and 9), and two refineries (in counties 2 and 5).

Since we assumed that switchgrass is harvested only in Summer and Fall, significant amounts of biomass inventory are carried over in field storage to meet year-round demand. Even though some portion of biomass is degraded in field storage due to chemical dry matter loss, our model prescribes inventory carryover only in the field storage echelon, because the capital cost of storage in other echelons (i.e., preprocessing facility and refinery) is relatively high compared

to the potential cost of dry matter loss. The amounts of dry matter loss in storages are 198,256 Mg of switchgrass, 821 Mg of mill residues, and 7,803 Mg of wood wastes, respectively.

A quarter-by-quarter plan for material flow is essential to support strategic decisions and can give useful information to Energy companies in support for tactical-level decisions, (e.g., in planning manpower and equipment needs in each period). However, to support tactical-level plans most effectively, the duration of a time period should be defined as a month, if not an even shorter time. In fact, time periods need not be of the same duration. Shorter time periods could be used to model the dynamics of harvesting and longer ones could be used to plan inventories and material flows at other times of the year.

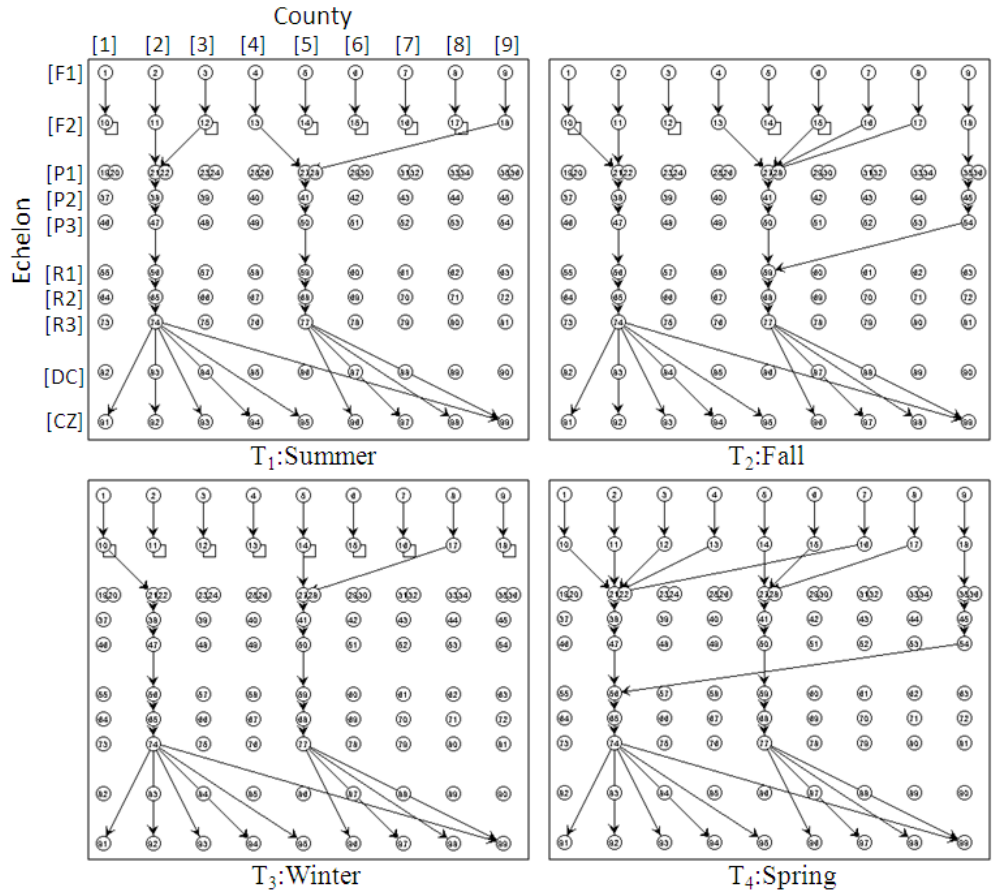

Figure 6 Material flow pattern for scenario 11 in each time period

3.4.2 Sensitivity Analysis

This section discusses results from several scenarios and analyzes factors that are significant to the biofuel SC.

3.4.2.1 Impact of the combination of feedstock cost and ETOH price

We compare performance measures (i.e., profit, revenue, and cost) that result in scenarios 1 - 18 to identify the impact of the combination of feedstock cost and ETOH price. Figure 7 shows that profit increases faster as feedstock cost reduces than as ETOH price increases. This reinforces the expectation that feedstock cost is a significant factor in determining profit. When the price of ETOH at the pump is less than \$0.66/L (i.e., scenarios 7 and 13), the biofuel SC is not economically viable; our model opens no facilities and prescribes no material flow.

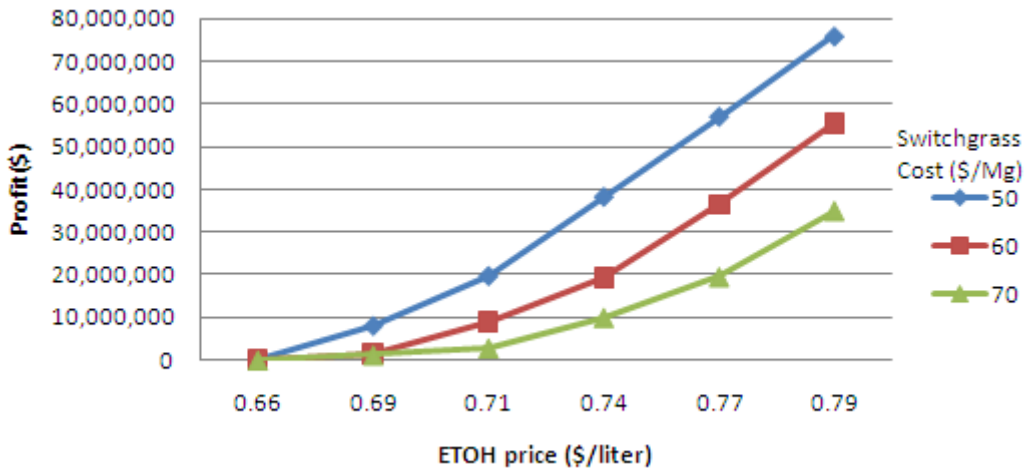

Figure 7 Relationship of profit to combinations of feedstock cost and ETOH price

Table 8 details results for scenarios 1 - 21 (ETOH production; profit; revenue; total cost; cost breakdown on detail processes; the percentage of land area used to grow switchgrass; and

the ratio of supply to demand). The refinery represents the most significant cost component in the SC. The cost to both purchase and collect feedstocks, is also a significant component. These results suggest that the components of the biofuel SC that offer the most leverage to improve the economic viability of the biofuel industry are the costs of refining and of feedstocks.

In terms of material flows, as ETOH price increases and feedstock cost reduces, the total amount of ETOH supplied increases so that the revenue increases. Our model prescribes the same amount of material flows for a set of scenarios (i.e., for those with a cost of switchgrass that is either $50/Mg ($60/Mg) with a price of ETOH that is above $0.74/L ($0.77/L)) because all demands are met in these instances and the amounts of processing materials are same. In addition, this also describes the reason why the total costs of scenarios 4, 5, and 6 are same and those of scenarios 11 and 12 are same.

Energy companies would be able to use a sensitivity analysis to evaluate the economic feasibility of generating biofuels in a selected region. Moreover, applying our model and method of analysis to a region or the entire country would provide useful information for government policy makers, for example, in estimating the subsidy levels required to induce investment in the bioenergy industry.

3.4.2.2 Impact of the combination of feedstock supply and ETOH demand

We now analyze the significance of combinations of feedstock supply and ETOH demand, studying scenarios based on scenario 11, which meets all customer-zone demands for biofuel. Scenarios 19, 20 and 21 are the same as scenario 11, except 19 tests a switchgrass yield of 17.30 Mg/ha (7.00 Mg/acre); 20 tests a switchgrass yield of 32.12 Mg/ha (13.00 Mg/acre); and 21 assumes that E20 can be used, doubling ETOH demand in comparison to scenario 11, which assumes that E10 is used.

Table 8 compares results for scenarios 19, 20 and 21 with those of scenario 11. As switchgrass yield decreases (i.e., scenario 19), the amount of ETOH produced, profit and land-area used decrease. Even though the amount of biomass is not sufficient to meet demand, some portion of land area is not used, because using it is not profitable. On the other hand, as switchgrass yield increases (i.e., scenario 20), the amount of ETOH produced and profit increase and the land area used decreases. As demand increases (i.e., scenario 21), the land area used increases slightly so that profit and the amount of ETOH produced also increase.

Energy companies can analyze land areas requirements to determine which farms they should contract to supply feedstocks in the most profitable way. For example, even though feedstock supply is not sufficient to meet demand under scenarios 19 and 21, some portion of land area in county 9 is not used because of prohibitive transportation cost.

3.5 Discussions

Our mathematical model, which deals with a multiple commodity flows, represents specific characteristics of several types of cellulosic biomass as well as the changes biomass undergoes in storage and processing in the various echelons of the SC. To our knowledge, this is the first model to deal with all echelons of the biofuel SC, including both upstream and downstream; the selection of technology and location for each opened facility; biomass moisture content; dry matter loss in storage; and single destination in the upstream of the biofuel SC.

Our case study demonstrates the use of our model as a decision support tool based on a set of data we have been able to gather from public sources to represent the biofuel industry in Central Texas. Our case study indicates that this region would be able to meet all local demand for ETOH only under certain scenarios that utilize E10. In particular, case studies provide informative results, identifying relationships that have not been investigated previously. Even though other factors (i.e., feedstock cost, feedstock yield, and ETOH demand) affect the

Table 8 Comparison of scenarios to analyze combinations of feedstock cost and ETOH price

Scenario No	ETOH Production (KL)	Profit ($)	Revenue ($)	Total Cost ($)	Cost breakdown (%)					Used /Given Area (%)	Supply /Demand (%)
					FS[c]	FT[d]	PR[e]	RF[f]	ET[g]		
1	129,116	57,829	115,710,527	115,652,698	16.45	7.96	8.93	66.25	0.39	34.6	24.5
2	253,460	7,764,851	236,229,115	228,464,264	18.56	8.07	8.96	63.93	0.48	63.2	48.1
3	427,976	19,542,153	414,222,801	394,680,649	20.60	7.77	8.79	61.70	1.13	96.6	81.2
4	527,157	38,188,889	529,117,456	490,928,558	21.41	6.68	8.76	61.81	1.34	98.6	100.0
5	527,157	57,085,950	548,014,508	490,928,558	21.41	6.68	8.76	61.81	1.34	98.6	100.0
6	527,157	75,983,002	566,911,560	490,928,558	21.41	6.68	8.76	61.81	1.34	98.6	100.0
7[b]	-	-	-	-	-	-	-	-	-	-	-
8	57,433	1,431,448	53,528,573	52,097,125	12.50	8.40	8.61	70.23	0.24	7.8	10.9
9	248,363	8,862,751	240,381,774	231,519,024	21.29	7.73	8.66	61.86	0.45	61.3	47.1
10	362,200	19,290,223	363,544,743	344,254,519	22.99	7.48	8.52	60.07	0.94	78.3	68.7
11[a]	527,157	36,558,390	548,014,508	511,456,118	24.56	6.41	8.41	59.33	1.29	98.6	100.0
12	527,157	55,455,442	566,911,560	511,456,118	24.56	6.41	8.41	59.33	1.29	98.6	100.0
13[b]	-	-	-	-	-	-	-	-	-	-	-
14	41,369	1,221,476	38,556,930	37,335,453	8.18	9.09	8.49	73.97	0.24	1.0	7.8
15	54,609	2,816,185	52,854,063	50,037,878	12.90	8.38	8.50	69.97	0.24	6.6	10.4
16	246,584	9,973,884	247,499,534	237,525,650	23.80	7.51	8.38	59.88	0.44	60.4	46.8
17	362,200	19,495,985	376,528,484	357,032,498	25.74	7.21	8.22	57.92	0.90	78.3	68.7
18	527,157	34,952,361	565,901,332	530,948,971	27.22	7.34	8.05	56.20	1.19	98.6	100.0
19	380,859	28,144,334	395,926,768	367,782,433	23.85	7.72	8.39	59.06	0.97	96.6	72.2
20	527,157	41,958,476	548,014,508	506,056,032	23.99	6.40	8.47	59.96	1.18	94.8	100.0
21	529,493	39,925,524	550,439,343	510,513,818	24.72	6.48	8.46	59.69	0.65	98.8	50.2

(a): a basic scenario to compare other scenarios about supply/demand changes; (b): non-profitable; (c): FS-Feedstock;
(d): FT-Feedstock Transportation; (e): PR-Preprocessing; (f): RF-Refinery; and (g):ET-ETOH Transportation

economic viability of the SC, ETOH price appears to be the most significant factor to economic viability; moreover, based on profits, the overall SC structure prescribed could be much different from one based on minimizing cost.

Biofuel manufacturers can use our mathematical model to plan the most profitable SC design and estimate the profit that a particular region might generate. In addition, government policy makers can employ our model to identify policies most likely to support a viable biofuel industry, for example, through a combination of providing subsidies and attracting private investment.

CHAPTER IV

A SOLUTION METHOD FOR BSCP MIP MODEL

This chapter describes a solution method to solve large instances of BSCP MIP model formulated in Chapter III. Section 1 describes a formulation of BSCP, an alternative to the embedded multi-commodity flow problem proposed by Chapter III. Section 2 explains our CG approach, which incorporates our new DP algorithm to generate flow-paths in the uncapacitated, embedded GFP subproblem. Section 3 describes the logic underlying POC. Section 4 evaluates the performance of our solution approach through computational tests.

4.1. Formulation

Our prior formulation of BSCP in Chapter III deals with material flows based on a multi-commodity viewpoint, defining each commodity in the upstream as the combination of biomass type and moisture content, which depends on location and time period. Since the number of flow variables is proportional to the number of commodities, reducing the latter would reduce the number of flow variables. In comparison, the alternative formulation we present in this paper reduces the number of commodities, downsizing the model in Chapter III and, therefore, enhancing solvability. This section describes a two-step procedure to define each commodity and the network that represents flows.

4.1.1 Commodity definition

We consider moisture content implicitly by invoking certain assumptions and defining transportation costs appropriately. The assumptions are that the moisture content of feedstock held in field storage facilities changes according to the weather conditions in each time period and that moisture content becomes negligible after drying during preprocessing. We adjust the cost of transporting each unit on each arc in each echelon before preprocessing to reflect the

moisture content of biomass in each time period. For example, 10 Mg biomass with a wet-basis moisture content of 10% is 9 Mg of dry biomass and 1 Mg of water. If feedstock is transported at a unit cost of $9.90 per Mg, the cost to transport this 10 Mg is $99.00. In contrast, if we define unit transportation cost to be $11.00 (=9.9*10/9) per Mg instead of $9.90, we can calculate transportation cost by considering only the dry biomass being transported; that is, 9 Mg*$11.00 per Mg = $99.00. Thus, by defining unit transportation costs based on the moisture content in each time period, we can deal with only the dry matter of biomass, so that a single commodity can be defined based only on biomass type.

4.1.2 Network definition

Even after defining commodities, several biomass types can be transported on the same route in the upstream. Our modeling alternative forms the underlying network by duplicating nodes and arcs in the upstream with respect to biomass types (e.g., switchgrass and wood wastes), so that only a single commodity flows on each (duplicate) arc. Figure 8 depicts an example of duplicating nodes and arcs with respect to biomass type *k*.

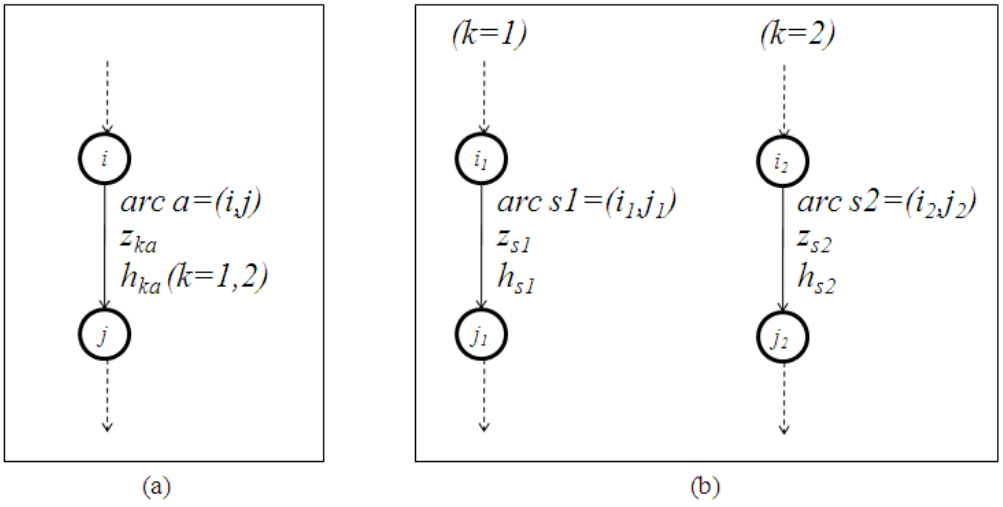

Figure 8 Network transformation by duplicating nodes and arcs: (a) original network, and (b) duplicate network

In Figure (8a), two biomass types flow on arc (i, j). We duplicate nodes i and j, creating i_1, i_2 and j_1, j_2, respectively, and all other arcs incident to nodes i and j as well. For example, arc (i, j) is duplicated to form (i_1, j_1) and (i_2, j_2), one for each type of biomass. The flow variable in the original network, z_{ka}, can be represented as z_{s1} and z_{s2} in the duplicate network, where *s1* and *s2* are arcs in the duplicate network (Figure (8b)) and a is an arc in the original network (Figure (8a)).

Flow is reduced on some arcs because dry matter loss occurs over time in storage and the conversion efficiency of a refinery is less than 100%. For example, a flow of 100 Mg of biofuel, produced at a conversion efficiency of 70%, reduces a flow of 142.86 (=100/0.7) Mg of biomass into the refinery. This requires the flow of 145.77 (=142.86/0.98) Mg of biomass into the storage facility, in which biomass is subject to 2% dry matter loss in a month. Such a flow problem can be modeled as a GFP.

This modeling alternative results in a single commodity, biomass, flowing in the upstream, but conversion introduces another commodity type, biofuel, in the downstream, we must deal with such a commodity type change. Flow balance constraints related to supply (i.e., farms) and demand (i.e., customers) nodes may not be satisfied at equality because of flow gains and losses associated with biomass storage and conversion. Moreover, whereas an exact formulation of GFP requires an equality condition for all constraints that represent flow balance, our prior formulation of BSCP involved some inequality constraints associated with supply and demand nodes to prescribe only profitable flow quantities. Therefore, to transform the embedded flow problem to the exact form of an embedded GFP, additional manipulation of the network is required, as we now describe. Figure 9 shows an example of the final form of a duplicate network, which is acyclic.

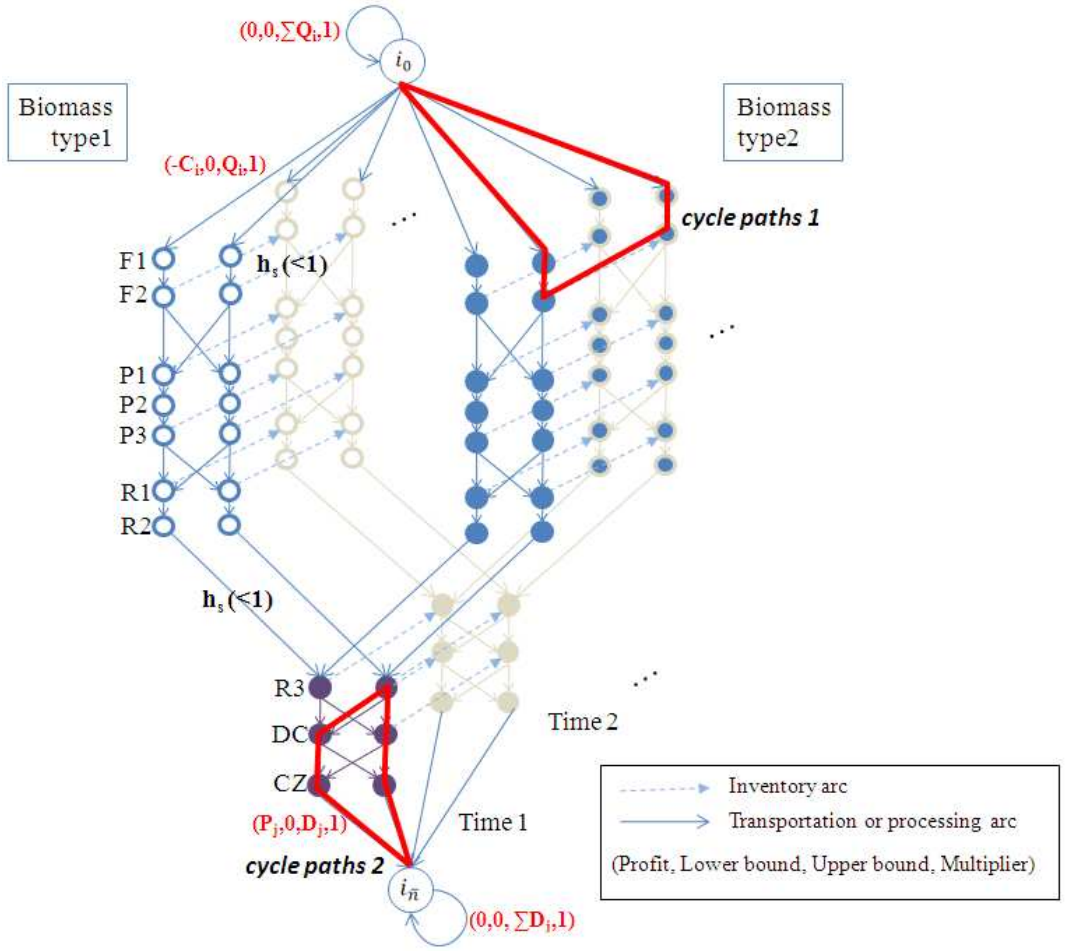

Figure 9 Example of a duplicate network

First, we convert each unit of flow to the same metric (e.g., a measure of energy) and revise related parameters (e.g., costs, prices, and multipliers) to correspond. Second, dummy start and end nodes (i.e., i_0 and $i_{\bar{n}}$, respectively, where $\bar{n} = n + 1$) are augmented, along with new directed arcs that connect node i_0 to each supply node and each demand node to $i_{\bar{n}}$. The cost associated with each new arc emanating from i_0 and pointing to a supply node is minus one times the cost of biomass at the supply point; its lower bound is zero and its upper bound is the capacity of the supply node to produce biomass. Similarly, the cost of each new arc from a

demand node to $i_{\bar{n}}$ is the biofuel selling price; its lower bound is zero; and its upper bound is the demand of the demand node.

Lastly, two special arcs, one from i_0 to i_0 and one from $i_{\bar{n}}$ to $i_{\bar{n}}$, are added to allow a feasible flow balance at i_0 and $i_{\bar{n}}$; i.e., the flow amount on arc (i_0, i_0) (arc ($i_{\bar{n}}$, $i_{\bar{n}}$)) must equal the amount of flow-out (flow-in) of node i_0 ($i_{\bar{n}}$) (Ahuja et al., 1993). Then, the modified network, in which supplies and demands of all nodes are zero, represents a generalized minimum-cost circulation problem (Wayne, 2002). A unit of flow on arc (*i*,*j*) will be h_{ij} at node *j*: if $h_{ij} < 1$, the arc is *lossy*; and if $h_{ij} > 1$, the arc is *gainy*.

4.1.3. Mathematical model

Table 9 defines the notation we use in our formulation, model 1 ((4.1)-(4.7)).

Table 9 Notation

Indices	
a : arc	$a \in A$
b : biomass type	$b \in B$
e : biofuel type	$e \in E$
f : facility	$f \in F$
i : duplicate node	$i \in N^d$
k : commodity	$k \in K$
l : layer (echelon)	$l \in L$
r : technology type	$r \in R_f$
s : duplicate arc	$s \in A^d$
t : time	$t \in T$
Sets	
A : Directed arcs	$:= A^+_{frt} \cup A^I_{frt}$
A^d : Duplicate directed arcs	$:= A^{d+}_{kfrt} \cup A^{dI}_{kfrt}$
A^d_a: Arcs that are duplicate from original arc a	
A^+_{frt} (A^-_{frt}): Directed arcs in period *t* that start or end at node *frt*	
A^I_{frt} : Arc for which flow represents inventory held at facility *f* of type *r* from period *t* to period *t+1*	
A^{d+}_{kfrt} (A^{d-}_{kfrt}): Duplicate arcs in period *t* that start (end) at duplicate node *kfrt*	

Table 9 Continued

A_{kfrt}^{dI} : Duplicate arc that represents inventory held at facility *f* of type *r* from period *t* to period *t+1*

A_T : Directed arcs associated with transportation $:= A_{frt}^{+}, f \in F_{F2} \cup F_{P3} \cup F_{R3} \cup F_{DC}, r \in R_f, t \in T$

B : Feedstock(biomass) types supplied by facility (e.g., farm) $f \in F_{F1}$

F_l : Candidate locations for facilities in echelon *l*, feedstock supply site (F_{F1}, F_{F2}), or customer zone (F_{CZ}), $l \in L$

F_{WH0} : Warehouses where biomass is held before preprocessing $:= F_{F2} \cup F_{P1}$

F_{WH1} : Warehouses where biomass is held $:= F_{WH0} \cup F_{P3} \cup F_{R1}$

F_{WH2} : Warehouses (i.e., storage tanks) where biofuel is held $:= F_{R3} \cup F_{DC}$

F_{WH}: Warehouses $:= F_{WH1} \cup F_{WH2}$

F_{PR} : Process facilities (preprocessing, refinery) $:= F_{P2} \cup F_{R2}$

F_{OP} : Facilities $:= F \backslash (F_{F1} \cup F_{CZ})$

F_{UP} : Upstream facilities $:= F_{F1} \cup F_{F2} \cup F_{P1} \cup F_{P2} \cup F_{P3} \cup F_{R1}$

F_{DOWN} : Downstream facilities $:= F_{R2} \cup F_{R3} \cup F_{DC} \cup F_{CZ}$

F : All facilities $:= F_{UP} \cup F_{DOWN}$

K_1: Feedstock types (i.e., commodities) $:= \{(f, t, b)\}, f \in F, t \in T, b \in B$

K_2: Biofuel commodity $:= \{e\}, e \in E$

K: Commodities $:= K_1 \cup K_2$

L : Echelons, $\{F1, F2, P1, P2, P3, R1, R2, R3, DC, CZ\}$ (*F1*: farm, *F2*: field storage, *P1-3*: preprocessing facilities, *R1-3*: conversion facilities, *DC*: distribution center, and *CZ*: customer zone)

N^d : Duplicate nodes

R_f : Types of technologies at facility *f*, $f \in F_l$

T : Time Periods

Parameters

C_a^T : Fixed cost of selecting arc *a*

C_{fr}^O : Fixed cost of opening facility *f* of technology type *r*

C_s : Revenue (>0) or cost (<0) for a unit flow on arc *s*

D_s : Upper bound on flow on arc *s*, associated with demand of the starting node *kfrt* of arc *s*

h_s : Multiplier associated with arc s

Q_a^T: Upper bound on flow on arc *a*

Q_f^F : Capacity limit of facility *f* (biomass storage, preprocessing, refinery, or biofuel storage)

Q_s : Upper bound on flow on arc *s*, which is associated with supply capacity of the end node *kfrt* of arc *s*

V_{fr} : Variable cost per unit of capacity of opening facility *f* of technology type *r*

Table 9 Continued

V_s^T : Variable cost for a unit of flow on arc *s* (variable transportation cost on transportation arc, variable holding cost on inventory arc)
Decision Variables
q_{fr} : Capacity of facility *f* of technology type *r*, $f \in F_{OP}, r \in R_f$
x_{fr} : 1 if facility *f* of type *r* is opened, 0 otherwise, $f \in F_{OP}, r \in R_f$
y_a: 1 if arc *a* is used, 0 otherwise, $a \in A$
z_s : Flow amount on duplicate arc s, $s \in A^d$

We now present model 1:

Model 1: $Z^* = \text{Max} - \sum_{f\in F_{OP}} \sum_{r\in R_f} C_{fr}^O x_{fr} - \sum_{f\in F_{OP}} \sum_{r\in R_f} V_{fr} q_{fr} - \sum_{a\in A} C_a^T y_a + \sum_{s\in A^d} C_s z_s$ (4.1)

s.t.

$$\sum_{r\in R_f} x_{fr} \le 1 \qquad f \in F_{OP} \qquad (4.2)$$

$$-Q_f^F x_{fr} + q_{fr} \le 0 \qquad f \in F_{OP}, r \in R_f \qquad (4.3)$$

$$\sum_{a\in A_{frt}^+} y_a \le 1 \qquad f \in F_{P1} \cup F_{R1}, r \in R_f, t \in T \qquad (4.4)$$

$$-Q_a^T y_a + \sum_{s\in A_a^d} z_s \le 0 \qquad a \in A_T \qquad (4.5a)$$

$$-q_{fr} + \sum_k \sum_{s\in A_{kfrt}^{dI}} z_s \le 0 \qquad f \in F_{WH}, r \in R, t \in T \qquad (4.5b)$$

$$-q_{fr} + \sum_k \sum_{s\in A_{kfrt}^{d+}} z_s \le 0 \qquad f \in F_{PR}, r \in R, t \in T \qquad (4.5c)$$

$$z_s \le Q_s \qquad s \in A_{kfrt}^{d-}, k \in K_1, f \in F_{F1}, r \in R_f, t \in T \qquad (4.6a)$$

$$z_s \le D_s \qquad s \in A_{kfrt}^{d+}, k \in K_2, f \in F_{CZ}, r \in R_f, t \in T \qquad (4.6b)$$

$$\sum_{s\in(i,j)} z_s - \sum_{s\in(j,i)} h_s z_s = 0 \qquad i \in N^d \qquad (4.6c)$$

$$x_{fr} \in \{0,1\} \qquad f \in F_{OP}, r \in R_f \qquad (4.7a)$$

$$y_a \in \{0,1\} \qquad a \in A_D \qquad (4.7b)$$

$$q_{fr} \ge 0 \qquad f \in F_{OP}, r \in R_f \qquad (4.7c)$$

$$z_s \ge 0 \qquad s \in A^d. \qquad (4.7d)$$

Objective (4.1) is to maximize the present worth of total system profit: the first term gives the fixed cost of opening all facilities; the second, total variable cost associated with the capacity of facilities opened; the third, total fixed cost associated with selection of single destination; and the

last, total net profit from supplying biofuel. Constraint (4.2) is that each facility $f \in F_{OP}$ can employ, at most, one technology type $r \in R_f$. Constraint (4.3) restricts the capacity of facility $f \in F_{OP}$ by its capacity limit Q_f^F, if it is opened; otherwise, it allows no flow from facility *f*. Constraint (4.4) requires that each field storage (preprocessing) facility $f \in F_{F2}(F_{P3})$ use a transport link to a single destination (i.e., preprocessing (refinery)) facility $f \in F_{P1}(F_{R1})$ to facilitate flow management in the upstream. Constraints (4.5) impose flow capacity Q_a^T for arc $a \in A$ (4.5.a), q_{fr} for storage facilities (4.5.b), and q_{fr} for processing facilities (4.5.c). Constraints (4.6) correspond to an embedded GFP: (4.6.a) imposes the capacity Q_s (i.e., supply limit) of arc s associated with farm $f \in F_{F1}$ and $r \in R_f$ in period $t \in T$ to restricts biomass flow; constraint (4.6.b) restricts the inflow of biofuel according to the demand D_s of arc *s* associated with customer zone $f \in F_{CZ}$ and type $r \in R_f$ in period $t \in T$; and constraint (4.6.c) balances flow at each (duplicate) node *i* in the (duplicate) network, where h_s marks gains, losses or unchanged flow across arc *s*, which is incident from *i*. Constraints (4.7.a) and (4.7.b) invoke binary restrictions on decision variables x_{fr} and y_a, respectively. Constraints (4.7.c) and (4.7.d) restrict decision variables q_{fr} and z_s, respectively, to be non-negative.

4.2. CG for an embedded GFP

Figure 10 depicts the structure of the constraint matrix of model 1. Constraints (6), which represent an embedded GFP on the network, involve a large number of continuous flow variables. Since GFP can be solved effectively by specialized algorithms (e.g., Vaidya, 1989; Murray, 1992; Kamath and Palmon, 1995; Wayne, 2002; and Goldfarb and Lin, 2002), an embedded GFP can be used effectively as a sub-problem in a CG approach in this and many other applications. This paper devises a CG approach, treating the embedded GFP as a subproblem, to solve the linear relaxation of model 1.

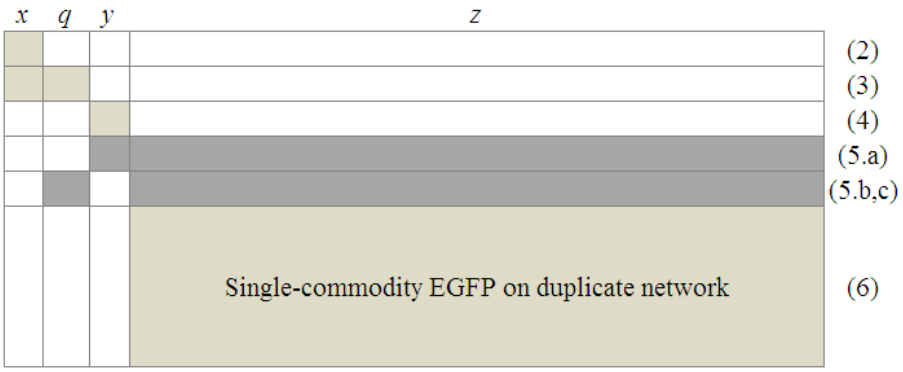

Figure 10 Structure of the constraint matrix

4.2.1 Tree- vs. path-flows in CG

If the flow upper bound of each arc is not considered, a feasible solution to the embedded GFP can be viewed as a path-flow; or, if considered, a tree flow, which can be represented as an aggregation of a set of path-flows. Jones et al. (1993) analyzed the impact of the number of sub-problem extreme points on the performance of CG in a multi-commodity flow problem. They reported that using path-flow solutions in multi-commodity flow subproblems is computationally superior to using tree flows, because a network admits fewer paths than trees, so that using path-flow solutions can result in substantially fewer master-problem iterations.

Following the earlier findings of Jones et al. (1993), we generate columns based on path-flows, imposing the upper bound constraint on each arc flow in the master problem rather than in the subproblem. Note that, since such an uncapacitated, embedded GFP is a linear program, the master problem of our decomposition scheme provides the same bound at each B&B node as does the linear relaxation of the original model 1.

4.2.2 Path-based formulation

This subsection presents our path-based formulation, model 2. Table 10 gives additional notation used in model 2.

Table 10 Notation for model 2

Indices
p : path, $p \in P$
Sets
P : Paths from α to β that satisfy flow balances (6.c)
$\bar{A}^1_{ps}$: Duplicate arcs from α to the one immediately preceding arc s on path p
$\bar{A}^2_{pi}$: Duplicate arcs from α to node i on path p
$\bar{A}_p$: Duplicate arcs on path p
$\bar{N}_p$: Duplicate nodes on path p
Parameters
C_p : Variable cost of a unit flow on path p $\coloneqq \sum_{s\in\bar{A}_p} H^p_s C_s, p \in P$
$H^p_s \coloneqq \prod_{j\in \bar{A}^1_{ps}} h_j,\ p \in P$ if $s \in \bar{A}_p$, 0 otherwise
Decision Variables
λ_p : Flow amount on path p, $p \in P$

To implement CG, generating columns from an uncapacitated, embedded GFP, we now transform the arc-based form of model 1 to the path-based form of model 2, the linear relaxation of which is the master problem in our CG decomposition:

Model 2: $Z^* = \text{Max} - \sum_{f\in F_{OP}} \sum_{r\in R_f} C^O_{fr} x_{fr} - \sum_{f\in F_{OP}} \sum_{r\in R_f} V_{fr} q_{fr} - \sum_{a\in A} C^T_a y_a + \sum_{p\in P} C_p \lambda_p$ (4.8)

s.t. (4.2), (4.3), (4.4), and (4.7a-c)

$$-Q^T_a y_a + \sum_{p\in P}\left(\sum_{s\in A^d_a} H^p_s\right)\lambda_p \le 0 \qquad a \in A_T \qquad (4.9a)$$

$$-q_{fr} + \sum_{p\in P}\left(\sum_{k\in K}\sum_{s\in A^{dI}_{kfrt}} H^p_s\right)\lambda_p \le 0 \qquad f \in F_{WH}, r \in R, t \in T \qquad (4.9b)$$

$$-q_{fr} + \sum_{p\in P}\left(\sum_{k\in K}\sum_{s\in A^{d+}_{kfrt}} H^p_s\right)\lambda_p \le 0 \qquad f \in F_{PR}, r \in R, t \in T \qquad (4.9c)$$

$$\sum_{p\in P}\left(H^p_s\right)\lambda_p \le Q_s \qquad s \in A^{d+}_\alpha \qquad (4.10a)$$

$$\sum_{p\in P}\left(H^p_s\right)\lambda_p \le D_s \qquad s \in A^{d-}_\beta \qquad (4.10b)$$

$$\lambda_p \ge 0. \qquad p \in P \qquad (4.11)$$

Objective (4.8) is the same as objective (4.1) except the fourth term which is re-expressed based on path p instead of arc s. Similarly, constraints (4.9a-c) and (4.10a, b) correspond with

constraints (4.5a-c) and (4.6a, b), respectively, replacing flow variable z_s and relevant coefficient by using variable z_p and coefficient based on path p: flow z_s on arc s can be represented by $z_s = \sum_{p \in P} H_s^p \lambda_p$. Constraints (4.11) restrict decision variables λ_p to be non-negative. Note that all path-flows on path $p \in P$ in model 2 satisfy constraints (4.6c) in model 1. The coefficient of λ_p (i.e., H_s^p) implies that each unit flow comprising λ_p induces flow H_s^p on arc s in path p. The linear relaxation of model 2 is the so-called master problem of a type II CG (Wilhelm, 2001) that uses an uncapacitated, embedded GFP price-out subproblem, to identify improving columns.

Model 2 can be transformed to form a type III CG, using Dantzig-Wolfe Decomposition, by adjusting the coefficient of λ_p on each arc $s \in \bar{A}_p$ from H_s^p to $H_s^p \pi^p$, where π^p is an extreme flow amount on path p, and by incorporating convexity constraint, $\sum_{p \in P} \lambda_p = 1$. However, our preliminary computational tests have shown that using type II CG gives better results for BSCP, so this paper reports its use. Our strategy is to solve the subproblem to determine an optimal unit path-flow from i_0 to $i_{\bar{n}}$, generating a column that enters the restricted master problem (RMP), which induces a set of path-flows to determine optimal, profitable flow quantities for the linear relaxation of model 2.

4.2.3 Subproblem to generate paths

Table 11 defines additional notation that we use to formulate the subproblem. Path-based model 2 involves a flow variable (λ_p) for each of many paths. The simplex optimality criterion indicates that entering path p as a column in the master problem basis will improve the current solution if $(wa_p - C_p) < 0$ and that, if $(wa_p - C_p) \geq 0$ for all paths $p \in P$, the current master problem solution is optimal, where w is a dual vector of model 2; a_p , $(0, \dots, (\sum_{s \in A_a^d} H_s^p), \dots, (\sum_{k \in K} \sum_{s \in A_{kfrt}^{dI}} H_s^p), \dots, \left(\sum_{k \in K} \sum_{s \in A_{kfrt}^{d+}} H_s^p\right), \dots, (H_s^p))^T$, is a column vector associated with variable λ_p; and C_p is the objective function coefficient associated with

variable λ_p, where $C_p = \sum_{s \in \bar{A}_p} H_s^p C_s$.

Table 11 Notation for sub-problem

Indices
u : row, $\quad u \in U$
Sets
$\hat{A}_p^u$: Arcs in path p with non-zero entries in row u of the constraint matrix of model 1
$\hat{R}_s$: Rows with non-zero entries in the column associated with arc s of the constraint matrix of model 1
U : Rows of the constraint matrix of models 1 and 2
Parameters
a_p : Column vector of coefficient associated with variable λ_p in the constraint matrix of model 2, $\quad p \in P$
w_u : Dual variable associated with row u in model 2, $\quad u \in U$
Decision Variables
X_p : 1 if path p is used, 0 otherwise, $\qquad p \in P$

Each non-zero element of a_p can be interpreted as sum of certain H_s^p values corresponding to arcs in path p with non-zero entries in row u of the constraint matrix of model 1. Thus, by using $\hat{A}_p^u$, which denotes a set of arcs in path p with non-zero entries in row u of the constraint matrix of model 1, each non-zero element of a_p can be expressed as a generalized form, $(\sum_{s \in \hat{A}_p^u} H_s^p)$. So, a_p can be re-expressed:

$$a_p = \left(0, \dots, \sum_{s \in \hat{A}_p^u} H_s^p, \dots, \sum_{s \in \hat{A}_p^{|U|}} H_s^p\right)^T \tag{4.12}$$

Let $X \coloneqq \{X_p \in \{0,1\}: \sum_{p \in P} X_p = 1, p \in P\}$. For a given vector of dual variable values, w^k, each associated with a row in the RMP, at iteration k in our CG procedure,

SUB: $Z_{sub}(w^k) = min\{(w^k a_p - C_p)X_p : X_p \in X\}$

$= min\left\{\left[(w_1^k, \dots, w_{|U|}^k)\left(0, \dots, \sum_{s \in \hat{A}_p^u} H_s^p, \dots, \sum_{s \in \hat{A}_p^{|U|}} H_s^p\right)^T - \sum_{s \in \bar{A}_p} H_s^p C_s\right] X_p : X_p \in X\right\}$

$= min\left\{\left[\sum_{s \in \bar{A}_p} H_s^p \left(\sum_{u \in \hat{R}_s} w_u^k - C_s\right)\right] X_p : X_p \in X\right\}$

$$= \min \{\left(\sum_{s\in \bar{A}_p} C_s^p\right) X_p : X_p \in X\}, \quad (4.13)$$

where $C_s' \coloneqq \sum_{u\in \hat{R}_s} w_u^k - C_s$; and $C_s^p \coloneqq H_s^p C_s'$. The final form of the subproblem is similar to the shortest path problem with cost C_s^p on arc *s*. However, since arc cost C_s^p depends on the path *p* in which arc s is included, and each arc may have a different cost in association with each possible path, solving subproblem (4.13) is not trivial. Therefore, we propose a new DP algorithm in the following section.

4.2.4 BRA to solve the sub-problem

This section presents a definition of the shortest distance from each node to end node $i_{\bar{n}}$, based on arc costs C_s^p, and describes BRA to solve the sub-problem.

4.2.4.1 Problem structure

Given that the series of *n(p)* nodes on path *p* is $i_0 - i_1 - i_2 \ldots - i_{n(p)-1} - i_{\bar{n}}$, model *SUB* can be re-expressed:

$$Z_{sub}(w^k) = min\ \{(C_{i_0 i_1}^p + C_{i_1 i_2}^p + \cdots + C_{i_{n(p)-1} i_{\bar{n}}}^p) X_p : X_p \in X\}. \quad (4.14)$$

By expanding cost parameters, C_s^p, using the definition $C_s^p = C_s' \prod_{j\in \bar{A}_{ps}^1} h_j$, we obtain

$$Z_{sub}(w^k) = min\{(C_{i_0 i_1}' + h_{i_0 i_1} C_{i_1 i_2}' + \cdots + (\prod_{j\in \bar{A}_{p,i_{n(p)-1}}^2} h_j) C_{i_{n(p)-1} i_{\bar{n}}}') X_p : X_p \in X\}. \quad (4.15)$$

Then, we can rearrange coefficients of this objective function based on arc multipliers:

$$Z_{sub}(w^k) = \min \{[C_{i_0 i_1}' + h_{i_0 i_1}\{C_{i_1 i_2}' + \cdots + h_{i_{n(p)-2} i_{n(p)-1}}\{C_{i_{n(p)-1} i_{\bar{n}}}'\} \ldots\}]\ X_p : X_p \in X\}. \quad (4.16)$$

To establish the dynamic programming recursion, let $f[i_{\bar{n}}] \coloneqq 0$;

$f[i_{n(p)-1}] \coloneqq C_{i_{n(p)-1} i_{\bar{n}}}'$;

$f[i_{n(p)-2}] \coloneqq C_{i_{n(p)-2} i_{n(p)-1}}' + h_{i_{n(p)-2} i_{n(p)-1}} C_{i_{n(p)-1} i_{\bar{n}}}' = C_{i_{n(p)-2} i_{n(p)-1}}' +$

$h_{i_{n(p)-2} i_{n(p)-1}} f[i_{n(p)-1}]$;

in general, $f[i_j] \coloneqq C_{i_j i_{j+1}}' + h_{i_j i_{j+1}} f[i_{j+1}]$, so that $f[i_0] \coloneqq C_{i_0 i_1}' + h_{i_0 i_1} f[i_1]$.

We can now simplify Equation (4.16):

$$Z_{sub}(w^k) = min\{[C'_{i_0 i_1} + h_{i_0 i_1}\{C'_{i_1 i_2} + \cdots + h_{i_{n(p)-2} i_{n(p)-1}}\{f[i_{n(p)-1}]\} \ldots\}]\, X_p : X_p \in X\}$$

$$= min\{f[i_0] X_p : X_p \in X\}. \quad (4.17)$$

By using Equation (4.17), we can define the shortest distance from each node *j* to the end node $i_{\bar{n}}$

Definition 4.1. The shortest distance from node i_j to end node $i_{\bar{n}}$, $f[i_j]$, is defined recursively by

$$f[i_j] := \begin{cases} 0 & if\ i_j = i_{\bar{n}} \\ \min\limits_{(i_j, i_{j+1}) \in A_{i_j}^{d+}} \{C'_{i_j i_{j+1}} + h_{i_j i_{j+1}} f[i_{j+1}]\} & if\ i_j \in N \backslash \{i_{\bar{n}}\}, \end{cases} \quad (4.18)$$

where $C'_{i_j i_{j+1}} = \sum_{u \in \hat{R}_{i_j i_{j+1}}} w_u - c_{i_j i_{j+1}}$.

A related property is stated by Property 2.

Property 4.2. If $P[i_j] = i_j - i_{j+1} - \cdots - i_{\bar{n}}$ is the shortest path from node i_j to $i_{\bar{n}}$, sub-path $P[i_{j+1}] = i_{j+1} - \cdots - i_{\bar{n}}$ is the shortest path from node i_{j+1} to $i_{\bar{n}}$.

Proof. This property follows from the recursion and the dynamic programming principle of optimality. ■

4.2.4.2 Computing algorithm

Based on Definition 4.1 and Property 4.2, we now describe BRA to find the shortest path from each node to end node $i_{\bar{n}}$. First, we note prior work on the conventional shortest-path problems. Ahuja et al. (1993) described two groups of solution algorithms: label-setting algorithms, which designate one label as permanent at each iteration, and label-correcting algorithms, which mark all labels as temporary until the final iteration. A few fundamental label-setting algorithms are the reaching algorithm on acyclic networks with the worst-case complexity of *O(m)*, and Dijkstra's algorithm on cyclic networks with nonnegative arc lengths, with $O(n^2)$. Several label-correcting algorithms (e.g., FIFO label-correcting algorithm with

worst-case complexity *O(nm)* (Bellman, 1958)) have been developed for solving the shortest path problem on networks with arbitrary costs and arbitrary topology.

Zhu and Wilhelm (2012) reviewed the literature (e.g., Desrochers and Soumis, 1988; and Dumitrescu and Boland, 2003) related to the resource-constrained shortest path problem and proposed a three-stage approach for solving it as a sub-problem in CG. However, the conventional shortest path problem is different from ours, in which the distance from each node to end node $i_{\bar{n}}$ is defined based on arc costs C_s^p.

Let $G = (N^d, A^d)$ be the duplicate network of the embedded GFP, where N^d is the set of nodes and A^d is the set of directed arcs. Note that G is acyclic. For a given $G = (N^d, A^d)$, order nodes topologically (Step 1), initialize $f[i_{\bar{n}}] = 0$ (Step 2) and $f[i] = M$ for each $i \in N^d$ (Step 3), where M is a big number, and update distance label of each predecessor of node $i_{\bar{n}}$ (Step 4). Then, process node j in decreasing topological order by updating distance label of each predecessor i of node j (Step 5). For each node j, scan incoming arcs. For each arc $(i, j) \in A_j^{d-}$, if $f[i] > C'_{ij} + h_{ij} f[j]$, set $f[i] = C'_{ij} + h_{ij} f[j]$. Figure 11 gives a formal description of our BRA.

Next, we prove the correctness of BRA by proving that whenever it processes node j in step 5, the optimal distance label of node j has already been determined so that BRA is a label-setting algorithm.

Proof. Suppose that BRA has processed nodes $i_{\bar{n}}, i_n, \dots, i_k$ and their distance labels are optimal. Next, BRA processes node i_{k-1}. Let the shortest path from node i_{k-1} to the node $i_{\bar{n}}$ be $i_{k-1} - i_h - \cdots - i_{\bar{n}}$ where $i_h > i_{k-1}$. By property 4.2, the path $i_h - \cdots - i_{\bar{n}}$ must be the shortest path from node i_h to $i_{\bar{n}}$. Since BRA processes nodes in decreasing topological order and $(i_{k-1}, i_h) \in A_{i_h}^{d-}$, node i_h is included in $\{i_k, \dots, i_{\bar{n}}\}$ and the distance label of node i_h, $f[i_h]$, is equal to the shortest distance of the path from node i_h to $i_{\bar{n}}$ by hypothesis. While processing node i_h, arc

(i_{k-1}, i_h) must be scanned and the distance label of node i_{k-1}, $f[i_{k-1}]$, set equal to $C'_{i_{k-1},i_h} + h_{i_{k-1}i_h} f[i_h]$; that is, the shortest distance from node i_{k-1} to $i_{\bar{n}}$ (i.e., path $i_{k-1} - i_h - \cdots - i_{\bar{n}}$). Therefore, when BRA processes node i_{k-1}, its distance label is already optimal. Even if alternative optima exist, the optimal distance label of node i_{k-1}, $f[i_{k-1}]$, is not affected because the shortest distance associated with each alternative shortest path is the same as $f[i_{k-1}]$. ■

Line No.	Step	Operation
1	1.	Determine the topological ordering of nodes in G
2	2.	$f[i_{\bar{n}}] \leftarrow 0$ and $succ[i_{\bar{n}}] \leftarrow \emptyset$
3	3.	$f[i] \leftarrow M$ and $succ[i] \leftarrow \emptyset$ for each $i \in N^d \backslash \{i_{\bar{n}}\}$
4	4.	For each arc incident to node $i_{\bar{n}}$: $(i, i_{\bar{n}}) \in A^d$
5		$f[i] \leftarrow C'_{i,i_{\bar{n}}}$ and $succ[i] \leftarrow i_{\bar{n}}$
6		End for
7	5.	For each node $j \in N^d \backslash \{i_{\bar{n}}\}$ in the decreasing topological ordering
8		For each incoming arc $(i, j) \in A_j^{d-}$
9		$temp \leftarrow C'_{ij} + h_{ij} f[j]$
10		If $f[i] > temp$
11		Then $f[i] \leftarrow temp$ and $succ[i] \leftarrow j$
12		End for
13		$j \leftarrow next\ node$
14		End while

Figure 11 BRA

Now, we analyze the worst case complexity of BRA.

Proposition 4.3. The worst case complexity of BRA is *O(m)*, where *m* is the number of arcs.

Proof. Step1, ordering nodes topologically can be done in *O(m)* (Ahuja et al., 1993). Step 2 runs in *O(1)*. Steps 3 and 4 run in *O(n)* and *O(m)*, respectively, where *n* is the number of nodes. Step 5 examines each arc just once (lines 7 and 8) and each line within step 5 runs in *O(1)* (lines 9-11 and 13) so that total runtime of step 5 is *O(m)*. Since $m \geq n$ according to the network structure of the embedded GFP, the worst-case complexity of BRA is *O(m)*. ■

Based on the shortest path solution found by BRA, we can construct a column representing flow on path p, which starts as a unit flow as it emanates from node i_0, flows as $H_s^p = \prod_{j \in \bar{A}_{ps}^1} h_j$ units on each arc s in path p, and ends as $\prod_{j \in \bar{A}_{pi_{\bar{n}}}^2} h_j$ units at node $i_{\bar{n}}$. The coefficient associated with variable λ_p in each row of model 2 represents this flow amount (i.e., H_s^p) on each arc s in path p.

4.2.5 Acceleration techniques

We employ two techniques to accelerate CG convergence. The first is to incorporate extra dual cuts. Liang and Wilhelm (2007) generalized extra dual cuts, describing that inserting a polynomial number of extra dual cuts into RMP upon initialization restricts the dual space, potentially accelerating CG convergence. Alvelos and Valerio de Carvalho (2007) incorporated cycle paths as extra dual cuts to generate several additional, feasible flow paths by forming a linear combination of the cycle and flow paths generated by the subproblem.

Similarly, this study uses two portions of the SC network to generate cycle paths as extra dual cuts: the first group of cycle paths is generated based on the upstream along cycles, $i_0 \rightarrow$ farm in period t $(i_1 \in F_{F1}) \rightarrow$ field storage facility of i_1 in period t $(i_2 \in F_{F2}) \rightarrow$ ... $\rightarrow$ field storage of i_1 in period t+j $\left(i_{j+2} \in F_{F2}\right) \rightarrow$ farm in the same location as i_1 in period t+j $\left(i_{j+3} \in F_{F1}\right) \rightarrow i_0$, where $t \in T$ and $1 \leq j \leq |T| - t$; and the second, based on the downstream along cycles, $i_{\bar{n}} \rightarrow$ customer zone $(i_1 \in F_{CZ}) \rightarrow$ distribution center $(i_2 \in F_{DC}) \rightarrow$ biofuel storage facility in a biorefinery $(i_3 \in F_{R3}) \rightarrow$ distribution center in a location different to i_2 $(i_4 \in F_{DC}) \rightarrow$ customer zone in a location different to i_1 $(i_5 \in F_{CZ}) \rightarrow i_{\bar{n}}$. Figure 2 depicts examples of such cycle paths.

The second technique is based on the conjecture that incorporating multiple columns found to be improving at each of CG iteration may lead to faster CG convergence. After solving the subproblem once, we incorporate several paths, each with positive reduced cost, rather than

incorporating only the best column. However, dual variables from RMP might not provide a good primal (master problem) solution (i.e., far from an optimal solution) in initial stages of the CG search, so that, if several paths were incorporated, some may never be entered into the RMP basis and would only increase runtime because they would have to be priced out at each iteration.

Therefore, we incorporate only a small portion of paths that have positive reduced cost in initial iterations. We control the number of paths that are made available to RMP, based on the ratio of their reduced costs to the most favorable reduced cost. For example, if the criterion for this ratio is set to 1%, only those improving paths whose reduced costs are within 1% of the most favorable reduced cost are incorporated. As the number of CG iterations increases, our algorithm increases this criterion linearly from 1% up to 100%, with the expectation that dual variable values will provide more useful information as the search approaches the optimal RMP solution. We determine the iteration number at which the criterion becomes 100%, through preliminary experiments for each instance (see Section 4.4.2).

4.3. Partial objective constraint

This section introduces POC, an inequality based on the portion of objective function (1) associated with binary variables. Subsections describe the definition of POC, its properties, and a method to obtain right-hand-side values for POC.

4.3.1 Definition of POC

Let $H = \{(f,r): f \in F_{OP}, r \in R_f\}$, $I_x \subseteq H$, and $I_y \subseteq A$. Binary variables, the indices of which are in I_x or I_y, are employed to construct POC. Then, objective function (1) can be re-expressed:

$$Z^* = \text{Max}\,\{R - B\}, \tag{4.19}$$

where $R = -\sum_{(f,r)\in H\setminus I_x} C^O_{fr} x_{fr} - \sum_{(f,r)\in H} V_{fr} q_{fr} - \sum_{a\in A_D\setminus I_y} C^T_a y_a + \sum_{s\in A^d} C_s z_s$; and

$B = \sum_{(f,r)\in I_x} C_{fr}^{O} x_{fr} + \sum_{a\in I_y} C_a^T y_a$.

At the optimal solution of model 1, $(\boldsymbol{q}^*, \boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*)$, the optimal objective function value Z^* can be represented as

$$Z^* = R^* - B^*, \tag{4.20}$$

where $R^* = -\sum_{(f,r)\in \mathrm{H}\setminus I_x} C_{fr}^{O} x_{fr}^* - \sum_{(f,r)\in \mathrm{H}} V_{fr} q_{fr}^* - \sum_{a\in A\setminus I_y} C_a^T y_a^* + \sum_{s\in A^d} C_s z_s^*$ and

$B^* = \sum_{(f,r)\in I_x} C_{fr}^{O} x_{fr}^* + \sum_{a\in I_y} C_a^T y_a^*$.

Letting Z_{inc} denote the objective function value of the incumbent solution of mixed integer program found during the B&B procedure, $R^* - B^* = Z^* \geq Z_{inc}$, so that

$$B^* \leq R^* - Z_{inc} \leq UB\_POC. \tag{4.21}$$

By using this relationship in (21) between a portion of the optimal objective function value, B^*, and *UB_POC*, we define POC based on B in (19) and *UB_POC* in (21):

$$B = \sum_{(f,r)\in I_x} C_{fr}^{O} x_{fr} + \sum_{a\in I_y} C_a^T y_a \leq UB\_POC, \tag{4.22}$$

Note that different POCs can be defined based on the definitions of subsets I_x and I_y.

4.3.2 POC properties

Here, we describe several properties of POC.

Property 4. POC may cut off some portion of the B&B tree but not the optimal integral solution.

Proof. If the value of *UB_POC* were greater than B^* and it be decreased, POC will tighten the restriction on binary variables in sets I_x and I_y so that some feasible integral solutions of model 1 can be infeasible to POC and can be cut off by POC. To prove the second part of Property 4, we show that, incorporating POC, re-optimizing model 1 at the optimal solution prescribes the same optimal solution as that of model 1. Let $(\boldsymbol{q}^*, \boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*)$ be the optimal solution of model 1. Then, the optimal objective function value of model 1 is $R^* - B^* = R^* - (\sum_{(f,r)\in I_x} C_{fr}^{O} x_{fr}^* + \sum_{a\in I_y} C_a^T y_a^*)$. Since POC does not involve any variables associated with R^*, incorporating POC

into model 1 does not affect R^*. Also, POC restricts $\sum_{(f,r)\in I_x} C_{fr}^O x_{fr} + \sum_{a\in I_y} C_a^T y_a$ by using the upper bound of $\sum_{(f,r)\in I_x} C_{fr}^O x_{fr}^* + \sum_{a\in I_y} C_a^T y_a^*$, UB_POC. Thus, the optimal solution of binary variables in sets I_x and I_y, (i.e., $x_{fr}^*, (f,r) \in I_x$;and $y_a^*, a \in I_y$) is still feasible to POC. Therefore, POC of Equation (4.22) does not cut off the optimal integral solution. ■

Property 5. POC may tighten the bound provided by the linear relaxation of model 1.

Proof. This is trivial. Let Z_{LP}^* be the optimal objective function value of the linear relaxation of model 1 and $Z_{LP\_POC}^*$ be that of the linear relaxation of model 1 after incorporating POC. Since POC can cut off the optimal solution of the linear relaxation of model 1, $Z_{LP\_POC}^* \leq Z_{LP}^*$. ■

However, even though POC offers these favorable properties, an appropriate value of *UB_POC* must be determined to be effective in tightening bounds without cutting off the optimal integral solution. The following section presents a method to determine an appropriate value of *UB_POC*.

4.3.3 A method to obtain *UB_POC*

We consider an upper bound $UB\_R^* \geq R^*$:

$$\sum_{(f,r)\in I_x} C_{fr}^O x_{fr} + \sum_{a\in I_y} C_a^T y_a \leq R^* - Z_{inc} \leq UB\_R^* - Z_{inc} \tag{4.23}$$

Upper bound UB_R^* can be calculated by solving problem P_{UB}, which is the same as model 1, except the objective function is changed from $max\ \{R - B\}$ to $max\ \{R\}$.

$$\mathbf{P_{UB}}\text{: } Z_{UB}^* = max\ \{R \mid \text{s.t. } (4.1) - (4.7)\} \tag{4.24}$$

Proposition 6. $R_{(20)}^* \leq Z_{UB}^*$, where $R_{(20)}^*$, defined under Equation (4.20), is a portion of the optimal objective function value of model 1 and Z_{UB}^* is the optimal objective function value of P_{UB}.

Proof. By way of contradiction, suppose that there exist optimal solutions $\boldsymbol{u}^*$ of model P_{UB} and $\boldsymbol{v}^*$ of model 1, respectively, such that $R_{(20)}(\boldsymbol{v}^*) = R_{(20)}^* > Z_{UB}^* = R_{(24)}(\boldsymbol{u}^*)$, where $R_{(20)}(\boldsymbol{v}^*)$ is

the portion (defined as $R_{(20)}$) of the objective solution value of model 1 associated with its optimal solution $\boldsymbol{v}^*$, and $R_{(24)}(\boldsymbol{u}^*)$ is the objective function value of model P_{UB} evaluated at its optimal solution $\boldsymbol{u}^*$. Since the constraints of model 1 constitute model P_{UB}, $\boldsymbol{v}^*$ is also a feasible solution with respect to model P_{UB}. So, this implies that $R^*_{(20)} = R_{(20)}(\boldsymbol{v}^*) = R_{(24)}(\boldsymbol{v}^*) \leq R_{(24)}(\boldsymbol{u}^*) = Z^*_{UB}$, where $R_{(24)}(\boldsymbol{v}^*)$ is the objective function value of P_{UB} associated with feasible solution $\boldsymbol{v}^*$. The inequality follows from the fact that $\boldsymbol{v}^*$ is a feasible solution to model P_{UB}, a maximizing problem, and $\boldsymbol{u}^*$ is its optimal solution. This contradicts the assumption that $R^*_{(20)} > Z^*_{UB}$. ■

Let $Z^*_{UB\_lp}$ be the optimal objective function value of the linear relaxation P_{UB_lp} of P_{UB}. Then, $Z^*_{UB\_lp} \geq Z^*_{UB}$. To determine the value of Z^*_{UB}, problem P_{UB}, a mixed-integer program, must be solved. Therefore, even though $Z^*_{UB\_lp} \geq Z^*_{UB}$, it is attractive to use $Z^*_{UB\_lp}$ as an upper bound on R^*, because we can obtain it easily by solving a linear programming problem as described in Figure 12.

(1) Solve the linear relaxation of model 1 at the root node of the B&B tree.
(2) Set the objective coefficients of binary variables in selected sets I_x and I_y to zero, creating an instance of problem P_{UB_lp}.
(3) Solve problem P_{UB_lp}, using the Dual Simplex method, starting with the optimal root-node solution as the initial feasible solution.

Figure 12 Procedure to calculate $\boldsymbol{Z^*_{UB\_lp}}$

We now re-express POC:

$$\sum_{(f,r)\in I_x} C^O_{fr} x_{fr} + \sum_{a\in I_y} C^T_a y_a \leq Z^*_{UB\_lp} - Z_{inc}. \tag{4.25}$$

$Z_{inc} \geq 0$ in model 1 because a solution in which all decision variables are zero is feasible. In addition, whenever a new integral incumbent solution is found during the B&B procedure, the right-hand side of inequality (4.25) can be reduced, tightening POC.

4.3.4 Selection of I_x and I_y

This paper generate two types of POCs, each based on a particular selection of subsets I_x and I_y.

POC1: $\sum_{(f,r)\in H} C_{fr}^{O} x_{fr} + \sum_{a\in A} C_a^T y_a \leq UB\_POC_1$,

includes all binary variables (with $I_x = H$ and $I_y = A$); and

POC2: $\sum_{(f,r)\in H} C_{fr}^{O} x_{fr} \leq UB\_POC_2$,

includes only binary variables x_{fr} (with $I_x = H$ and $I_y = \emptyset$), which are associated with facility locations and technology types and have large objective function coefficients in comparison with those of binary variables y_a.

4.4. Computational tests

The objectives of our computational tests are to evaluate the efficacy of our solution approach and benchmark against a state-of-the-art commercial solver. We employ CPLEX 12.1 and C++ with Callable Library under the Windows 7 64-bit operating system with an Intel(R) Core(TM)2 Quad CPU Q9650 @ 3.00 GHz and a RAM of 8GB.

Our experiment is based on a case study reported by An et al. (2011b), which involves nine counties in the Central Texas Region. This region is representative of many others for which cellulosic biomass must be supplemented with energy crops to its meet fuel demand. An et al. (2011b) conducted a sensitivity analysis for several economic factors, including cost and supply of feedstock (e.g., switchgrass, whose moisture content ranges generally from 20 to 60%), and price and demand of ETOH. All parameter values used in our computational tests are from Chapter III.

Table 12 shows 15 test instances, which we generated based on two factors: the numbers of farms and time periods (e.g., quarters, bi-months, months). Our tests involve five levels (9, 12, 15, 25 and 34) of the number-of-farms factor, and three levels (4, 6 and 12) - representing quarters, bi-months and months- of the number-of-time-periods factor.

Table 12 Test instances

No	Name	# Farms	# Periods	# Rows	Variables				Embedded GFP Network	
					Binary *x*	Binary *y*	Binary Total	Continuous	# Nodes	# Arcs
1	F9T4	9	4	2,813	81	648	729	5,150	902	5,067
2	F9T6	9	6	4,205	81	972	1,053	7,760	1,352	7,677
3	F9T12	9	12	8,381	81	1,944	2,025	15,590	2,702	15,507
4	F12T4	12	4	4,178	108	1,152	1,260	8,594	1,202	8,484
5	F12T6	12	6	6,248	108	1,728	1,836	12,938	1,802	12,828
6	F12T12	12	12	12,458	108	3,456	3,564	25,970	3,602	25,860
7	F15T4	15	4	5,760	135	1,800	1,935	12,902	1,502	12,765
8	F15T6	15	6	8,615	135	2,700	2,835	19,412	2,252	19,275
9	F15T12	15	12	17,183	135	5,400	5,535	38,942	4,502	38,805
10	F25T4	25	4	12,589	225	5,000	5,225	33,502	2,502	33,275
11	F25T6	25	6	18,845	225	7,500	7,725	50,352	3,752	50,125
12	F25T12	25	12	37,613	225	15,000	15,225	100,902	7,502	100,675
13	F34T4	34	4	20,788	306	9,248	9,554	60,250	3,402	59,942
14	F34T6	34	6	31,130	306	13,872	14,178	90,510	5,102	90,202
15	F34T12	34	12	62,156	306	27,744	28,050	181,290	10,202	180,982

The following subsections describe the design of our experiments, report the runtime to solve the linear relaxation of model 2 at the root node of the B&B search tree, and present our overall evaluation based on the 15 test instances described in Table 12.

4.4.1. Test procedure

Our approach uses a CG approach to solve the linear relaxation of model 2 but not to prescribe bounds at nodes in the B&B search tree, because our preliminary tests showed that using CG to prescribe bounds at all nodes is not as fast as CPLEX B&B. We conjecture that CPLEX is faster - at least in part - because it employs an early termination criterion that allows

the dual simplex algorithm to stop before reaching an optimal solution. In contrast, CG must use the primal simplex algorithm, which must be solved until finding at least a near-optimal solution. CPLEX also enjoys the advantages of having been coded by professional programmers and having been refined over a long period of time.

We believe that our experience is not unique: most studies since the mid 2000s have not compared their branch-and-price algorithms with commercial solvers. Rather, their computational tests compared different branch-and-price algorithms (e.g., Senne et al., 2005; Villa and Hoffman, 2006; Grønhaug et al., 2010; Gutierrez-Jarpa et al., 2010; Irnich, 2010; and Brunner et al., 2011). Some papers that have compared branch-and-price performance with a commercial solver (Brunner et al., 2011; and Mestry et al., 2011) reported only the gaps provided by two codes within a given runtime limit. We speculate that the capabilities of commercial solvers may have been improved during last decade to the point at which only branch-and-price codes written by professional programmers can be competitive; otherwise, they cannot be competitive, even though they may offer certain theoretical advantages.

Therefore, rather than using CG to prescribe bounds, this paper presents an approach, which solves BSCP faster than state-of-the-art solver CPLEX 12.1, using our BRA in CG at the root node of the search tree and CPLEX defaults with POC during the B&B search.

Figure 13 structures our solution procedure, elements of which are discussed in sections 3 and 4. In step 1, CG solves the linear relaxation of model 2 using CPLEX to solve RMP and RBA to solve the sub-problem at each CG iteration. Step 2 calculates $Z^*_{UB\_lp}$ for POC(s) by solving P_{UB_lp} as described in section 4. Step 3 generates POC, using $Z^*_{UB\_lp}$ from Step 2. Steps 2 and 3 are iterated for selected subsets I_x and I_y. By incorporating POC1 or POC2, step 4 modifies model 1, forming model 1'. Step 5 uses CPLEX B&B logic to solve model 1' at nodes after than the root node so that CPLEX defaults control the branching process. Note that CPLEX

does not allow a right-hand-side constant to be changed after its branching logic starts. Thus, our computational tests do not update Z_{inc} to strengthen *UB_POC* when a new incumbent solution is found; rather, we maintain $Z_{inc} = 0$ or the incumbent solution value solved by using CPLEX heuristics in Step 2-1.

Step1. Solve the linear relaxation of model 2 using CG.
Step2. Modify objective function coefficients and solve P_{UB_lp} using CG, starting with the optimal root node solution as the initial feasible solution to obtain the *UB_POC*.
(Step2-1. Solve model 1 by using CPLEX heuristics to strengthen UB_POC using incumbent solution.)
Step3. Generate POC using *UB_POC*. Iterate steps 2 and 3 for selected subsets I_x and I_y, generating associated POC(s).
Step4. Incorporate POC(s) into the model 1 only at the root node, forming model 1'.
Step5. Solve the augmented model 1' by using CPLEX B&B logic.

Figure 13 Solution procedure

4.4.2 Solving the linear relaxation of model 2 using CG

Table 13 compares runtimes required to solve the linear relaxation of model 2 using both CPLEX and our CG approach. The first column gives the names of test instances and the second column presents the optimal objective function value of the linear relaxation of model 2, which is the same value as that of model 1, for each instance. Columns 3 and 4 (5 and 6) give CPLEX (CG) results (i.e., runtime and number of simplex iterations). Time Reduction (columns 7 and 8) gives two measures to compare runtimes of CPLEX and CG: 'A-B' is the CPLEX runtime minus that of CG; and '100*B/A' gives the percentage of CPLEX runtime required by our CG approach. The last two columns give CG iteration numbers: CG iteration number for 100% criterion (columns 9), after which 100% improving columns found are incorporated into RMP as described in section 3.5; and one for the total number of CG iterations (columns 10) to obtain the optimal solution.

Table 13 Comparison of CPLEX and CG for the linear relaxation of model 2

Name	Optimal value	CPLEX		CG		Time Reduction		CG iteration number	
		Time (sec)	# Simplex Iter.	Time (sec)	# Simplex Iter.	A-B (sec)	100*B/A (%)	100% criterion	Total
F9T4	1.719960E+07	0.22	2,000	0.16	2,410	0.06	71.4%	3	23
F9T6	1.844870E+07	1.03	7,861	0.55	5249	0.48	53.0%	3	37
F9T12	1.173080E+07	4.29	22,256	2.48	19,333	1.81	57.8%	1	49
F12T4	2.344030E+07	0.90	7,130	0.44	4,513	0.46	48.4%	3	26
F12T6	2.462720E+07	2.43	13,923	1.23	8,379	1.20	50.7%	9	48
F12T12	9.964422E+06	8.53	30,950	5.49	6,377	3.04	64.4%	1	90
F15T4	1.143080E+07	1.12	7,519	0.28	2,846	0.84	25.0%	3	22
F15T6	1.240550E+07	3.09	14,875	1.00	4,959	2.09	32.3%	1	44
F15T12	3.520170E+07	28.55	69,378	22.33	87,803	6.23	78.2%	21	113
F25T4	1.248940E+07	4.04	13,784	0.69	4,411	3.35	17.0%	3	30
F25T6	1.362290E+07	14.24	32,384	3.32	15,249	10.92	23.3%	3	55
F25T12	1.477820E+07	124.38	60,410	32.01	86,287	92.37	25.7%	1	95
F34T4	1.007000E+07	10.73	21,208	1.21	6,537	9.52	11.3%	2	23
F34T6	3.541140E+07	64.60	72,608	15.62	49,942	48.99	24.2%	3	56
F34T12	1.624580E+07	328.28	187,251	66.74	13,732	261.54	20.3%	1	111

Our CG approach solves all test instances of the linear relaxation of model 2 faster than CPLEX. As the instance size increases, the ratio of CG runtime to that of CPLEX decreases; that is, the runtime advantage of CG increases with instance size in these tests. Note that the CG iteration number of 100% criterion, which was determined by preliminary computational tests to manage the number of improving columns incorporated, is generally less than 20% of the total CG iteration number.

4.4.3 Solving BSCP using CG and POC(s)

Table 14 presents several solution values to generate POC(s). The first column again notes the instance. The first three columns give optimal solution values of (linear relaxation) model 1 and their gaps. Columns 5 and 6 give the left-hand-side values of POC1 and POC2 at the optimal solution of model 1. Columns 7 and 8 present the optimal solution value of model $P_{UB}$with respect to POC1 and POC2. By using the values in columns 7 and 8, and the trivial

incumbent solution $Z_{inc} = 0$, we determine the right-hand-side values of POC1 and POC2 (columns 9 and 10). Columns 11-13 give the incumbent solution value solved by CPLEX heuristics, their gaps, and their runtimes. Finally, columns 14 and 15 present the right-hand-side values of POC1' and POC2' strengthened by the incumbent solution value in columns 11 and 12.

The gaps between Z_{LP} (column 2) and Z_{IP} (column 3) range from 41% to 453%. Since POC1 and POC2 use the trivial incumbent solution $Z_{inc} = 0$, each of their right-hand-side values is the same as the optimal solution values of model P_{UB}. The gaps between Z_{IP} (column 3) and the incumbent solution values solved by CPLEX heuristics (column 12) range from 38% to 100%. For instances F15T4 and F34T4, CPLEX heuristics found only the trivial incumbent solution (i.e., $Z_{inc} = 0$).

Table 15 compares the runtimes of CG with POC(s) against the benchmarking CPLEX B&B. The first column gives the names of test instances. A group of seven columns, CPLEX, POC1, POC2, POC1&2 (i.e., both POC1 and POC2), POC1', POC2', and POC1'&2', gives runtimes for each of the two methods we tested: without using CPLEX cuts and with using CPLEX cuts. Note that the runtimes associated with POC(s) involve the runtimes of CG.

In the first group of columns in Table 15 (i.e., without CPLEX cuts), our solution approach based on POC is faster than CPLEX with a few exceptions: two instances (F12T12 and F34T6) for POC1; four instances (F9T4, F9T6, F25T4 and F34T6) for POC2; and three instances (F9T4, F15T6 and F25T4) for POC1&2. Figure 14 graphs the ratio of runtime of POC1, POC2, and POC1&2 to CPLEX runtime on each test instance without using CPLEX cuts. A ratio less than 1.0 indicates a method that is faster than CPLEX. POC1 is less effective than

Table 14 Solution values to generate POC1 (POC2) and strengthened POC1' (POC2')

Name	Z_{LP}	Z_{IP}	GAP[(a)]	LHS*[(b)] of POC1	LHS*[(b)] of POC2	$R^*(24)\_1$[(c)]	$R^*(24)\_2$[(c)]	RHS of POC1	RHS of POC2	Z_{inc}_H[(d)]	GAP_H[(e)]	Time for Z_{inc}_H	RHS of POC1'	RHS of POC2'
F9T4	1.71996E+7	9.78245E+6	76%	4.70541E+6	4.69581E+6	1.87632E+7	1.87628E+7	1.87632E+7	1.87628E+7	5.20740E+6	47%	1.4	1.35558E+7	1.35554E+7
F9T6	1.84487E+7	9.52891E+6	94%	4.78806E+6	4.78006E+6	1.98643E+7	1.98641E+7	1.98643E+7	1.98641E+7	5.87931E+6	38%	5.2	1.39850E+7	1.39848E+7
F9T12	1.17308E+7	5.14870E+6	128%	4.70211E+6	4.69581E+6	1.30093E+7	1.30092E+7	1.30093E+7	1.30092E+7	2.13364E+6	59%	36.2	1.08757E+7	1.08756E+7
F12T4	2.34403E+7	1.42319E+7	65%	4.96325E+6	4.94855E+6	2.51494E+7	2.51490E+7	2.51494E+7	2.51490E+7	4.95891E+6	65%	2.4	2.01905E+7	2.01901E+7
F12T6	2.46272E+7	1.45254E+7	70%	4.79046E+6	4.78006E+6	2.62423E+7	2.62420E+7	2.62423E+7	2.62420E+7	7.08008E+6	51%	8.6	1.91622E+7	1.91619E+7
F12T12	9.964422E+6	4.01397E+6	148%	4.70321E+6	4.69581E+6	1.07575E+7	1.07575E+7	1.07575E+7	1.07575E+7	1.09415E+5	97%	59.2	1.06481E+7	1.06481E+7
F15T4	1.14308E+7	4.30164E+6	166%	4.70961E+6	4.69581E+6	1.19664E+7	1.19663E+7	1.19664E+7	1.19663E+7	0.00000E+0	100%	2.1	1.19664E+7	1.19663E+7
F15T6	1.240550E+7	5.67554E+6	119%	4.70801E+6	4.69581E+6	1.31053E+7	1.31052E+7	1.31053E+7	1.31052E+7	3.50819E+5	94%	17.1	1.27545E+7	1.27544E+7
F15T12	1.477820E+7	7.62659E+6	94%	4.62327E+6	4.61157E+6	1.57969E+7	1.57968E+7	1.57969E+7	1.57968E+7	1.23092E+7	51%	436.0	2.51008E+7	2.51006E+7
F25T4	1.24894E+7	4.79253E+6	161%	4.71591E+6	4.69581E+6	1.30858E+7	1.30857E+7	1.30858E+7	1.30857E+7	4.03922E+6	16%	10.3	9.04658E+6	9.04648E+6
F25T6	1.36229E+7	6.33165E+6	115%	4.71381E+6	4.69581E+6	1.43774E+7	1.43773E+7	1.43774E+7	1.43773E+7	3.74868E+5	94%	44.1	1.40025E+7	1.40024E+7
F25T12	3.74989E+7	2.66001E+7	41%	9.32388E+6	9.30938E+6	3.98109E+7	3.98107E+7	3.98109E+7	3.98107E+7	2.11335E+6	72%	1658.9	1.36836E+7	1.36835E+7
F34T4	1.00700E+7	1.81969E+6	453%	4.63587E+6	4.61157E+6	1.05612E+7	1.05610E+7	1.05612E+7	1.05610E+7	0.00000E+0	100%	21.1	1.05612E+7	1.05610E+7
F34T6	3.54114E+7	2.27045E+7	56%	9.32758E+6	9.30938E+6	3.74207E+12	3.74204E+7	3.74207E+7	3.74204E+7	6.98688E+6	69%	557.0	3.04338E+7	3.04335E+7
F34T12	1.624580E+7	8.36782E+6	94%	4.62697E+6	4.61157E+6	4.62697E+6	4.61157E+6	4.62697E+6	4.61157E+6	1.88906E+6	77%	458.6	1.54430E+7	1.54429E+7

(a) GAP = 100 * (Z_{LP} -Z_{IP})/ Z_{IP}

(b) LHS* of POC1 (POC2): LHS of POC1 (POC2) at optimal solution

(c) $R^*(24)\_1$ (_2): $R^*(24)$ for POC1 (POC2)

(d) Incumbent solution value by using CPLEX heuristics

(e) GAP_H = 100 * (Z_{IP} - Z_{inc}_H)/ Z_{IP}

Table 15 Comparison of runtimes

Name	Without CPLEX cuts (sec)							With CPLEX cuts (sec)						
	CPLEX	**POC1**	**POC2**	**POC1&2**	**POC1'**	**POC2'**	**POC1'&2'**	**CPLEX**	**POC1**	**POC2**	**POC1&2**	**POC1'**	**POC2'**	**POC1'&2'**
F9T4	4.1	4.0	4.2	4.2	4.0	3.9	4.0	39.7	11.0	25.3	18.3	10.84	25.2	18.3
F9T6	10.5	8.9	10.7	9.2	9.1	10.0	8.5	62.5	32.8	71.2	45.0	32.59	71.0	44.5
F9T12	46.6	37.5	25.9	26.8	39.8	22.6	23.7	227.2	223.7	175.1	184.7	222.56	161.0	184.2
F12T4	60.8	28.9	30.6	34.4	28.9	29.9	33.9	118.8	120.3	133.8	116.3	119.7	133.9	115.5
F12T6	47.7	29.9	45.8	32.3	30.7	44.2	30.8	163.9	329.9	288.2	252.6	328.5	360.8	222.6
F12T12	205.0	337.3	106.9	91.2	337.6	105.8	91.4	850.9	683.0	672.8	943.4	683.0	673.6	945.0
F15T4	14.9	12.2	12.1	10.1	12.2	12.1	10.1	229.0	147.1	173.5	192.2	147.1	173.5	192.2
F15T6	37.4	39.7	25.6	38.6	38.9	24.1	40.3	373.2	544.6	489.9	393.6	543.1	490.8	397.3
F15T12	724.2	716.1	714.3	464.3	716.3	718.2	465.5	3997.1	3765.5	3744.8	4845.3	4528.6	3696.1	2605.7
F25T4	142.9	98.3	263.9	162.4	96.6	228.4	239.8	1546.0	5546.2	2290.3	3273.1	1993.3	2719.21	3121.4
F25T6	651.7	478.1	580.2	553.3	473.7	580.2	555.1	4007.7	>7200	5724.6	3262.2	>7200	6732.7	>7200
F25T12	2199.7	1189.0	2129.4	1176.7	1277.4	2138.3	1141.0	>7200	>7200	>7200	>7200	>7200	>7200	>7200
F34T4	1184.1	526.2	642.3	581.1	526.2	581.1	526.2	>7200	>7200	>7200	>7200	>7200	>7200	>7200
F34T6	2589.8	4907.5	3397.1	2164.0	4936.2	5893.0	2891.9	>7200	>7200	>7200	>7200	>7200	>7200	>7200
F34T12	>7200[(a)]	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200	>7200

(a) >7200.0: optimal solution was not found within the time limit of 7200 sec

POC2 and POC1&2 in our tests. It is interesting to note that the runtime of POC1&2 is between or better than those of POC1 and POC2, with the exceptions of F12T4. Average ratios of runtime of our methods to that of CPLEX are 90% for POC1, 91% for POC2, 76% for POC1&2, respectively, implying that POC1&2 outperforms, on average, CPLEX, POC1, and POC2.

The strengthened POC2' is faster than POC2 for most instances, with the exception of F15T12 and F34T6. In contrast, the strengthened POC1' and both POC1'&2' do not reduce runtimes in our tests. From the results in the second group of columns in Table 15, solving with CPLEX cuts is slower than solving without CPLEX cuts for all instances. Especially, for four instances (F25T12, F34T4, F34T6, and F34T12), it did not find the optimal solution within 7200 seconds.

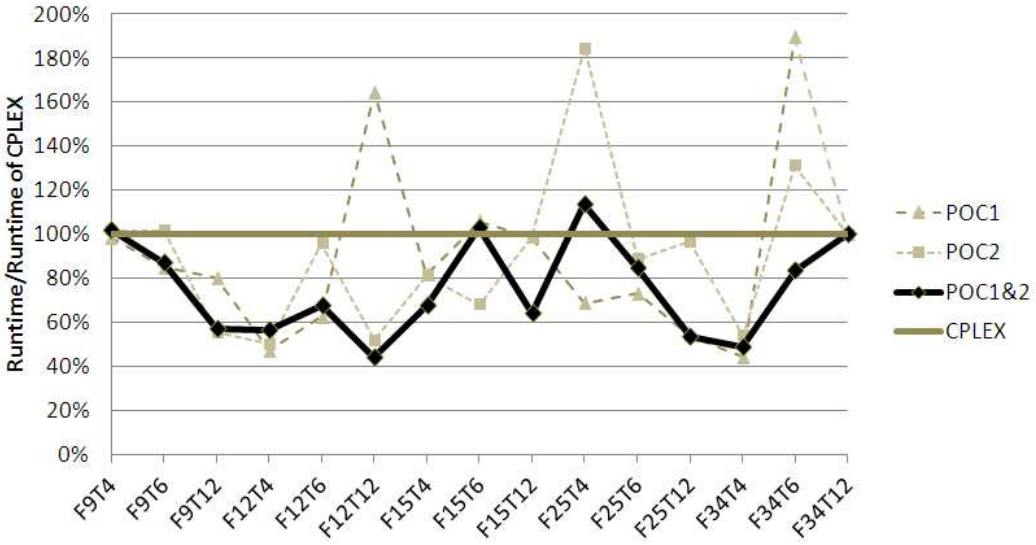

Figure 14 Comparison of four solution methods

To investigate the impact of the right-hand-side value of POC1, we solved the instance F9T6 by using strengthened right-hand-side values of POC1. Table 16 compares the performance of POC1 with various right-hand-side values. The first column gives the test

number and the second, the name of POC. Columns 3 and 4 present the right-hand-side values and their gaps. Columns 5-7 provide performance measures (i.e., runtime, simplex iterations and nodes examined).

Row 1 gives the result using CPLEX without POC1. Row 2 is associated with POC1 and row 3, with POC1', which has a strengthened right-hand-side value by using CPLEX heuristics. From row 4 to 6, we strengthened the right-hand-side value of POC1. Row 7 gives the result by using the optimal right-hand-side value of POC1. As the right-hand-side value of POC1 decreases up to the optimal solution, the runtime, simplex iterations, and nodes examined decrease as well. This result implies that the performance of POC(s) depends on the quality of the right-hand-side value and such inequalities may be helpful in accelerating B&B.

Table 16 Comparison of various strengthened right-hand-side values of POC1 for instance F9T6

No	POC	RHS of POC1	GAP_RHS	Runtime (sec)	Iteration	Nodes
1	-	-	-	10.50	62935	1781
2	POC1	19864300	315%	8.9	55641	1126
3	POC1_H[(a)]	13984993	192%	8.9	55641	1126
4	POC1_a	10000000	109%	7.5	49333	1018
5	POC1_b	8000000	67%	6.3	40349	884
6	POC1_c	6000000	25%	6.3	40349	884
7	POC1_*[(b)]	4788056	0%	6.2	39839	725

(a) By using CPLEX Heuristics
(b) Optimal

4.4.4 Analysis of the performance of POC(s)

This section analyzes and compares the performances of POC1, POC2, and POC 1&2 based on a few typical test instances (i.e., F12T4, F12T6, and F12T12). Each following subsections compare the detail performances of those solution methods, without using CPLEX cuts, for several cases: the performance of POC1&2 is better than those of POC1 and POC2; that of POC1&2, between those; and that of POC1&2, worse than those.

4.4.4.1 The performance of POC1&2 is better than those of POC1 and POC2 (F12T12)

Figure 15 compares best bound, best integer, and the number of iterations based on four solution methods for instance F12T12 over the entire B&B procedure. In terms of runtime, POC1&2 is best among all methods; POC2 is second; CPLEX, third; and POC1, the worst. The best bound of CPLEX decreases slower than others and the gap of CPLEX decreases slower than others. Note that, by default, CPLEX has a relative MIP gap of 10^{-4} (0.01%) and an absolute MIP gap of 10^{-6}, and the number of simplex iterations in this paper does not include the iterations caused by CPLEX strong branching.

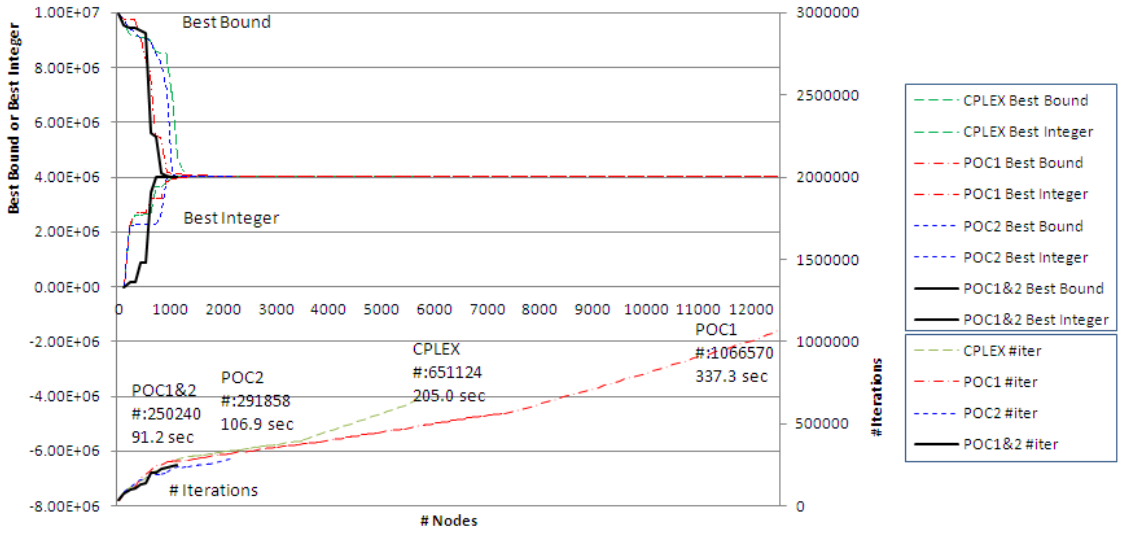

Figure 15 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T12, Scope: the entire B&B search)

One interesting observation is that all methods show very small gaps between best bound and best integer after about 1500th node. In the initial stage of the B&B search, binary variables associated with opening facility (i.e., x_{fr}), which have relatively large objective coefficients, might be selected as a branching variable because CPLEX branching rule may tend

to branch first on a variable which has a large impact on the objective function value. Followed by those binary variables, binary variables y_a associated with selecting arc, which have small objective coefficient (e.g., 300) compared to x_{fr} (e.g., 4,511,168 and 60,000), may be selected as a branching variable. Even though fixing y_a to zero may make some impact on material flows, the impact of fixing it to one on the solution and the objective function values may be very small. In addition, the number of binary variables y_a (e.g., 3,564 for F12T12) is much larger than that of x_{fr} (e.g., 108 for F12T12). Thus, in some unfortunate case, while fixing y_a variables, it could take long time to reach the relative optimality gap less than 0.01%.

Figure 16 enlarges Figure 15 for the range between 1st and 1000th node in the B&B search. POC1&2 shows the smallest best bound in the initial stage (up to 100th node) of the B&B search. Even though the best integer of POC1&2 increases slower than other methods, the gap between best bound and best integer decreases faster than others (see the result on 800th node in Figure 3). This earlier convergence of POC1&2 than others might result in earlier termination. However, even though the gap of POC1 between best bound and best integer decreases faster than CPLEX, POC1 searched larger number of nodes, resulting in larger runtime and simplex iterations than CPLEX. This might be associated with an unlucky B&B search path as well as large B&B search space according to the large number of y_a variables.

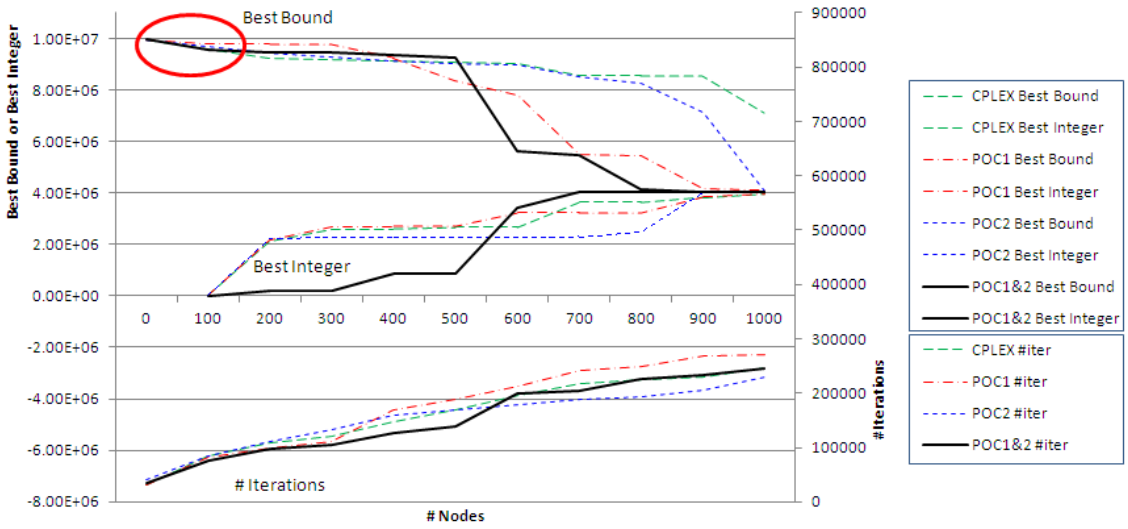

Figure 16 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T12, Scope: between 1st and 1000th node in the B&B search)

4.4.4.2 The performance of POC1&2 is between those of POC1 and POC2 (F12T6)

Figure 17 compares best bound, best integer, and the number of iterations based on four solution methods for instance F12T6 over the entire B&B procedure. Overall convergence pattern is similar to that of F12T12. POC1 is the best among all methods; POC1&2 is second; POC2, third; and CPLEX, the worst. The gap of CPLEX converges slower than others.

Even though the number of nodes searched of POC2 is greater than that of CPLEX, the number of simplex iterations of CPLEX is greater than that of POC2, resulting in that runtime of CPLEX is greater than that of POC2. This may be because the nodes in B&B search path of CPLEX require more simplex iterations than those in the B&B search path of POC2.

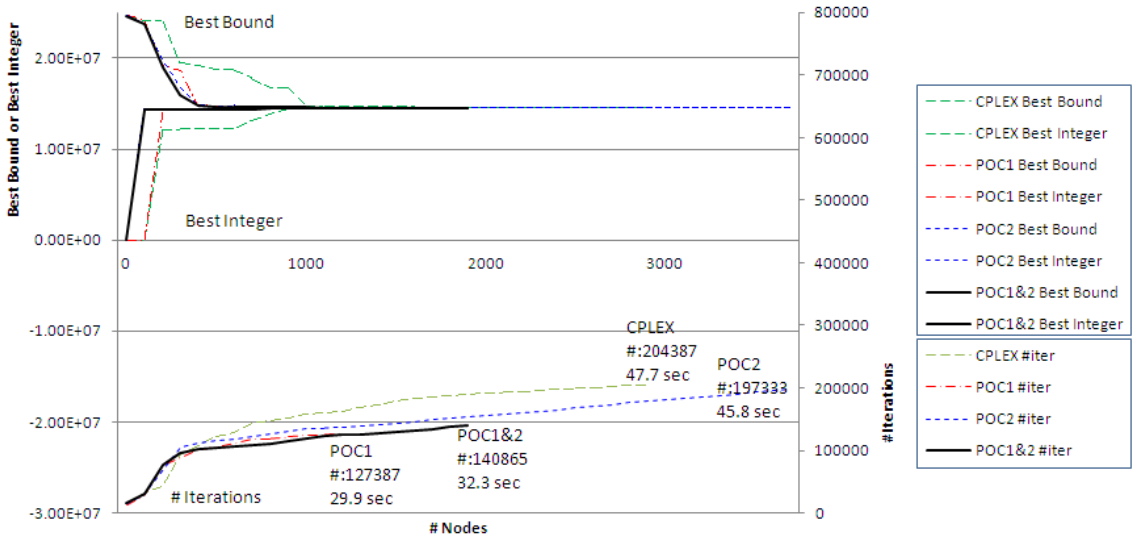

Figure 17 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T6, Scope: the entire B&B search)

Figure 18 enlarges Figure 17 for the range between 1st and 1000th node in the B&B search. POC1&2 shows the smallest best bound and the largest best integer in the initial stage (up to 100th node) of the B&B search. Even though the gap between best bound and best integer of POC1&2 decreases faster than others, the gaps of POC1, POC2, and POC1&2 become similar on 400th node. After about 500th node, POC1 shows the smallest gap. This faster convergence of POC1 than others may result in earlier termination. One interesting observation is that the number of iterations of CPLEX is smallest among all methods in the initial stage of B&B search before 300th node, but becomes largest after 450th node.

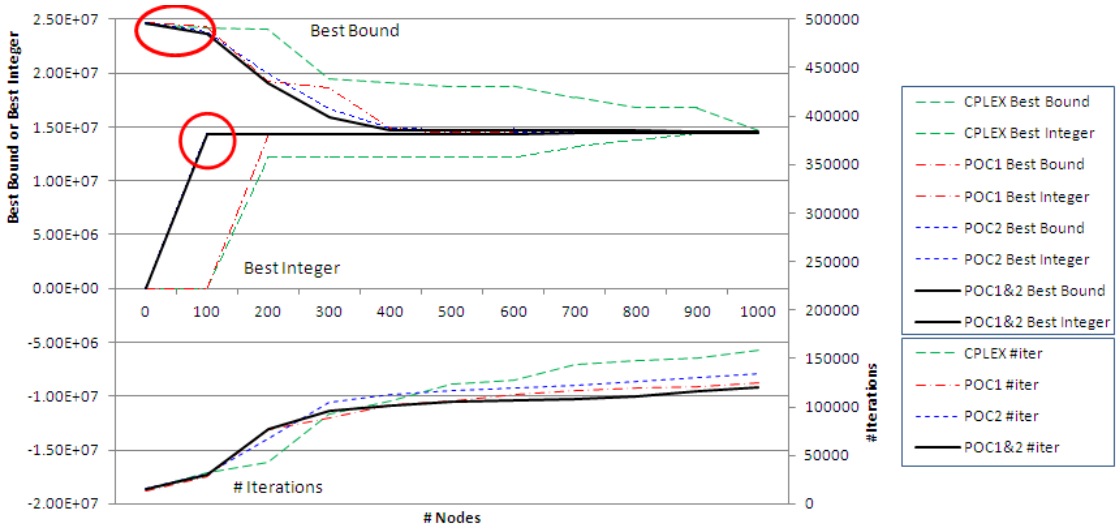

Figure 18 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T6, Scope: between 1st and 1000th node in the B&B search)

4.4.4.3 The performance of POC1&2 is worse than those of POC1 and POC2 (F12T4)

Figure 19 compares best bound, best integer, and the number of iterations based on four solution methods for instance F12T4 over the entire B&B procedure. POC1 is best among all methods; POC2 is second; POC1&2, third; and CPLEX, the worst. Even though CPLEX shows the smallest gap in the initial stage of B&B search, the number of nodes searched of CPLEX is greater than others, resulting in that runtime of CPLEX is the largest among all methods: i.e., it took longer time for CPLEX to reach the relative optimality gap less than 0.01%. This might be due to the unlucky path of CPLEX B&B search.

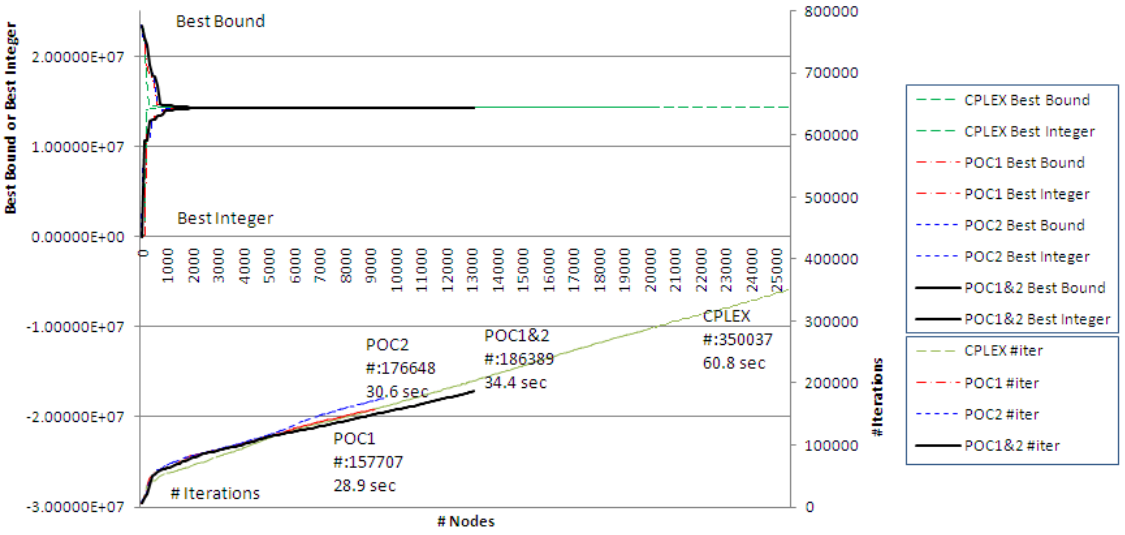

Figure 19 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T4, Scope: the entire B&B search)

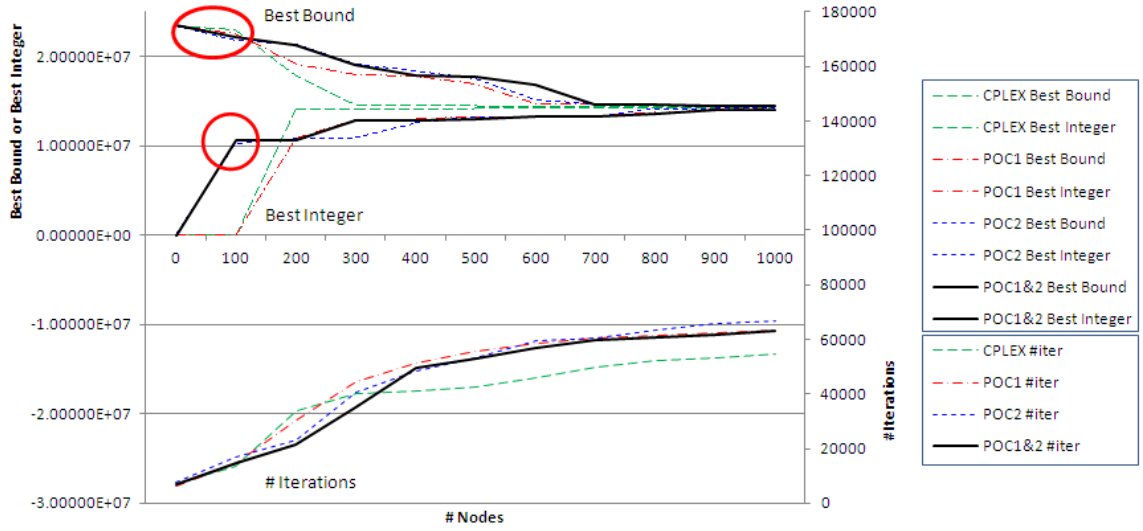

Figure 20 Comparison of Best Bound, Best Integer, and # Iterations based on four solution methods (Instance: F12T4, Scope: between 1st and 1000th node in the B&B search)

Figure 20 enlarges Figure 19 for the range between 1st and 1000th node in the B&B

search. Even though POC1&2 shows the smallest best bound and the largest best integer in the initial stage (up to 100th node) of the B&B search, the gap between best bound and best integer of POC1&2 decreases slower than others after 100th node: i.e., the initial advantage was not helpful to acquire faster convergence. This may result in the larger runtime of POC1&2 than POC1 and POC2.

4.4.4.4 Discussions on the performance of POC

Since the analysis showed that the performance of POC is associated with CPLEX strong branching, here, this paper briefly reviews studies for B&B. Benichou et al. (1971) proposed the concept of *pseudo-cost*, remarking that "though the results are unstable, the comparison show the interest of such a strategy to speed up … ". Linderoth and Savelsbergh (1999) conducted a computational study of B&B search strategies, noting that "there is no one search strategy that will work best on all problem instances". Fischetti and Lodi (2003) presented *local branching*, reporting that their method improved the performance in 23 out of 29 cases. Achterberg et al. (2005) described *reliability branching*, noting that "we did not base our conclusions on performances of single instances and discuss those in detail. We rather rely on average numbers of over all instances". Glankwamdee and Linderoth (2006) studied *lookahead branching*, noting that "the intuition behind our study is to view strong branching as a greedy heuristic for selecting the branching rule". Achterberg and Berthold (2009) proposed *hybrid branching*, comparing the performances of several rules based upon the geometric means of runtime and the number of nodes examined. Ostrowski et al. (2011) introduced *orbital branching*, comparing the performance of branching rules based upon the number of best runtimes. Karamanov and Cornuejos (2011) considered branching on general disjunctions, showing that their method outperforms on average the benchmarking method, with several extreme cases in which their method is worse. Based up on this limited literature review, it

seems that studies for B&B have not focuses on the details about exceptional cases, but reported the average performance.

Linderoth and Savelsbergh (1999) noted that the ability of the selection of a suitable subset of variables on which to perform a number of dual simplex iterations impacts greatly on the effectiveness of strong branching. Thus, we may expect that using POC1&2 may be better than using a single POC (i.e., POC1 or POC2). In effect, on average, POC1&2 outperforms others. However, from Glankwamdee and Linderoth (2006)'s perspective that considers strong branching as a greedy heuristic, it may not be guaranteed that providing better information to strong branching always outperform providing worse information.

CHAPTER V

SIMULATION MODEL: PART1. BIOMASS MODULE LOGISTICS SYSTEM

This chapter introduces a simulation model for a biomass module logistics system which uses large biomass packages (i.e., modules) of sufficient size and density to provide maximized legal highway loads and quick load/unload times. Such a system was being tested conceptually at Texas A&M University, but modeling was used to predict system performance prior to constructing prototypes.

Section 1 presents a background of biomass logistics systems. Section 2 describes the conceptual biomass module system. Sections 3 and 4 give details of IBSAL modeling and a simulation model, respectively. Section 5 gives simulation results and analyzes the sensitivity of each performance factor to the system.

5.1 Background

The conceptual logistics system described here has similarities with the cotton logistics system, which uses large packages commonly called cotton modules. Cotton is normally harvested and collected in a dump basket on the harvester. A dump trailer transports collected material to a module builder located at the edge of a field. The module builder compresses the cotton into large modules (e.g., 2.4 m wide * 2.4 m high * 9.8 m long) (Ravula et al. 2008). The cotton modules, protected from weather with a plastic cover, wait on a field to be transported to gins by module haulers. A recent development is cotton harvesters that form smaller modules on-board the harvester, eliminating the need for the dump trailer, the module builder and the associated labor (Taylor, 2007). Ravula et al. (2008) suggested that it may be possible to apply operational strategies from a cotton logistics system to a biomass logistics system, describing the similarities between those systems.

In many industries, logistics systems use large packages to deliver materials efficiently. Shipping companies use large containers to facilitate material handling and to protect shipping products. Even though its adoption required huge investment and negotiated standards, cost savings from improved material handling efficiency exceeds those investment costs. As a result, factories are located far from customers and low-cost products are shipped around the globe (Levinson, 2008). A disadvantage of the container shipping system is the accumulation of empty containers at the customer locations and the need to find return loads or ship empty containers.

Chapter II provides a review about operational level studies of biofuel SC, most of which estimated performance of biomass logistics systems. In particular, since a bale system has been commonly used to transport hay, several earlier studies (Jenkins et al., 1984; Sokhansanj et al., 2006; Petrou and Mihiotis, 2007; Kumar and Sokhansanj, 2007; Morey et al., 2010; Suh and Suh, 2010) employed a bale system to estimate important measures (e.g., cost and energy) associated with delivering biomass. Table 17 describes highlights of those studies.

Jenkins et al. (1984) estimated logistics costs of several types of biomass (e.g., corn and sorghum stover, wheat and barley straw, and rice straw) in California of the U.S.A. They developed a cost equation that includes a variable cost of transportation, a fixed cost of transportation, and collection and processing cost. The collection and processing costs were $28.02/Mg for corn and sorghum stover, $29.58/Mg for wheat and barley straw, and $35.35/Mg rice straw, respectively. For the straw, the fixed transportation cost was $6.14/Mg and the variable transportation cost was $0.061/Mg-km; and, the fixed and variable transportation costs of the corn and sorghum stover were $9.78/Mg and $0.078/Mg-km, respectively.

Table 17 Estimates of biomass logistics cost in prior studies

Category	Authors	Year	Biomass	Process	Cost			Remark
					Collection & Processing ($/Mg)	Transportation ($/Mg-km or $/Mg)	Total ($/Mg)	
Bale	Jenkins et al.	1984	Rice straw	swath-rec bale-transport-store-tub grind	41.50	0.061 $/Mg-km	41.5+0.061*d	d: distance from farm to a refinery Include storing operation
			Wheat or barley straw	swath-rec bale-transport-store-tub grind	35.72	0.061 $/Mg-km	35.72+0.061*d	
			Corn or Sorghum stover	swath-roll bale-transport-store-tub grind	37.80	0.078 $/Mg-km	37.8+0.078*d	
	Tatsiopoulos IP, Tolis AJ	2003	Cotton stalks	cut-collect(by 3rd party)-transport-drying & baling-store-transport	-	-	58.5[(a)]	Include storing operation
				cut-collect(by farmers)-transport-drying & baling-store-transport	-	-	33.8[(a)]	
	Sokhansanj et al.	2006	Corn stover	combine-shred-bale-stack-load-transport-unload-stack-grind	39.65	13.76 $/Mg	53.57	Distance winding factor: 1.4 32-160 km
	Petrou and Mihiotis	2007	Cotton stalks	cut-bale-transport-store-transport	18.51	0.2 $/Mg-km	62.53	Include payment to contractors
	Kumar and Sokhansanj	2007	Switchgrass	swath-rake-bale(square)-stack-tarp-transport-stack-grind	36.69	-	41-45	Transporting distance: 3-77 km
				swath-rake-bale(square)-stack-tarp-grind-transport	33.70	-	41-55	
				swath-rake-bale(round)-stack-tarp-transport-stack-grind	35.21	-	39-48	
				swath-rake-bale(round)-stack-tarp-grind-transport	32.22	-	39-54	
	Sokhansanj et al.	2010	Corn stover	harvest-bale(square)-stack-transport	38.01	9.98 $/Mg	48.35	Distance: 70 km
	Morey et al.	2010	Corn stover	shred-rake-bale-transport-grind-roll press-transport	43.08	6.40 $/Mg	48.48	Include storing operation Distance: 42 km

Table 17 Continued

Category	Authors	Year	Biomass	Process	Cost			Remark
					Collection & Processing ($/Mg)	Transportation ($/Mg-km or $/Mg)	Total ($/Mg)	
Bale	Suh and Suh	2010	Corn stover	bale-transport-grind	-	-	12-42	Refinery Capacity: 0-2500 million liters/year
				bale-grind-pellet-transport(truck)-grind	-	-	18-33	
				bale-grind-pellet-transport(rail)-grind	-	-	27-30	
Silage	Turhollow et al.	1996	Herbaceous crops	chop-truck in field-transport	7.2	8.37-13.98 $/Mg	15.57-21.28	
				chop with wagon-transport	7.3	8.37-13.98 $/Mg	15.58-21.29	
	Kumar and Sokhansanj	2007	Switchgrass	chop-ensile-transport	27.18	-	38-59	Transporting distance: 3-77 km

(a): converted from euro to U.S. dollar by using the exchange rate of 1.3 for U.S. dollar per 1 euro

The logistics system for cotton stalks investigated by Tatsiopoulos and Tolis (2003) comprises chopping, transporting to an intermediate warehouse, biomass treatment (i.e., drying only, drying-baling, and drying-pelletizing) and storing in an intermediate warehouse, and transporting to a power plant. They compared costs of several logistics systems considering several technologies of biomass treatments in an intermediate warehouse and the type of the ownerships of transportation vehicles (i.e., third party companies and farmers). In the case of drying-baling, the estimated logistics costs were $58.5/Mg when using third party machines and $33.8/Mg when using farmers' machines. Petrou and Mihiotis (2007) evaluated the commercial price of biomass delivered by several collection methods of cotton stalks: uprooting and baling; cutting and baling; and cutting and transfer in bulk. The first and second methods resulted in similar prices and the third one was worse than others.

Corn stover has been studied as feedstock based on a bale system. Sokhansanj et al. (2006) developed the IBSAL simulation framework and evaluated a bale system for corn stover feedstock. Their model comprises combine, shredding, baling, stacking, truck travel, stacking, and grinding. The estimated cost was $53.57/Mg, including the transportation cost of $13.76/Mg.

Morey et al. (2010) estimated the logistics cost of corn stover based on a bale system in Minnesota. The logistics system investigated consists of shredding, raking, baling, transportation to storage site, tub-grinding, compacting, and transportation to conversion plant by a semi-truck. The total cost for delivering bulk corn stover to conversion plants is estimated at $48.48/Mg.

Suh and Suh (2010) in Minnesota developed a simulation model to compare five logistics options for delivering corn stover based on the combination of a process type and a transportation method (see Table 1). Their model does not include harvesting and collecting

operations. The costs were estimated at $12-56/Mg and show that using pellet operation is better than other options.

Logistics systems to deliver switchgrass also have been evaluated. Kumar and Sokhansanj (2007) employed IBSAL to evaluate several logistics system for switchgrass as described in Table 1. The most economic logistics system was a loafing system, the cost of which was $30-45/Mg over several biorefinery capacities. Loafing is a form of large biomass package formed directly from a windrow of material. The loafing system was estimated to deliver biomass at lower cost than baling ($40-55/Mg).

In 1996, Turhollow et al. investigated the cost of a silage system for herbaceous crops. They estimated in-field cost at about $3-12/Mg and the transportation cost at about $8-14/Mg, assuming that the silage density on the truck was 249.89 or 416.48 kg/m^3 (15.6 or 26.0 lb/ft^3). Kumar and Sokhansanj (2007) also evaluated a silage system for switchgrass. The estimated total cost was $38-59/Mg including processing cost of $27.18/Mg.

As these studies have shown, the cost to deliver herbaceous biomass varies by feedstock type, location and technologies used for packaging the biomass. The cost estimates were determined using economic modeling and simulation. The USDA and DOE (2008) reported that biomass logistics, which includes harvesting/collecting, storing and pre-processing, constitutes as much as 20% of the current cost of supplying cellulosic ETOH. Reducing the logistics cost will be a key factor in successful commercialization of the cellulosic bioenergy industry.

5.2 Description of the conceptual biomass module system

The biomass module system investigated consists of several unique machines. Subsection 1 describes the overall system. Subsections 2, 3 and 4 give details about the individual machines required in the system. The module system was not intended for any specific biomass crop or growing region. The fundamental features of the system (high density

and rapid handling) are applicable to all biomass logistics systems. However, one feature of the system (gas tight packaging) was intended to allow the handling and storage of high moisture biomass. This feature is especially needed for high yielding biomass crops that are difficult to field dry to safe storage moisture (< 20%) and high rainfall regions.

Other machines were required to make up the entire logistics system, but are not described here because they are commercial machines implemented as IBSAL simulation elements. Those additional machines included a mower/conditioner and a tractor/semi-trailer combination.

5.2.1 Overall system

The conceptual module-based logistics system comprised a series of unit operations, including cutting and conditioning for field drying, chopping, forming modules, moving modules from a local field location and loading them onto trailers, highway transport to a bio-refinery, and unloading/transport within the bio-refinery. Figure 21 depicts these unit operations in the biomass module system. The shaded blocks represent machines that require development (module former and hauler) or modification (forage harvester).

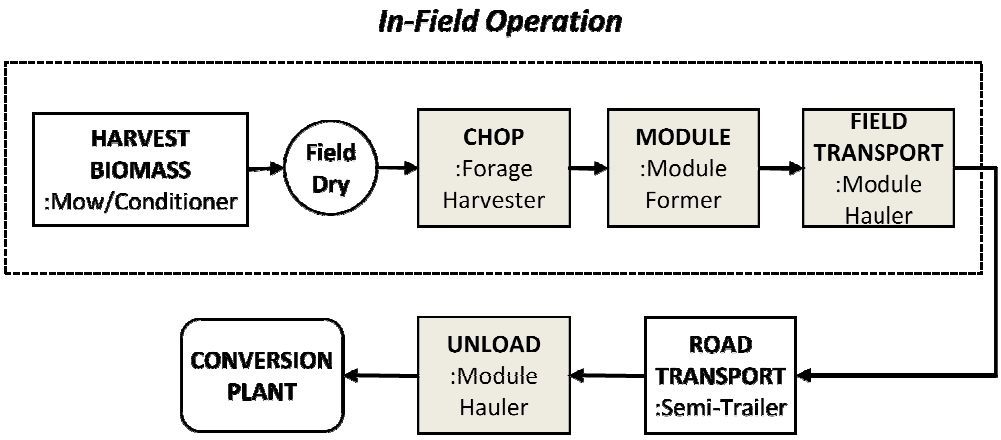

Figure 21 Schematic of the biomass module system: shaded blocks represent unit operations that are not currently available in the IBSAL simulation tool, and were developed in this study

Biomass chopped by a self-propelled forage harvester (SPFH) is collected in the module former, which is towed by the forage harvester. The module former compresses the collected biomass and generates biomass modules, which are enclosed in a plastic oxygen barrier. The biomass module produced is a large package that can be transported effectively by the module hauler and semi-trailer. A module hauler transports modules for a relatively short distance (expected to be less than one mile) from the field to the local field storage or loading site. For long distance transport to a refinery, a semi-trailer is used. Within the refinery, module haulers would be used again to unload and transport modules.

5.2.2 Modified forage harvester

The SPFH will pull a module former, so it will require more power than a conventional forage harvester of the same capacity. Commercially available forage harvesters are designed to pull a silage wagon, but the mass of the former and the on-board biomass will cause the draft to exceed significantly that of the silage wagon. Therefore, changes in the operational parameters of the SPFH when used in the biomass module system were required. These included the additional power to handle the module former and operational factors (e.g., speed and setup time) associated with efficiency.

In a conventional forage chopping system, some portion of chopped biomass can be lost when the stream of chopped material is not retained in the haul trailers (overflow on filled trailers, poor stream direction by operator, etc.). The modified forage harvester was expected to reduce such biomass losses incurred by directly coupling the material transfer between the forage harvester and the towed module former.

5.2.3 Module former

Since no such machine exists currently, the module forming operations were anticipated. While the forage harvester is operating, chopped biomass is blown into the towed module

former. The module former continuously compresses biomass into a constrained package. When the quantity of biomass reaches its predefined amount of 13.6 Mg wet weight (15 tons), the package would be closed and the next one started. The formed module length would vary with the moisture content of the biomass, but maximum length would be 7.3 m (24 ft.). The module former has the capacity to haul up to one and a half full length modules. The finished modules would be unloaded when the machine reaches the edge or the field or a turn-row. The module will wait at the field edge until transported by a module hauler. While unloading the formed module, the forage harvester is required to stop its operation so that unloading operation would decrease the system field efficiency.

Since a module former is not commercially available, several assumptions were employed to define its properties: power – 298.4 kW (400 hp), maximum volume - 47.6 m^3 (1,680 ft^3), module dry matter density target – 240.3 kg/m^3 (15 lb/ft^3); $450,000 purchase cost and plastic cost of $50/module. Module former properties (e.g., purchase cost, power, loading preparation time, unloading time, and module density) may be strongly related to its performance. A range of these values was investigated through a sensitivity analysis (see section 5.4.2).

5.2.4 Module hauler

The module hauler was patterned after a similar machine used in Australian to load cotton modules (which are 25% longer than US standard modules) onto flat-bed trailers (Figure 22). That machine can load and unload the modules quickly (2-3 minutes), and transports the cotton modules from the fields to the loading site. This machine straddles the trailers and uses its tilting, live bed to unload the module onto the trailer. To work with biomass modules, this concept must be modified to handle the greater mass and to lift two biomass modules at one time. Unloading is a reverse operation of loading.

Since a module hauler was also commercially unavailable, judgments were made to define its properties; power – 335.7 kW (450 hp), ability to load two maximum size modules; purchase cost - $375,000; and average travel speed - 11.3 km/hr (7 mph). These properties were also investigated with a sensitivity analysis (see section 2.4.2).

Figure 22 Australian module hauler loading a cotton module from the field

5.3. IBSAL modeling

This section presents details associated with the IBSAL simulation models developed. The IBSAL modeling tool is described, followed by descriptions of new IBSAL elements developed to represent the conceptual machines of the biomass module system, properties of those machines considered in a sensitivity analysis, and the model of the entire module system.

5.3.1 Overview of IBSAL

IBSAL is a collection of simulation elements programmed in ExtendSim, a simulation package (Imagine That, Inc., 2010). This study used a version of IBSAL obtained in 2010 and compatible with ExtendSim versions 7.0 and 8.0. The authors have collaborated with ORNL

personnel to verify and improve IBSAL, and have access to a version of IBSAL not yet released to the public. Some elements have been updated (e.g., moisture content variation logic, dry matter loss logic and etc.) from the version used in previous studies (i.e., Sokhansanj et al., 2006; Kumar and Sokhansanj, 2007; Sokhansanj et al., 2008).

Inputs are spatial information such as farm size, yield, and transportation distances, harvesting schedule in a weekly basis, daily weather data, and machine data. Outputs are cost, energy input, carbon emission, quantities of biomass lost and delivered, and operation time.

In IBSAL, simulation items are generated based on the harvesting schedule (i.e., discrete event simulation). Each simulation item represents a quantity of biomass calculated based on a unit land size and specified crop yield. Thus, the amount of biomass in each simulation item is set by the user. For example, if the input values of a unit land size and a crop yield are 50 ha and 5 Mg/ha respectively, each simulation item represents 250 Mg of biomass. The generated simulation item passes through all simulation elements in the model, in which several attributes (e.g., moisture content) are updated and some resources (e.g., cost and energy) used to process the biomass are calculated.

The simulation model checks the input weather condition to determine how conditions affect processes (change in moisture content) or are suitable for machine operations (rainfall prevents operation). If machine resources are not sufficient to process the simulation items arrived, the simulation items wait in the queue in each simulation elements until the weather condition become better or a machine resource is available. In particular, some collecting machines (e.g., forage harvester and baler) include logic for a field-drying operation. If the moisture content of a simulation item is greater than the target value (e.g., 20%), the simulation item will wait in the queue up to some predefined maximum waiting time.

IBSAL simulation models calculate several performance measures (e.g., cost, required

energy, a quantity of CO_2 emissions, and the tonnage of processed biomass) associated with supplying biomass to a refinery. The cost includes capital, maintenance, tax, interest, and labor costs. The required energy is calculated based on the power of each machine and its working time. The quantity of CO_2 emissions is assumed to be proportional to the energy used. The mass of biomass is that portion remaining after considering the dry matter loss in each operation.

The IBSAL libraries provide several simulation elements that represent machines or processes. These are categorized as harvesters, tractors, transporters, loaders, processors, and storages. Simulation elements in each group have similar structure, so a new or modified element can utilize much of the existing structure and algorithm. An additional functionality added at Texas A&M was the ability to capture and use historical weather from the NOAA National Climatic Data Center, rather than the TMY2 weather data used by the original IBSAL package.

5.3.2 New IBSAL elements

Three new IBSAL elements were developed; a forage harvester working with a towed module former, the module former, and a module hauler. Each are described individually.

5.3.2.1 Modified forage harvester

In the conceptual logistics system analyzed in this study, operations of a forage harvester and a module former are strongly coupled: i.e., those two machines must work together. In the ORNL version of IBSAL, some elements can have a strong relationship, for example between a tractor and a trailer, each unit operation element in IBSAL deals only with one machine and has no logic to relate with a strongly coupled machine. To consider such strong relationship in detail, the forage harvester element required information from the module former (e.g., maximum load and loading/unloading time) and vice versa (e.g., operating time). These two elements were developed separately, although mechanisms were added to pass the necessary information.

To create the SPFH compatible with the module former, the existing IBSAL element for a commercial forage harvester was modified. This new simulation element was developed to incorporate several additional aspects:

(1) The simulation item was changed from a bulk of biomass based on a unit land area to a module, the size of which was calculated based on the maximum volume and the maximum weight limitations defined by the module former,

(2) operations can proceed only when both a forage harvester and a module former are available, and

(3) the total processing time was the sum of the setup time of a machine, the operating time of chopping, and the unloading time of modules.

To specify the property of the modified SPFH, a user can set several parameters, for example, power, purchase cost, dry matter loss, efficiency, speed, daily working hours, and the maximum moisture content for operation to begin. In particular, several parameters (i.e., power, purchase cost, dry matter loss and efficiency) may be affected by the required modification of a commercially available SPFH so that this paper considered those in the sensitivity analysis (see Section 5.4.2).

5.3.2.2 Module former

The simulation element of a module former was built from other transporter elements in IBSAL. The module former element uses the same processing time as the forage harvester. It contains logic for calculating system performance measures based on the processing time. It also contains logic that utilizes the moisture content attribute of the biomass in each simulation unit to determine the dimensions and mass of the module formed and the frequency of unloading finished modules.

A module former in a simulation model can be defined by setting several parameter

values of its simulation element such as purchase cost, power, efficiency, preparation time, unloading time, module density, the maximum volume of a module. This paper examined several values of purchase cost, power, efficiency, preparation time, unloading time and module density through a sensitivity analysis (see Section 5.4.2).

5.3.2.3 Module hauler

The module hauler element was developed based on the IBSAL element of a commercially available truck. The modifications in logic are that it can deal with two modules at once, and that it can load and unload two modules in one operation by itself without using other loading/unloading machines.

Similar to existing transporter IBSAL elements, several parameters (e.g., purchase cost, power, speed, loading/unloading time, the number of operators and machines, min/max transportation distances, and weather conditions) are used to define a module hauler that is used in a simulation model. The sensitivity analysis considered several parameters including purchase cost, power, speed, and loading/unloading time (see Section 5.4.2).

5.4 Simulation model

5.4.1 Simulation scenario

The simulation consisted of a conceptual series of sequential operations that generated the simulation units and proceeded through delivery to a conversion plant. Grass-type biomass was available from some farm land area. Crops were cut by mower/conditioners based on a predefined harvest-schedule. The cut biomass was allowed to field dry until the required moisture content was acquired or the maximum waiting time reached. A forage harvester chopped the biomass and a module former generates biomass modules. The generated modules were laid down on the field and were transported by a module hauler to a local field-storage site. Modules stored in a local field-storage site were transported to a refinery by semi-trailer. The

semi-trailer transported two modules at once. After arriving to a refinery, a second module hauler was used to unload and move the module within a refinery. No additional preprocessing operations were included in the simulation. Figure 23 shows the simulation model developed with existing and modified IBSAL elements.

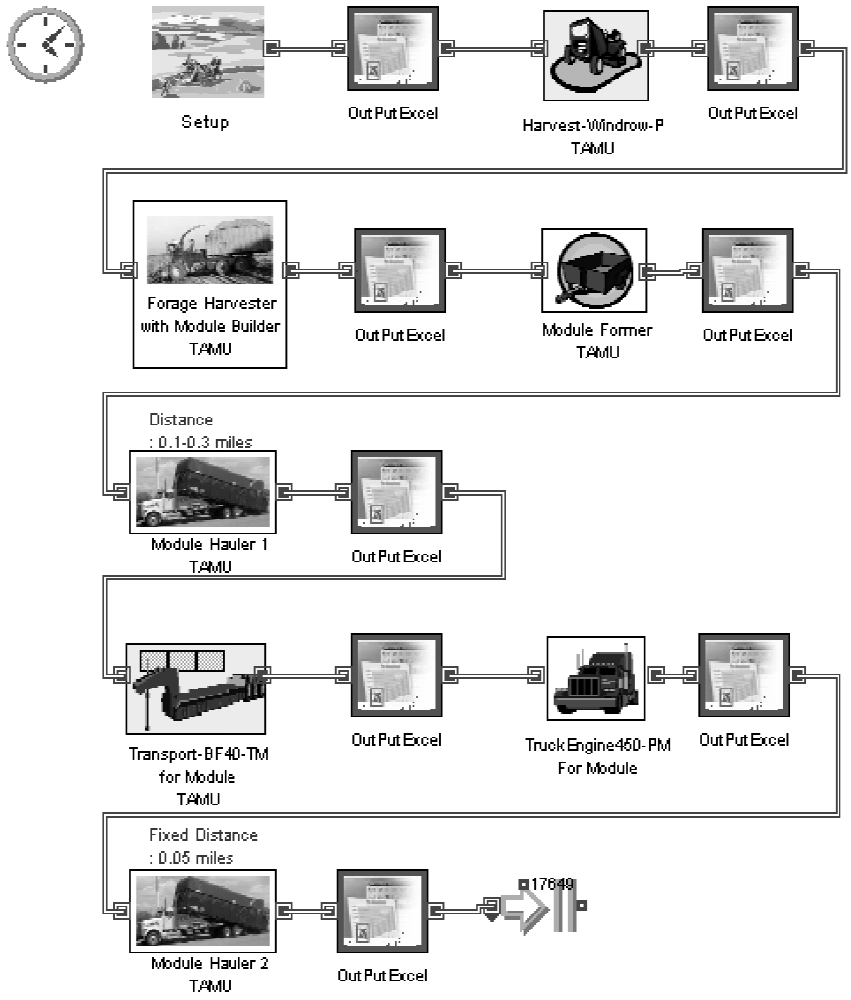

Figure 23 Module-based biomass simulation model: the OutPutExcel elements record the process data into a spreadsheet.

Daily weather data for 1992 in College Station, TX was obtained from National Climatic Data Center provided by the U.S. National Oceanic and Atmospheric Administration. The data

for 1992 was selected because the total precipitation in 1992 of 606 mm is about the median of the total precipitations in College Station, which range from 250 mm in 2010 to 1141 mm in 1994. The dry matter yield of biomass was evaluated at 11.2 or 22.4 Mg/ha (5 or 10 ton/ac). The farm area was 20,250 ha (50,000 ac) for the yield of 11.2 Mg/ha (5 ton/ac), and 10,125 ha (25,000 ac) for 22.4 Mg/ha (10 ton/ac). Thus, the annual biomass supply was the same to both yield cases (i.e., 250,000 Mg).

5.4.2 Sensitivity analysis

Since the required machines in the module-based logistics system are not available commercially, several values critical to the estimating cost and capacity are not known and must be considered for their ultimate impact on the design of such machines. Several factors related to the system performance were identified and tested with a range of values through a sensitivity analysis. For each property, a median value (selected as a most likely value based on the experience of the authors), an optimistic and a pessimistic value were selected. Table 18 gives the properties and values considered in the sensitivity analysis.

While a commercially available forage harvester is considered to have 10% dry matter loss (Sokhansanj et al., 2008), the dry matter loss was judged to be improved as a result of the close coupling of the forage harvester and the module former (6% for median or 2% for optimistic). The impact on field efficiency for the towed SPFH/module former combination compared to a conventional forage harvester (85%) was considered to be zero in the optimistic scenario, 10% in the median case and a reduction of 25% in the pessimistic situation. While the basic forage harvester modeled in IBSAL has a power of 376.7 kW (505 hp), additional power of 37.3, 59.7, and 74.6 kW (50, 80, and 100 hp) were considered. Since the modified forage harvester may require additional cost, three additional costs (i.e., +30,000, +60,000, and

+90,000) were considered based on the purchase cost of the basic forage harvester (i.e., $161,200).

Table 18 Machine properties and values used to evaluate the biomass module system

No	Machine	Property	Values	Unit
1	Forage Harvester	Purchase cost	191,200, 221,200, 251,200	$
2		Dry matter loss	0.02, 0.06, 0.10	Decimal fraction
3		Efficiency	0.60, 0.75, 0.85	Decimal fraction
4		Power	376.7 + 37.3, 59.7, 74.6 (505 + 50, 80, 100)	kW (hp)
5	Module Former	Purchase cost	350,000, 450,000, 550,000	$
6		Power	223.8, 298.4, 373.0 (300, 400, 500)	kW (hp)
7		Preparation time	10, 30, 60	sec
8		Unloading time	0.5, 2, 4	minute
9		Module density	192.2, 240.3, 288.4 (12, 15, 18)	kg/m^3 ($lb/ft^{3)}$
10	Module Hauler	Purchase cost	250,000, 375,000, 500,000	$
11		Power	261.1, 335.7, 410.3 (350, 450, 550)	kW (hp)
12		Average travel speed	4.8, 11.3, 16.1 (3, 7, 10)	km/hour (miles/hour)
13		Loading time	1, 2, 3	minute
14		Unloading time	1, 2, 3	minute

Purchase cost of a module former ($350,000 - 550,000) and the required power (223.8 - 373.0 kW {300 - 500 hp}) were estimated based on similarity to the function and complexity of the John Deere round module cotton picker. Since 20 to 30 modules may be generated from a unit of the plastic cover material, and time is required to refill the consumable plastic periodically, the required preparation time was judged to range from 10 to 60 seconds/a module. The time required to unload a formed module was assumed to be from 0.5 to 4 minutes. Module density produced was judged to be from 192.2 to 288.4 kg/m^3 (12 to 18 lb/ft^3).

A module hauler would be more costly than the machines used to handle and load cotton modules in Australia because of the greater weight being handled. Assumptions representing a module hauler included the following: the purchase cost would be from $250,000 to $500,000; the power was from 261.1 to 410.3kW (350 to 550 hp); the average travel speed was from 4.83 to 16.09 km/hr (3 to 10 miles/hr); and the loading/unloading times were from1 to 3 minutes.

5.5 Results

5.5.1 Sensitivity analysis procedure

The simulation included sufficient numbers of each machine to harvest, collect and deliver all biomass within a harvesting time period (i.e., from August 1 to December 31) so that no biomass remains in the system after finishing the simulation run. The machine numbers for each type in Table 19 were determined manually with multiple simulation runs to select the appropriate number.

Three scenarios were considered to investigate the relationship between performance factors, transportation distance and crop yield; 40.2 km (25 miles) and 11.2 dry Mg/ha (5 dry ton/ac) for scenario I; 80.47 km (50 miles) and 11.2 dry Mg/ha (5 dry ton/ac) for scenario II; and 40.23 km (25 miles) and 22.4 dry Mg/ha (10 dry ton/ac) for scenario III. The change in the costs, energy consumption and productivity between the median and best or worst property values was examined to determine which factors resulted in the greatest improvement or degradation as the property was changed. For each scenario, 29 cases of data setup were simulated (i.e., the median case for all factors, and 28 cases varying individually each of the fourteen factors to the best and worst values). The total number of test cases was 87. Each case was named by using the combination of scenario number, best/worst, and a performance factor number (coded as in Table 2). For example, the case of ‘I-B-1’ meant that scenario was I and the value of factor 1 (forage harvester dry matter loss) was at the best value. Best value was interpreted as giving the

lowest cost/energy consumption. Median cases were named as 'I/II/III-M-0'.

Table 19 The number of machines used in each case of scenario I

Cases	Windrower	Forage Harvester	Module Former	Module Hauler in fields	Trailer	Module Hauler in plant
I-M-0	3	8	8	7	8	2
I-B-1	3	8	8	7	9	2
I-B-2	3	8	8	7	8	2
I-B-3	3	8	8	7	8	2
I-B-4	3	8	8	7	8	2
I-B-5	3	8	8	7	8	2
I-B-6	3	8	8	7	8	2
I-B-7	3	8	8	7	8	2
I-B-8	3	8	8	7	8	2
I-B-9	3	8	8	7	8	2
I-B-10	3	8	8	7	8	2
I-B-11	3	8	8	7	8	2
I-B-12	3	8	8	6	8	2
I-B-13	3	8	8	7	8	2
I-B-14	3	8	8	7	8	2
I-W-1	3	8	8	7	8	2
I-W-2	3	9	9	7	8	2
I-W-3	3	8	8	7	8	2
I-W-4	3	8	8	7	8	2
I-W-5	3	8	8	7	8	2
I-W-6	3	8	8	7	8	2
I-W-7	3	8	8	7	8	2
I-W-8	3	8	8	7	8	2
I-W-9	3	8	8	9	12	2
I-W-10	3	8	8	7	8	2
I-W-11	3	8	8	7	8	2
I-W-12	3	8	8	12	8	2
I-W-13	3	8	8	7	8	2
I-W-14	3	8	8	7	8	2

5.5.2 Simulation results

This section shows the simulation results for the basic scenario, followed by scenarios with varying transportation distance and biomass yield. In all cases, comparisons were made between the simulated results for the scenario with median property values and the changing of

single properties. Because the study focus was to determine the influence of machine operational factors, no attempt was made to vary multiple properties simultaneously. All cost results were reported based in year 2007 US dollars for making our estimates comparable to the goals of the DOE, which were estimated in year 2007 US dollars as well.

5.5.2.1 Basic scenario

Scenario I was considered to be the base scenario. Tables 20 and 21 give the estimates of the performance measures for best and worst cases, respectively.

Table 20 Estimates of the performance measures in scenario I median case and best values

Cases	Cost ($/Mg)			Change of Cost ($/Mg)	Energy (MJ/Mg)	Change of Energy (MJ/Mg)	CO_2 (kg/Mg)	Change of CO_2 (kg/Mg)
	Collection & Processing	Transportation	Total					
I-M-0	19.02	9.60	28.62	-	478.74	-	32.83	-
I-B-1	18.61	9.60	28.21	-0.41	478.74	0.00	32.83	0.00
I-B-2	18.61	9.59	28.20	-0.41	471.72	-7.02	32.35	-0.48
I-B-3	17.71	9.65	27.36	-1.25	455.84	-22.90	31.26	-1.57
I-B-4	18.88	9.60	28.47	-0.14	472.64	-6.10	32.41	-0.42
I-B-5	18.16	9.60	27.76	-0.86	478.74	0.00	32.83	0.00
I-B-6	18.89	9.60	28.49	-0.13	456.33	-22.41	31.29	-1.54
I-B-7	18.84	9.60	28.44	-0.18	475.54	-3.20	32.61	-0.22
I-B-8	18.22	9.61	27.83	-0.79	464.40	-14.34	31.84	-0.98
I-B-9	18.15	8.16	26.32	-2.30[(a)]	440.75	-37.99[(a)]	30.22	-2.60[(a)]
I-B-10	19.02	9.48	28.50	-0.12	478.74	0.00	32.83	0.00
I-B-11	19.02	9.41	28.43	-0.19	470.65	-8.10	32.27	-0.56
I-B-12	19.02	9.40	28.42	-0.20	473.88	-4.86	32.49	-0.33
I-B-13	19.02	9.50	28.52	-0.09	476.48	-2.26	32.67	-0.15
I-B-14	19.02	9.50	28.52	-0.09	476.48	-2.26	32.67	-0.15

(a) the property with the greatest change in that performance factor.

The total cost to capture and deliver biomass was $28.26/Mg for the median case (I-M-0) and ranges from $26.36 to $32.33/Mg. The range of energy consumed was from 440.75 to

540.10 MJ/Mg with a median case of 478.74 MJ/Mg. The median case of CO_2 emission was 32.83 kg/Mg and varies from 30.22 to 37.83 kg/Mg. The numbers of each machine were selected to process all available biomass during the simulation time, so biomass tonnages were the same in all cases except the cases, I-B-2 and I-W-2, which considered the change in dry matter loss. Thus, estimates of the biomass tonnage are not reported.

Table 21 Estimates of the performance measures in scenario I median case and worst values

Cases	Cost ($/Mg)			Change of Cost ($/Mg)	Energy (MJ/Mg)	Change of Energy (MJ/Mg)	CO_2 (kg/Mg)	Change of CO_2 (kg/Mg)
	Collection & Processing	Transportation	Total					
I-M-0	19.02	9.60	28.62	-	478.74	-	32.83	-
I-W-1	19.43	9.60	29.03	0.41	478.74	0.00	32.83	0.00
I-W-2	19.50	9.66	29.17	0.55	488.01	9.27	33.46	0.64
I-W-3	21.84	9.55	31.39	2.78	529.28	50.54	36.29	3.47
I-W-4	19.16	9.60	28.76	0.14	484.84	6.10	33.24	0.42
I-W-5	19.88	9.60	29.48	0.86	478.74	0.00	32.83	0.00
I-W-6	19.15	9.60	28.75	0.13	501.15	22.41	34.36	1.54
I-W-7	19.29	9.59	28.88	0.27	483.57	4.83	33.16	0.33
I-W-8	20.09	9.58	29.68	1.06	498.03	19.28	34.15	1.32
I-W-9	20.42	11.91	32.33	3.72[(a)]	540.10	61.35[(a)]	37.03	4.21[(a)]
I-W-10	19.02	9.71	28.73	0.12	478.74	0.00	32.83	0.00
I-W-11	19.02	9.78	28.80	0.19	486.84	8.10	33.38	0.56
I-W-12	19.02	10.49	29.51	0.89	500.36	21.62	34.31	1.48
I-W-13	19.02	9.69	28.71	0.09	481.00	2.26	32.98	0.15
I-W-14	19.02	9.69	28.71	0.09	481.00	2.26	32.98	0.15

(a) the property with the greatest change in that performance factor.

In general, all fourteen factors were consistent in having similarly ranked changes in both the best and worst scenarios. Increasing the module DM density achieved by the module former (factor 9) from 240.3 to 288.3 kg/m^3 (15 to 18 lb.ft^3) had the largest impact on cost, energy, and CO_2 emissions for both the best and worst comparisons. The module density

affected the transportation, collection and processing costs. The efficiency of the forage harvester (factor 3) was the second most important factor in both best and worst cases. In the best case, reducing the time to unload a module from the module former (2.0 to 0.5 min) gave the third best improvement in cost, but for the worst case, reducing travel speed from 11.3 to 4.8 km/hr resulted in the third greatest increase in cost. For all three machines, the power requirements did not have a great influence on the cost, but did on the energy consumption and CO_2 emissions.

5.5.2.2 Impact of variations of a transportation distance and a biomass yield

Since the sensitivity of the performance factors could be changed in different biomass supply system environments, variations of transportation distance and biomass yield were considered. Scenario II considered the increase of the transportation distance from 40.2 to 80.5 km (25 to 50 miles). Scenario III dealt with the increase of the biomass yield from 11.2 to 22.4 Mg/ha (5 to 10 Mg/ac). In scenario III, the total biomass amount was kept the same as in scenario I, so the harvested area decreased by 50%.

Table 22 shows estimates of the performance measures in several scenarios. As the transportation distance increased (from scenario I to II), the cost, energy, and CO_2 emission increase. On the other hand, as the biomass yield increases (from scenario I to III), the cost, energy, and CO_2 emission decrease.

In terms of collection and processing cost, the results of scenario I and II were same. This was because an identical amount of biomass was collected, and the increased transportation distance did not affect the collection and processing cost. However, since the influence of field efficiency increases in the higher yield of scenario III, the collection and processing cost decreased significantly. For the transportation cost, the results of scenario II were much larger than the results of scenario I and III. Even though the transportation costs of scenario III are

slightly less than those of the scenario I, those differences are not meaningful, because those differences could occur due to the slight change of the number of generated modules associated with different predicted biomass moisture content. The simulation results show that the increase of transportation distance affects only the transportation cost, and the increase of biomass yield influences the costs associated with collection and processing.

Table 22 Estimates of the performance measures in several scenarios

Scenarios			Scenario I		Scenario II		Scenario III	
			Value	Case	Value	Case	Value	Case
Cost ($/Mg)	**Collection & Processing**	**min**	17.71	I-B-3	17.71	II-B-3	11.52	II-B-9
		max	21.84	I-W-9	21.84	II-W-9	13.79	II-W-9
	Transportation	**min**	8.16	I-B-9	14.28	II-B-9	8.13	II-B-9
		max	11.91	I-W-9	20.84	II-W-9	11.89	II-W-9
	Total	**min**	26.36	I-B-9	32.43	II-B-9	19.65	II-B-9
		max	32.33	I-W-9	41.26	II-W-9	25.69	II-W-9
		median case	28.62	I-M-0	35.81	II-M-0	21.97	II-M-0
Energy (MJ/Mg)	**Total**	**min**	440.75	I-B-9	586.63	II-B-9	327.61	II-B-9
		max	540.10	I-W-9	753.04	II-W-9	427.25	II-W-9
		median case	478.74	I-M-0	650.27	II-M-0	365.80	II-M-0
CO2 (kg/Mg)	**Total**	**min**	30.22	I-B-9	40.22	II-B-9	22.46	II-B-9
		max	37.03	I-W-9	51.63	II-W-9	29.30	II-W-9
		median case	32.83	I-M-0	44.59	II-M-0	25.08	II-M-0

Figure 24 and 25 depict the impact of the transportation distance and the biomass yield on the sensitivity of the performance factors such as cost and energy, respectively. Each bar represents the change of the estimate of the performance measure from the median case to the best (or worst) case of each of all factors (1 – 14). So, the length of the bar represents the measure of the significance of each factor to the system performance. Since the CO_2 emission is calculated based on the used energy, that chart is not included.

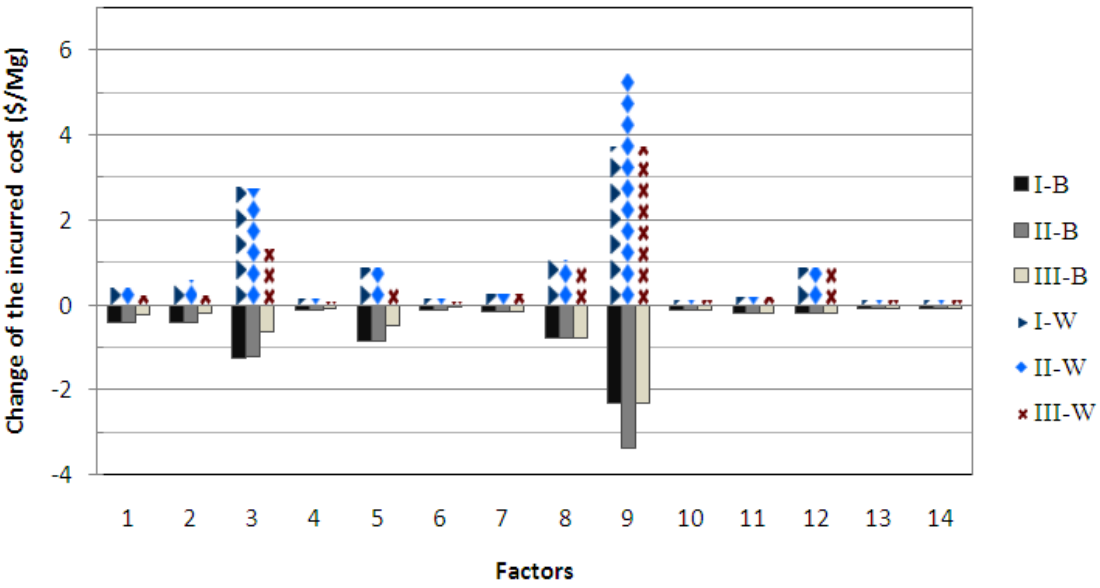

Figure 24 Comparison of the impact of the variation of the transportation distance and the biomass yield to the sensitivity of the performance factor: cost

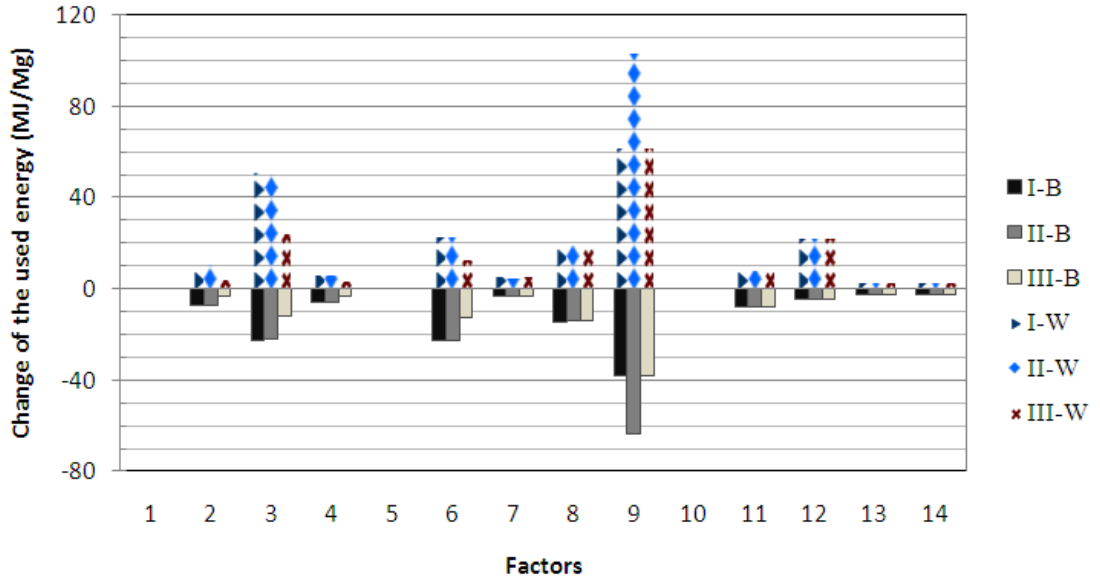

Figure 25 Comparison of the impact of the variation of the transportation distance and the biomass yield to the sensitivity of the performance factor: energy

When the transportation distance increases (i.e., from I-B/W to II-B/W), only the significance of a module density (factor 9) increases. This result implies that module density is the critical factor to improve the system performance.

For the cases under scenario III in which a biomass yield increases, the significance of several factors decrease. While the changes of the significance of factors 10, 11, and 12 related to a module hauler were very small, the significance of factors associated with a forage harvester (1, 2, 3, and 4) and a module former (5, 6, 7, 8 and 9) decrease considerably. This is because the increase in biomass yield may affect operations associated with harvesting and collecting in a field, and the transportation distance was equal in scenarios I and III.

For the most significant factors, the reduction in cost and energy between the median and best values was of lesser absolute magnitude than resulted from including the worst values. Those factors having small impact on the cost and energy consumption had nearly equal change when varied between the best and worst values. These observations indicate that the selected median values were reasonable for the operation of the conceptual system.

5.6 Discussion

The analysis of the overall system performance as a function of the considered factors provides important guidance in the establishing of design specifications for the three machines critical to the conceptual system considered. The dry matter density of the biomass modules was shown to be the factor most affecting cost of delivering the biomass. The worst case value for density (192 kg/m^3) is roughly equivalent to the densities achieved in commercial balers. For the biomass module system to achieve lower costs, the density will need to be near to 240 kg/m^3.

The cost to supply a unit Mg of herbaceous biomass to a conversion plant ranged from \$19.65 to \$41.26/Mg with a median cost of \$28.62/Mg. This cost range is very competitive with other biomass logistics systems estimated in the literature. This result indicated that development

of the biomass module system would be justified, and provided design specifications to be met. Such a feedstock supply system could overcome some barriers for successful commercialization of the cellulosic biofuel industry.

In general, field efficiency impacts significantly on accomplishing an economical agricultural system. Since a forage harvester and a module former were strongly coupled in the module system, the turning, preparation time and unloading would result in lower field efficiency. Moreover, techniques used to form the biomass modules will have a large effect on influence on field efficiency. For example, the length of a plastic on a roll that must be periodically loaded will affect the preparation time, and field size, shape and biomass moisture content will affect the number of modules produced and the locations where they must be unloaded. The design of the coupled forage harvest and towed module former must incorporate mechanisms that allow for high field efficiencies.

Other factors can provide additional value and must be included in the design considerations. However, the potential gains are more limited, and design compromises can be made on those factors.

The IBSAL simulation tool proved useful in conducting the sensitivity analysis for the operational factors of the conceptual system. The IBSAL framework is modularized very well so that developing new elements did not affect existing elements. Moreover, IBSAL elements in each machine category have similar structures (e.g., harvester and transporter). Therefore, the development of new elements on IBSAL framework was relatively straightforward. However, since a simulation item in IBSAL is defined based on the biomass amount per a unit land area, it was required to split those simulation units into multiple units of a module size to consider the detail operations in a module system. This meant that all subsequent elements in the operational sequence had to be compatible with those smaller simulation units.

CHAPTER VI

SIMULATION MODEL: PART2. EVALUATION OF ALTERNATIVE BIOMASS LOGISTICS SYSTEMS

This chapter evaluates economic benefits of a biomass module system, comparing it to bale and silage systems. Section 1 presents a background of IBSAL studies for biomass logistics systems. Section 2 describes several biomass logistics system including biomass module, silage, and bale systems. Section 3 provides simulation models based on IBSAL. Section 4 gives simulation results and compares several performance measures of several biomass logistics systems.

6.1. Background

Several studies based on IBSAL have reported projected costs of various logistics systems. The first study was Sokhansanj et al. (2006) that estimated a bale system for corn stover by using IBSAL. They developed an IBSAL simulation model that comprises combine, shredding, baling, stacking, truck travel, stacking, and grinding. Their model resulted in $53.57/Mg for the operations in their bale system, which consists of the CP cost at $39.81/Mg and the transportation cost at $13.76/Mg.

In 2007, Kumar and Sokhansanj employed IBSAL to evaluate and compare several biomass logistics systems (i.e., bale, silage, and loaf systems), considering switchgrass as a feedstock. Their study invoked several assumptions that the yield of switchgrass is 11 dry Mg/ha; a given farm land area supplies switchgrass to a refinery which is located at the center point of the land considered; the utilization rate of the land for growing switchgrass is about 10%; and a road winding factor is 1.4. From their evaluation results (see Table 23), a loafing system was the best system in terms of both CP and transportation costs. They estimated

transportation costs based on the maximum transportation distance with the assumption that the minimum transportation distance is 3 km, noting that a transportation distance is proportional to a refinery capacity. Even though the silage system studied by Kumar and Sokhansanj (2007) has relatively small CP costs for several scenarios, transportation costs were higher than other logistics systems. This implies that the silage system may not be an appropriate method for a large-capacity refinery.

Table 23 Predicted logistics costs for switchgrass by IBSAL (Kumar and Sokhansanj, 2007)

System	Process	Cost			Transportation Machine		
		Collection & Processing ($/Mg)	Transportation ($/Mg)+ ($/Mg/km)	Total ($/Mg)+ ($/Mg/km)	Biomass Density (kg/m^3)	Speed (km/h)	Cost ($/h)
Bale	swath-rake-bale(square)-stack-tarp-transport-stack-grind	33.43	3.05+0.111*d[(a)]	36.26+0.111*d	128	24	50.46
	swath-rake-bale(square)-stack-tarp-grind-transport	32.76	1.18+0.173*d	33.94+0.173*d	64-96	24	51.68
	swath-rake-bale(round)-stack-tarp-transport-stack-grind	31.95	3.05+0.111*d	34.78+0.111*d	128	24	50.46
	swath-rake-bale(round)-stack-tarp-grind-transport	31.28	1.18+0.173*d	32.46+0.173*d	64-96	24	51.68
Silage	chop-ensile-transport	27.18	5.66+0.258*d	32.84+0.258*d	64-96	24	51.68
	swath-rake-chop-pile-transport	14.81	5.66+0.258*d	20.47+0.258*d	64-96	24	51.68
Loaf	swath-rake-loaf-grind-transport	22.33	1.18+0.173*d	23.51+0.173*d	64-96	24	51.68

(a) *d* is a maximum transportation distance

However, since the detail processes of the logistics systems in prior studies are not exactly the same and they dealt with different biomass crops as well as different transportation distances, comparing their results directly may be inappropriate.

Based on the previous cost estimates which are reported by previous studies, the currently available logistics systems may not be suitable for delivering herbaceous biomass crops to refineries due to high logistics cost incurred by low density of biomass materials as well as high moisture content. The USDA and the DOE (2008) reported that biomass logistics, which includes harvesting/collecting, storing and pre-processing, constitutes as much as 20% of the current cost of supplying cellulosic ETOH.

6.2. Description of biomass logistics systems

Figure 26 depicts the detail processes of the logistics systems considered in this chapter: i.e., biomass module, silage, and bale systems. The operations involved in the logistics systems are classified into three categories (i.e., in-field operation, road-transportation, and in-plant operation) to reveal distinction between each other more clearly. The in-field operation includes several operations from harvesting to field-transportation for delivering crops to local storage sites.

The road-transportation covers a few operations associated with transportation on roads such as loading and unloading as well as actual road-transportation. Only the bale system involves a process that is classified as the in-plant operation (i.e., grinding). The CP includes operations both in the fields and in a refinery.

Most herbaceous biomass has high moisture content (e.g., 20 – 60%) so that decreasing moisture content before transportation may be needed to improve transportation efficiency. Therefore, it is assumed that field-drying operations after windrowing will be conducted in all systems. Another assumption invoked is that biomass will be stored on the fields without using any facilities so that the storage cost could be negligible. Moreover, the materials chopped by a forage harvester are assumed to be used in a conversion plant without any additional size-reduction process in both the biomass module and the silage systems.

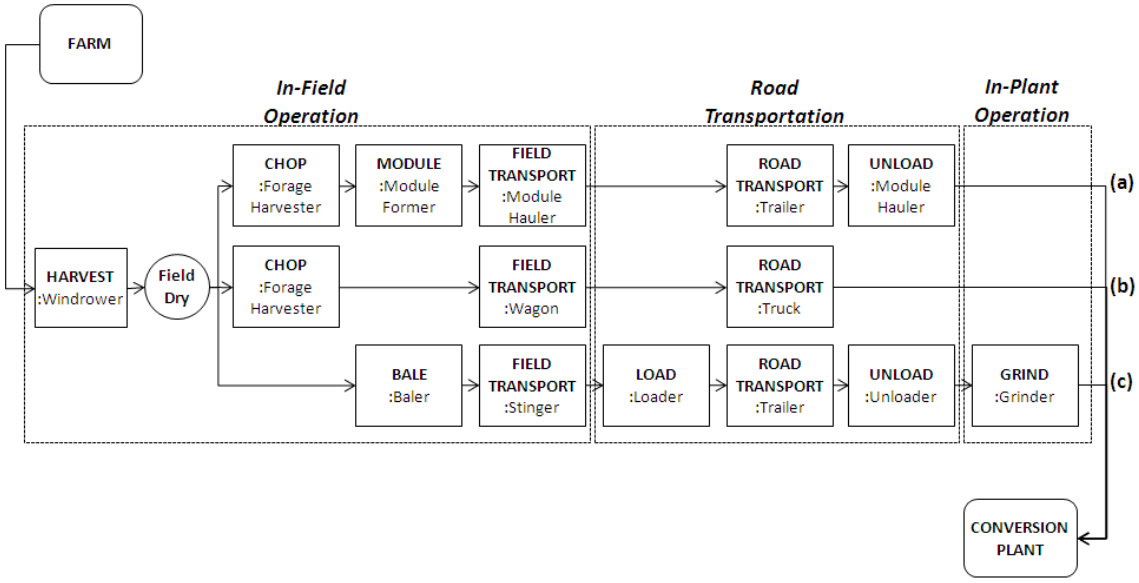

Figure 26 Process diagram for each of biomass logistics systems: (a) biomass module system, (b) silage system, and (c) bale system

After windrowing, crops conditioned are dried on the fields under the atmosphere until the required moisture content (e.g., less than 20%) has been acquired or the maximum managerial waiting time has been reached. Therefore, the time for field-drying operation may vary according to weather conditions. The following subsections describe details of each biomass logistics system.

6.2.1 Biomass module system

The biomass module system is a conceptual system using several conventional machines as well as specially designed new machines, as described in Chapter V. It begins from cutting and drying operations. After a SPFH chops crops, the chopped materials are collected in a module former which is pulled by a SPFH. Strongly coupled operations between a SPFH and a

module former may decrease the loss of biomass while transferring the chopped materials into a module former. However, to represent such a feature, a commercially available SPFH may be needed to be modified because pulling a module former may require more power and energy. Moreover, the field efficiency of the modified SPFH may be decreased due to the strongly coupled operations with a module former. For example, when a module former unloads a biomass module, a modified SPFH must stop its operation.

A module former is a new machine that compresses the collected biomass for increasing density and generates large biomass modules. The expected density will be about 240 km/m^3 (15 lb/ft^3). The size of each module is determined based on the maximum weight load and the maximum volume of a transportation machine. For example, when a biomass module has high moisture content, the module size may be limited by the maximum weight load of a transportation machine so that it could be smaller than the maximum volume of a transportation machine. To facilitate biomass handling and keep biomass from being affected by the atmosphere, biomass modules are covered in a plastic package, which requires additional cost (e.g., $50/module).

Biomass modules are handled and transported efficiently by a module hauler in the fields and by a semi-trailer on the roads. A module hauler is a new machine that is specially designed for handling modules in a field area. In addition, it is designed to load (unload) modules to (from) a semi-trailer quickly. A semi-trailer is used for long-distance transportation from a field storage location to a conversion plant. At a conversion plant, a module hauler is required to unload modules from a semi-trailer and to move them within it.

6.2.2 Silage system

The initial stage of the silage system considered is similar to the biomass module system as shown in Figure 26: i.e., after cutting and conditioning, biomass materials are dried on the

fields. Then, a SPFH chops biomass, collecting them into a wagon moving alongside of the SPFH. In the fields, wagons transport the chopped biomass. Then, for road-transportation, trucks are used. The density of biomass transported is the same as that of biomass after they are chopped by a SPFH, typically 80 kg/m^3 (5 lb/ft^3).

6.2.3 Bale system

Different to the module and the silage systems, the bale system does not contain chopping operation after field-drying. Instead, a baler collects biomass materials on the ground and generates bales, increasing density, in general, up to 128 km/m^3 (8 lb/ft^3). Then, a stinger collects bales in the fields and transport to field storage sites. A semi-trailer is used for road-transportation. A bale loader loads bales to a trailer and unloads them from it. After delivering biomass bales to a conversion plant, bales need to be ground for reducing their sizes enough to be processed in a conversion plant.

6.3. IBSAL modeling

This section explains IBSAL simulation models to evaluate biomass logistics systems. Each of subsections describes a basic scenario and common factors for simulation models, and presents simulation models for the biomass module, the silage, and the bale systems, respectively.

6.3.1 Simulation scenario

This subsection describes the scenario that was used in the simulation models developed. Grass-type biomass is supplied to a conversion plant from a farm land area located at College Station, Texas state of the U.S. Total annual mass supply is 250,000 dry Mg/year. Two cases of average biomass yield have been considered such as 11.2 and 22.4 Mg/ha (5 and 10 dry ton/ac) so that total crop supply area is 20,234 ha (50,000 ac) for 11.2 Mg/ha and 10,117 ha (25,000 ac) for 22.4 Mg/ha.

National Climatic Data Center operated by the U.S. National Oceanic and Atmospheric Administration provides historical weather data for several locations in the U.S. and several other counties. Since the precipitation of 606 mm in 1992 is approximately the median of historical data for the total year-precipitations in College Station, which range from 250 mm in 2010 to 1141 mm in 1994, the daily weather data for the year 1992 were selected to be used in the simulation models.

The in-field transportation distance was assumed to be 0.16-0.48 km. To investigate the impact of transportation distance, several transportation distances between the fields and a conversion plant (i.e., 16, 40, 80, and 161 km) have been considered. Enough number of machines was used to each operation to process all biomass within a harvesting time period (i.e., from August 1 to December 31) so that no biomass remains in the system after finishing the simulation run.

Simulation items are generated based on the harvesting schedule. Each simulation item represents some amount of biomass (e.g., 250 Mg) which is calculated based on a unit land area and a crop yield. A generated simulation item passes through all simulation elements in the model, in which several attributes (e.g., moisture content) are updated and some relevant measures (e.g., cost and energy) are calculated. Each simulation model follows the operation procedure illustrated in Figure 26.

6.3.2 Biomass module system

The simulation model of the biomass module system described in Chapter V was used again (see Figure 23 in section 5.4.1). The model comprises existing IBSAL elements (i.e., windrower and semi-trailer) and several new IBSAL elements (i.e., a modified SPFH, a module former, and a module hauler). After mowing and conditioning by the windrower element, simulation items will stay in the queue of the SPFH element to meet the requirement for the safe

level of moisture content or the maximum waiting time. Then, as long as machines in the subsequent simulation elements are available, simulation items will pass through them, calculating relevant measures (e.g., cost, energy used and dry matter loss). Table 23 provides detail specifications of machines used in the simulation models. Note that the parameter values for the new IBSAL elements have been selected to use the median values provided in Table 18 (see section 5.4.2), which were selected as a most likely value based on the experience of the authors.

6.3.3 Silage system

The simulation model of the silage system consists of series of a unit operation element provided by IBSAL (see Figure 27). All parameter values are provided by the database of IBSAL for commercially available machines. Note that the SPFH used in this model has less power, (i.e., 376.7 kW) and more dry matter loss (i.e., 0.25) compared to the modified SPFH used in the biomass module model (i.e., 436.4 kW and 0.06, respectively), and the density of biomass transported is 80.09 kg/m^3 which is less than that of the biomass module system, 240.28 kg/m^3 (see Table 24).

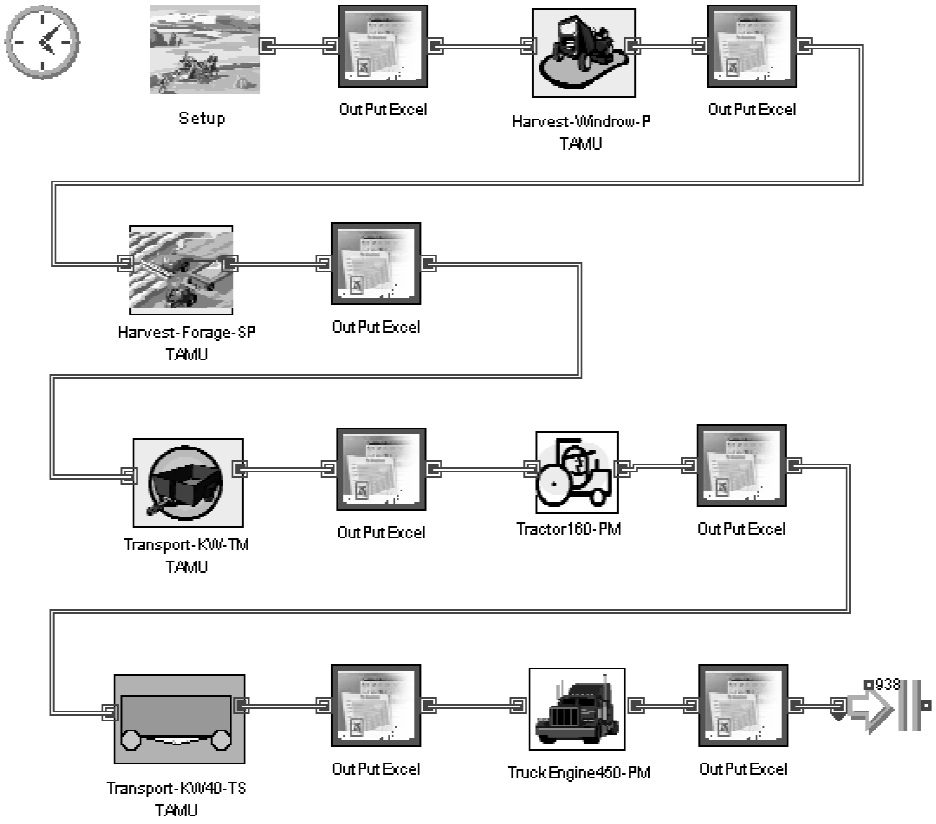

Figure 27 Simulation model for the silage system

Table 24 Specification of machines used in each simulation model

System	No.	Machine	Power (kW)	Width (m)	Speed (km/h)	Efficiency	Cost ($/hr)	Dry Matter Loss	Max Volume (m3)	Max Load (Mg)	Load/Unload Time (min)	Etc
Module	1	Windrower	89.52	4.6	12.9	0.8	106.80	0.05	-	-	-	
	2	SPFH with Module Former	436.4	4.3	9.7	0.75	228.40	0.06	-	-	-	Particle Size : 2.54 cm
	3	Module Former	298.4	-	-	-	150.70	0	47.57	15	0.5/2.0(/module)	Biomass Density : 240.28 (kg/m^3)
	4	Module Hauler1	335.7	-	11.8	-	122.40	0	47.57	15	1.0/1.0(/module)	Transport 2 modules Distance: 0.16-0.48 km
	5	Semi-trailer	410.3	-	80.5	-	23.93	0	95.14	30	1.0/1.0(/module)	Transport 2 modules
	6	Module Hauler2	335.7	-	11.3	-	122.40	0	47.57	15	1.0/1.0(/module)	Transport 2 modules Distance: 0.48 km
Silage	1	Windrower	89.5	4.6	12.9	0.80	106.80	0.05	-	-	-	
	2	SPFH	376.7	4.3	9.7	0.85	238.60	0.25	-	-	-	Particle Size : 2.54 cm
	3	Wagon	119.4	-	16.1	-	21.86	0.005	33.98	25	3.0/5.0	Biomass Density : 80.09 (kg/m^3) Distance: 0.16-0.48 km
	4	Truck	410.3	-	88.5	-	29.43	0.005	72.49	40	5.0/10.0	
Bale	1	Windrower	89.5	4.6	12.9	0.80	106.80	0.05	-	-	-	
	2	Baler	261.1	3.1	12.9	0.65	60.07	0.25	-	-	-	Biomass Density : 128.14 (kg/m^3) Bale Size: : 0.9*1.2*2.4 (2.72 m^3)
	3	Stinger	261.1	-	24.1	-	85.85	0.005	-	20	0.1/0.05(/bale)	Transport 8 bales Distance: 0.16-0.48 km
	4	Loader1	223.8	-	-	-	131.00	0.005	-	-	0.25/0.2(/bale)	
	5	Semi-trailer	410.3	-	80.5	-	23.93	0.005	-	30	0.25/0.2(/bale)	Transport 34 bales
	6	Loader2	223.8	-	-	-	131.00	0.005	-	-	0.25/0.2(/bale)	
	7	Grinder	288.7	-	-	0.80	275.90	0.1	-	-	-	Particle Size : 1.27 cm

6.3.4 Bale system

Figure 28 describes the simulation model of the bale system. The initial stage of the bale system model up to the windrower element is same to other simulation models. Then, the density of simulation items is set to be 128.14 kg/m^3 (8 lb/ft^3) through the baler element. The size of a rectangular bale that the model considered is 0.9 m*1.2 m*2.4 m (2.72 m^3). The stinger element estimates measures associated in-field transportation of bales. The bale loader elements are used before and after the semi-trailer element, which can transport 34 bales at once. Only the bale simulation model has an in-plant operation (i.e., grinding) for reducing the size of biomass materials. Similar to the silage simulation model, the bale simulation model comprises simulation elements existing in the current version of IBSAL.

6.4. Simulation results

This section describes the performance measures, which are estimated by simulation models, and the number of machines assigned for each scenario, and presents their estimates for several scenarios.

6.4.1. Performance measures

Simulation models estimate several performance measures (e.g., cost, required energy, a quantity of CO_2 emissions, and the net biomass yield) to supply biomass from farms to a conversion plant. The cost includes capital, maintenance, tax, interest, and labor costs. The required energy is calculated based on the power of each machine and the working time. The quantity of CO_2 emissions is calculated based upon that of the energy used. The net biomass yield is the amount of biomass after considering physical dry matter loss occurred through all operations in each system.

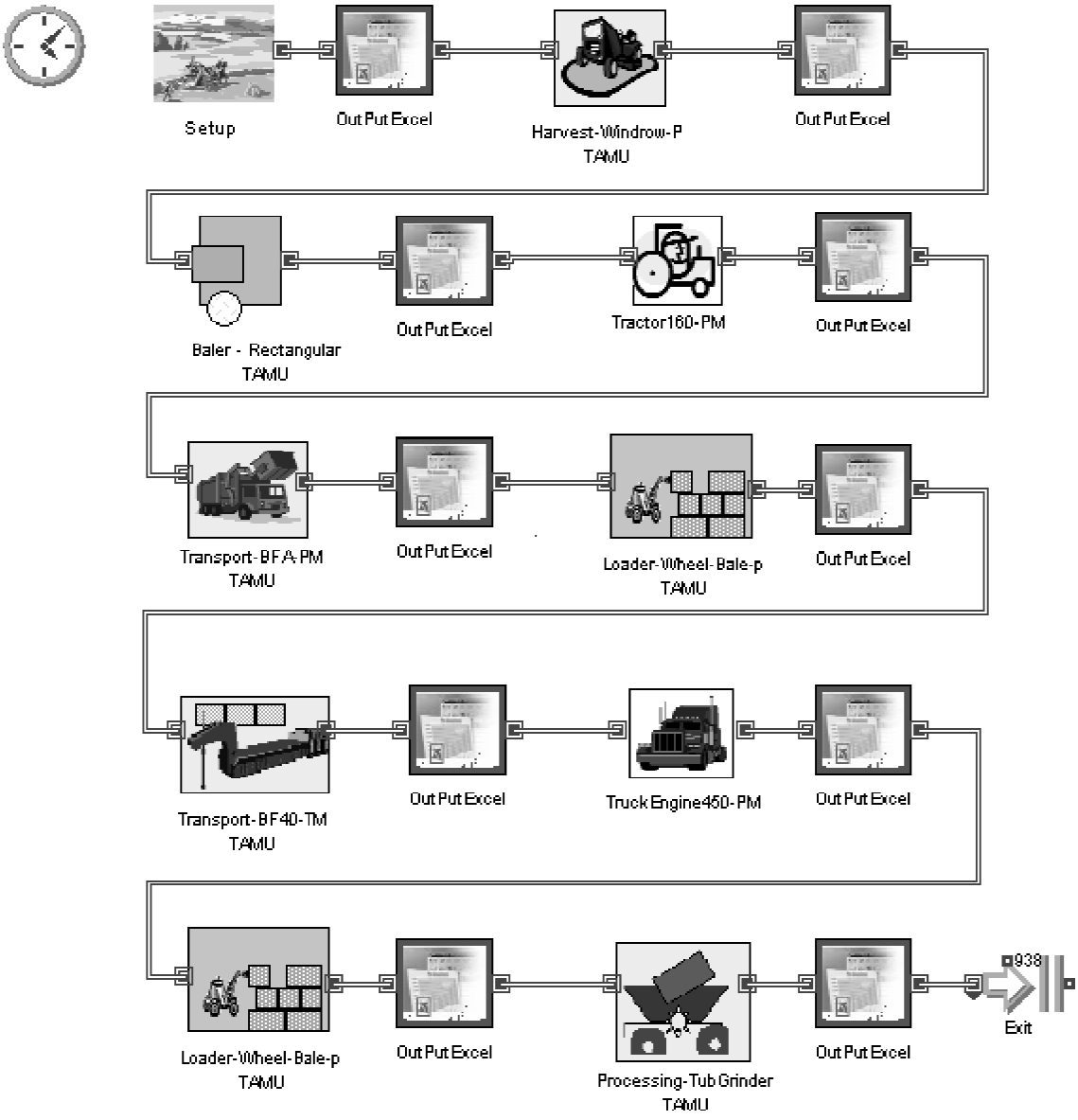

Figure 28 Simulation model for the bale system

6.4.2. The number of machines

Each simulation element used enough number of machines to process all biomass within a simulation time horizon (i.e., harvesting period) so that simulation models did not hold any simulation items after the simulation run finished. To decide the appropriate number of machines to be assigned, preliminary simulation runs were conducted. Table 25 gives the number of machines assigned to each simulation element under several scenarios. The transportation distance affects the required number of machines associated with transportation (e.g., a semi-trailer and a truck).

Table 25 The number of machines used in each machine under several scenarios

System	No.	Machine	Transportation Distance (km)							
			16		40		80		161	
			Y11[(a)]	Y22[(b)]	Y11	Y22	Y11	Y22	Y11	Y22
Module	1	Windrower	3	2	3	2	3	2	3	2
	2	SPFH with Module Former	8	4	8	4	8	4	8	4
	3	Module Former	8	4	8	4	8	4	8	4
	4	Module Hauler1	7	7	7	7	7	7	7	7
	5	Semi-trailer	5	5	9	9	15	15	29	29
	6	Module Hauler2	2	2	2	2	2	2	2	2
Silage	1	Windrower	3	2	3	2	3	2	3	2
	2	SPFH	4	2	4	2	4	2	4	2
	3	Wagon	14	14	14	14	14	14	14	14
	4	Truck	30	30	60	61	110	110	209	209
Bale	1	Windrower	3	2	3	2	3	2	3	2
	2	Baler	8	4	8	4	8	4	8	4
	3	Stinger	3	3	3	3	3	3	3	3
	4	Loader	2	2	2	2	2	2	2	2
	5	Semi-trailer	4	4	9	9	18	18	33	34
	6	Loader	2	2	2	2	2	2	2	2
	7	Grinder	8	8	8	8	8	8	8	8

(a): Crop yield of 11.2 Mg/ha (5 ton/ac)
(b): Crop yield of 22.4 Mg/ha (10 ton/ac)

As transportation distance increases, the required number of transportation machines

increased as well, while other machines were not affected. Similarly, several machines associated with in-field operations were affected by the crop-yield variation (e.g., a windrower, a SPFH and a module former).

6.4.3. Estimates of performance measures

This section provides the simulation result of each biomass logistics system. Comparison with the results of several scenarios describes the impact of variations of a transportation distance and a biomass yield on the performance of each biomass logistics system.

6.4.3.1 Total cost

Figure 29 depicts estimates of the logistics costs with respect to transportation distances (i.e., 16, 40, 80, and 161 km) as well as crop yields (11.2 and 22.4 dry Mg/ha). For a short transportation distance less than about 5 km, the silage system has lower cost than other systems. In contrast, the biomass module and the bale systems are better in all scenarios, in which transportation distance is greater than 5km, because the silage system appeared to be more sensitive (i.e., larger slope) to transportation distance than others. The biomass module system has the lowest logistics cost compared to other systems. Moreover, the sensitivity of the biomass module system to the transportation distance is lowest among all systems.

As crop yield increases from 11.2 to 22.4 dry Mg/ha, the total cost decreases in all systems. This is primarily due to the decrease of the CP costs (see Table 4). In particular, the change of the crop yield affects significantly the biomass module system: i.e., the reduction of the CP cost of the biomass module system ranges from 34 to 39% with respect to the crop-yield change from 11.2 to 22.4 dry Mg/ha; that of the silage system, from 13 to 27%; and that of the bale system, from 14 to 27%, respectively. The bale system showed the least impact on the CP cost from the crop-yield change because the grinding operation, which is a major component for the cost of the bale system, is not affected by a crop yield.

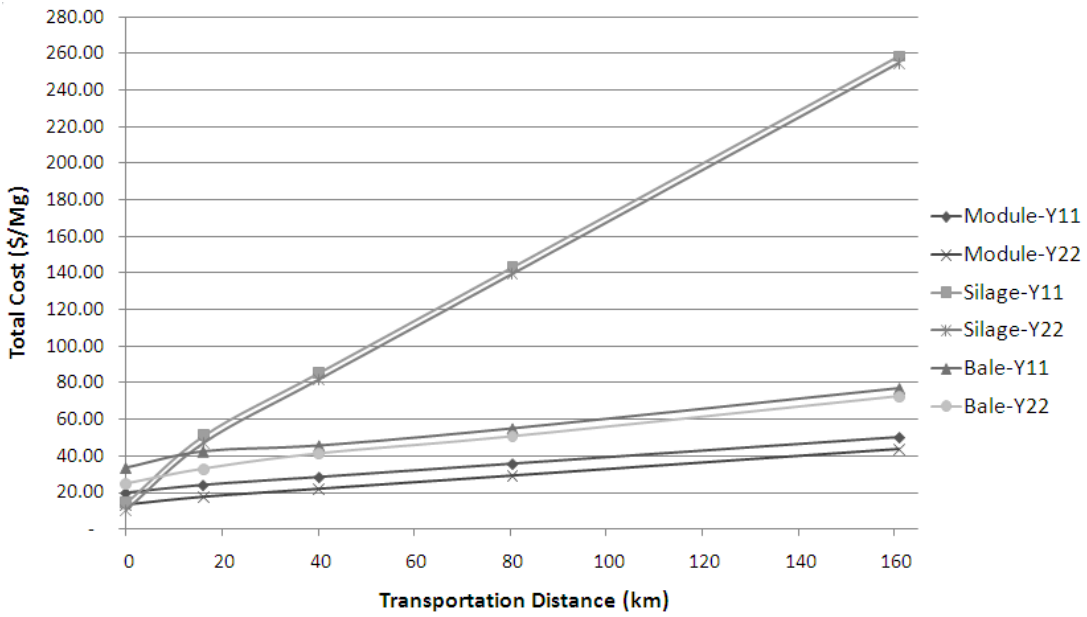

Figure 29 Total cost of several logistics systems under several scenarios

Based on the target that DOE set in the biomass multi-year program plan in 2010, the economic feasibility of the logistics systems can be speculated. For example, if the logistics cost per a unit biomass is under $38.59/Mg, which is the DOE goal by 2012, the biomass logistics system can be considered to be economically feasible. The results in Table 25 indicate that only a few cases of the logistics systems satisfy the criteria: that is, for the biomass module system, the cases of the distance 16-80 km for both crop yields of 11.2 and 22.4 Mg/ha satisfy this criterion; and for the bale system, only one case of the distance 16 km for a crop yield of 22.4 does.

Table 26 Estimates of performance measures of each biomass logistics systems regarding transportation distances and crop yield

System	Trans. Distance (km (mile))	Collection & Processing ($/Mg)		Transportation ($/Mg)		Logistics Cost ($/Mg)		Energy (MJ/Mg)		CO_2 ($kgCO_2$/Mg)		Net Biomass Yield (Mg)		Man Hour (Hour)	
		Y11[(a)]	Y22[(b)]	Y11	Y22	Y11	Y22	Y11	Y22	Y11	Y22	Y11	Y22	Y11	Y22
Module	16 (10)	19.19	12.59	5.03	5.03	**24.22**[*]	**17.62**[*]	376	263	26	18	197,534	197,536	19,527	14,420
	40 (25)	18.80	12.24	9.73	9.73	**28.53**[*]	**21.97**[*]	479	366	33	25	197,534	197,536	26,541	21,416
	80 (50)	18.20	11.62	17.52	17.52	**35.72**[*]	**29.14**[*]	650	537	45	37	197,534	197,536	38,229	33,078
	161 (100)	17.00	10.38	33.11	33.11	50.10	43.49	993	879	68	60	197,534	197,536	61,607	56,400
Silage	16 (10)	15.83	11.59	35.75	35.75	51.57	47.34	1,073	1,001	74	69	170,380	170,363	73,192	68,426
	40 (25)	19.72	15.52	66.41	66.41	86.13	81.93	1,867	1,795	128	123	170,380	170,363	119,866	115,162
	80 (50)	23.52	19.40	120.16	120.16	143.68	139.56	3,191	3,120	219	214	170,380	170,363	197,585	193,016
	161 (100)	31.25	27.12	227.66	227.66	258.91	254.78	5,839	5,769	400	396	170,380	170,363	353,285	348,640
Bale	16 (10)	36.74	26.66	8.08	8.08	44.82	**34.73**[*]	536	417	39	31	157,317	157,317	73,148	66,061
	40 (25)	32.30	27.39	14.43	14.43	46.73	41.82	706	586	51	42	157,317	157,317	82,217	75,229
	80 (50)	33.50	28.59	25.05	25.05	58.55	53.64	988	868	70	62	157,317	157,317	97,521	90,526
	161 (100)	35.90	31.00	46.29	46.29	82.19	77.29	1,552	1,432	109	100	157,317	157,317	128,128	121,157

(a): Crop yield of 11.2 Mg/ha (5 ton/ac)

(b): Crop yield of 22.4 Mg/ha (10 ton/ac)

*: meet the 2012 goal of DOE in the biomass multi-year program plan

6.4.3.2 Transportation cost

Figure 30 shows estimates of transportation cost for each logistics system. Since transportation cost does not increase as crop yield increases (see Table 26), Figure 30 has no data plot associated with the crop-yield change. The trend line equations are provided to estimate transportation cost for other transportation distances between the simulation data points. The slope and the intercept of the trend line equation can be considered as the variable and fixed costs of transportation, respectively.

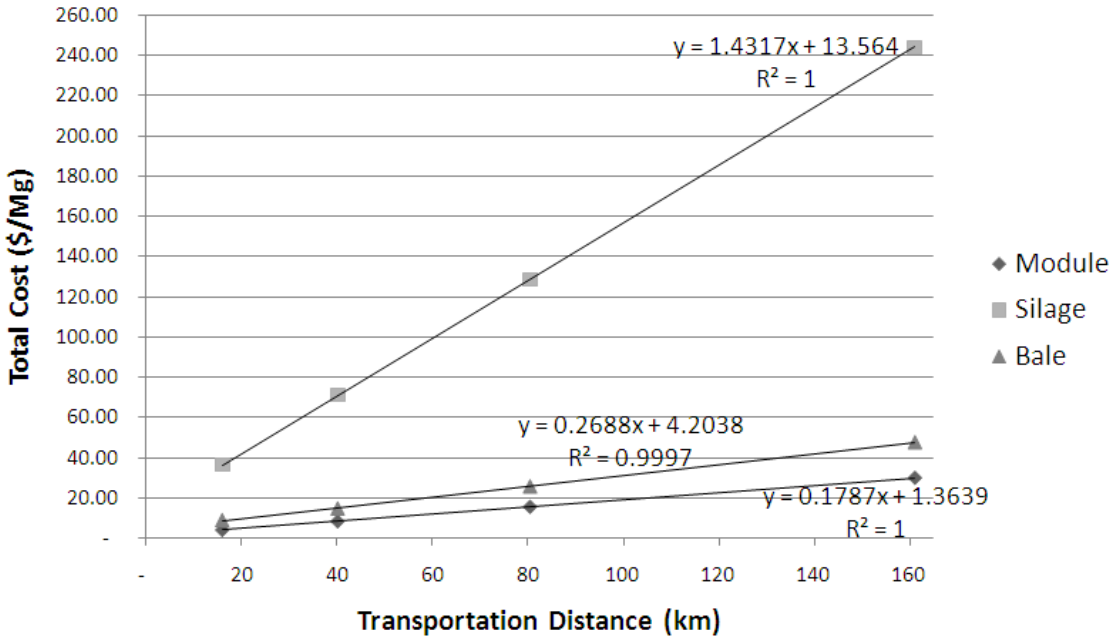

Figure 30 Transportation cost of several logistics systems regarding transportation distance

For all systems, each linear trend line strongly corresponds with its simulation data (i.e., $R^2 \geq 0.9997$). The biomass module system has the lowest slope and intercept, implying that it has the least impact from the transportation distance and may be appropriate for long-distance transportation compared to other systems.

6.4.3.3 Other performance measures

Table 25 gives all detail estimates for several performance measures. Estimates of energy used and CO_2 emissions are proportional to the total costs. The biomass module system was estimated to deliver the largest amount of biomass to a conversion plant among all logistics systems (i.e., the smallest dry matter loss). This may be mainly because it uses a plastic package to cover each module so that physical dry matter loss is very low (i.e., close to zero) while transporting modules. On the other hand, the bale system shows the smallest net biomass yield supplied (i.e., the largest dry matter loss) because it has more operations than others so that more physical dry matter loss can occur in such additional operations. For man-hour results, the biomass module system shows the best results. This is partly because it has simplified operations and delivers highly densified biomass.

6.4.3.4 Total expected profit

To our knowledge, any previous studies for biomass logistics system have not addressed a measure which deals with both perspectives of cost and net biomass yield simultaneously. Thus, This subsection presents total expected profit that deals with both perspectives of those simultaneously.

Simulation models estimated logistics cost based on a unit amount of biomass ($/Mg), a *unit logistics cost*, and a net biomass yield as reported in Table 4. Revenue from a unit biomass, *unit revenue*, may not be affected directly by the selection of logistics systems because revenue is based on the selling price of the final product (e.g., ETOH). In general, profit from a unit biomass, *unit profit*, can be calculated from unit revenue minus unit cost which includes unit logistics cost and unit cost of all other processes except logistics (e.g., cost of conversion process). Since total expected profit for a given farm land area can be estimated from unit profit times net biomass yield, it involves both perspectives of logistics cost and net biomass yield.

Even though various net biomass yields might incur different unit cost of all other processes except logistics due to economy of scale, to analyze the effects of logistics cost and net biomass yield on profit more clearly, this study simplified a procedure to estimate total expected profit. First, this study invoked an assumption that unit costs of all other processes except logistics are the same for various net biomass yields. Second, tax was not considered in the analysis. The following steps describe the procedure to estimate total expected profit by using logistics cost and net biomass yield:

1. In general, unit profit = unit revenue – unit cost.

2. By decomposing unit cost into unit logistics cost and unit cost of other processes,

    unit profit = unit revenue – unit logistics cost – unit cost of other processes.

3. Let UR_UC denote unit revenue – unit cost of other processes. Then, unit profit can be re-expressed:

    unit profit = UR_UC – unit logistics cost.

4. By using net biomass yield and unit profit, total expected profit is:

    total expected profit = net biomass yield * unit profit

    = net biomass yield * (UR_UC – unit logistics cost)

    = net biomass yield * UR_UC – net biomass yield * unit logistics cost.

Note that the UR_UC is associated with some exogenous variables including, for example, end product price, government subsidy, and cost of conversion technology.

Based on this relationship, Table 27 presents equations of total expected profit for several scenarios in each logistics system. To generate positive total profit from a given farm land area, UR_UC must be greater than unit logistics cost. Since net biomass yield is the gradient of total expected profit (i.e., the slope of the equation), it affects significantly total expected profit. Unit logistics cost affects the y-axis intercept of the equation (i.e., – net biomass yield *

unit logistics cost).

Table 27 Estimates of total expected profit based on net biomass yield and unit logistics cost

System	Trans. Distance (km)	Total Expected Profit ($)	
		Y11	Y22
Module	16	*197534*UR_UC- 197534*24.22*	*197534*UR_UC - 197534*17.62*
	40	*197534*UR_UC - 197534*28.53*	*197534*UR_UC - 197534*21.97*
	80	*197534*UR_UC - 197534*35.72*	*197534*UR_UC - 197534*29.14*
	161	*197534*UR_UC - 197534*50.10*	*197534*UR_UC - 197534*43.49*
Silage	16	*170380*UR_UC - 170380*51.57*	*170380*UR_UC - 170380*47.34*
	40	*170380*UR_UC - 170380*86.13*	*170380*UR_UC - 170380*81.93*
	80	*170380*UR_UC - 170380*143.68*	*170380*UR_UC - 170380*139.56*
	161	*170380*UR_UC - 170380*258.91*	*170380*UR_UC - 170380*254.78*
Bale	16	*157317*UR_UC - 157667*44.82*	*157317*UR_UC - 157667*34.73*
	40	*157317*UR_UC - 157317*46.73*	*157317*UR_UC - 157317*41.82*
	80	*157317*UR_UC - 157317*58.55*	*157317*UR_UC - 157317*53.64*
	161	*157317*UR_UC - 157317*82.19*	*157317*UR_UC - 157317*77.29*

To illustrate the use of equations in Table 27, suppose that switchgrass is harvested and supplied to a biofuel refinery. Mandil and Shihab-Eldin (2010) reported that refining cost for switchgrass is about $0.39/liter ($1.46/gallon) and ETOH yield could be up to 300 liters/Mg by using current biochemical conversion process. Figure 31 shows estimates of the total expected profit of each biomass logistics system for switchgrass under the condition of transportation distance of 16 km and crop yield of 11.2 Mg/ha. The x-axis value, UR_UC, can be determined by using ETOH yield, ETOH price, and refining cost: (a) if ETOH price is at $0.66/liter ($2.5/gallon), UR_UC is $81.13/Mg (=300liters/Mg * ($0.66/liter - $0.39/liter)); (b) at $0.79/liter ($3.0/gallon), $120.75/Mg; and (c) at $0.92/liter ($3.5/gallon), $160.38, respectively.

The biomass module system shows higher profit and a wider profitable range with respect to UR_UC than other systems. The bale system is better than the silage system, if UR_UC is less than $132.86/Mg (e.g., (a) and (b)); the silage system is better than the bale, if

not (e.g., (c)). This shows an example that a logistics system with lower logistics cost may not always be better than another system with higher logistics cost, implying that it may be necessary and valuable to consider both perspectives of logistics cost and net biomass yield simultaneously for evaluating biomass logistics system more appropriately.

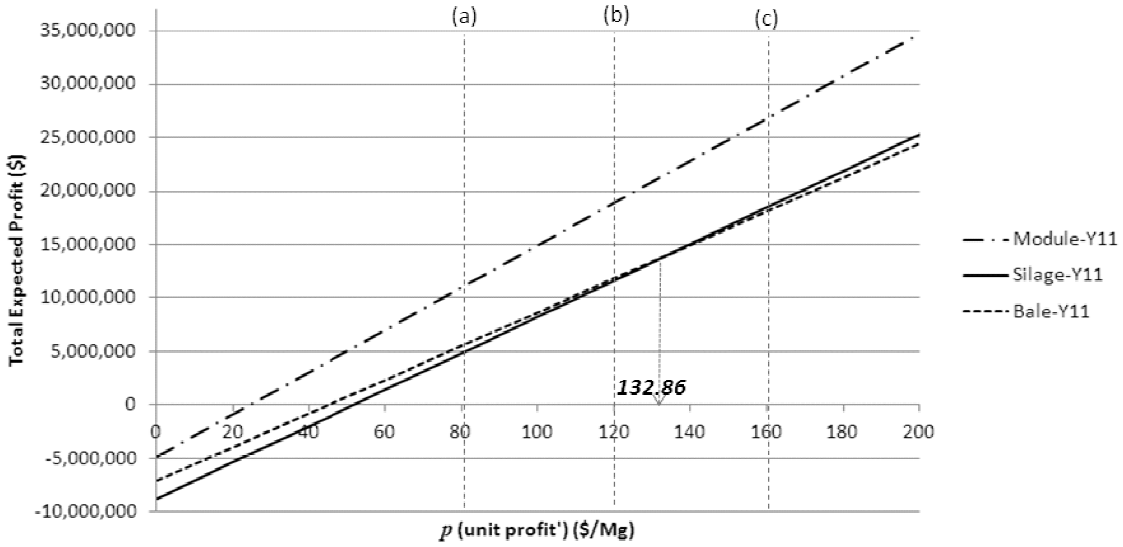

Figure 31 Example of the total expected profits: transportation distance of 16 km and crop yield of 11.2 Mg/ha

CHAPTER VII

CONCLUSIONS AND FUTURE RESEARCH

This dissertation provides an effective approach to design the SC of the cellulosic biofuel industry. It achieves its purpose in five related parts: literature review of the biofuel SC studies; formulation of BSCP with a case study on a region in Central Texas; an exact solution approach to solve large-scale instances of BSCP; development of new IBSAL simulation elements to model the biomass module system; and evaluation of alternative biomass logistics systems.

7.1 Conclusion and future research on the literature review of the biofuel SC studies

The first part of this dissertation reviews research on SCs for the biofuel and petroleum-based fuel industries as well as the literature on relevant, generic supply-chain models that have been published over the last decade. This paper identifies trends in generic SCM: internationalizing facility location, increasing use of IT, improving sustainability, and managing product perishability. It presents taxonomies of supply-chain studies that deal with biofuel and petroleum-based fuels, categorizing papers according to decision level (strategic, tactical, operational, and integrated) as well as process level (upstream, midstream, and downstream). By reviewing quantitative models available for petroleum-based fuels and generic SCM studies, this paper identifies gaps in biofuel research.

The biofuel industry is on the verge of growing explosively due to environmental regulations and renewable, sustainable energy needs. Currently, the industry is relying upon pilot plants to develop efficient processes to produce cellulosic biofuel and verify its economic viability. Operations Research can play a pivotal role in providing decision support to optimize the biofuel SC and to predict how relevant parameters affect system performance and economic

viability. The ultimate transition from pilot plant to a large-scale, commercial system will give rise to new issues that Operations Research models can address to enhance economic viability. Therefore, studies of the biofuel SC are indispensible and future research, as recommended by this paper, will contribute to the growth and viability of the biofuel industry.

Based on the perspectives that this paper provides, we now recommend fertile opportunities for research to contribute to biofuel SCM.

(1) Strategic level. In the near future, the U.S. government may begin to regulate the release of GHG from power plants, cars, and factories with the stated goal of reducing global warming. Even though studies have dealt with network design, including locations for feedstock, preprocessing facilities, and bio-refineries, no study has considered the impact of GHG emissions on design. To enable the biofuel industry to be more environmentally friendly, network design must deal with GHG emissions. Because demand for biofuel can be expected to grow dramatically in the future as government support increases and crude oil prices rise and because technologies are in a state of flux, facility location and capacity planning are very important topics. In particular, it may be possible to adapt capacity planning models that have been proposed for generic SCs. However, the multi-level facility network design problem is computationally challenging because of its large scale. Thus, appropriate solution methodologies are needed.

(2) Tactical level. One characteristic aspect of the biofuel SC is that feedstock (i.e., biomass) deteriorates over time. Inventory models are needed to quantify tradeoffs between the cost of biomass loss in storage and the cost of preprocessing capacity. Larger capacities would allow preprocessing facilities to process biomass more quickly, so that less storage capacity would be required. However, capacity is expensive and it may be more economical to build plants of lesser capacities and incur the costs of providing storage facilities and of biomass degradation. Another

characteristic of biofuel SC is that some important feedstocks (e.g., herbaceous crops) are harvested seasonally. Because feedstock availability is limited during some part of the year, storage or/and alternative biomass sources are required. This topic offers a fertile opportunity for researchers to optimize the SC appropriately. Paralleling inventory-control models that have been formulated for crude oil and oil products, new models are needed to prescribe inventory policies for biomass and biofuel based on market forces.

(3) Operational level. Several types of biomass can be used for feedstock; most are subject to restrictions such as seasonal harvesting and sustainability requirements. Since sufficient supplies of biomass must be developed to meet demand, optimizing harvesting operations to improve productivity is an important issue. We note that no planning models address the biofuel conversion processes, but several studies address the planning of petroleum-refinery operations and it may be possible to adapt them for biofuel applications. We expect that ongoing studies will identify efficient conversion processes in the near future, so that such planning models will soon be required.

(4) Integration of strategic, tactical, and operation levels. Integration of strategic-, tactical-, and operational-level decisions is an ongoing issue relative to the generic SC. In comparison, integrating the three levels is more important for the biofuel industry: biomass has low energy content so that any possible losses in the SC must be minimized to enable the economic viability of the industry. Accordingly, optimization approaches that are capable of solving large-scale instances effectively will play a key role in the economic vitality of the biofuel industry.

(5) Risk management. The biofuel industry is more vulnerable to risk than many other industries because feedstock yield depends on the weather and could be negatively affected by pests and diseases. In addition, biofuel must compete with petroleum-based fuels, the price of which is highly variable due to complex relationships between supply, demand, fuel trading

futures, and limited reserves. For example, the Green Hunter Energy biodiesel refinery in Houston, which had been the largest in the U.S., recently was offered for sale at an undisclosed price after the owner shut down the $70 million facility in the face of low demand. Such setbacks have resulted from the dramatic fall in petroleum-based fuel prices during the 2008-09 global recession (The Wall Street Journal, 2009a). This experience highlights the need for risk management in the biofuel industry.

(6) Impact on the petroleum industry. Since anticipated, new regulations on GHGs will likely increase the use of biofuel, demand will likely become more predictable, regardless of the price of petroleum-based fuel. Amounts of fuel produced in petroleum refineries will be correspondingly reduced; in fact, some facilities will be shut down. We expect that opportunities may evolve to allow biofuel distribution to be integrated with the existing downstream system for petroleum-based fuel. The compatibility of the fuel types must be considered in assessing the economics of such an integration.

(7) Adopting generic SCM research. The biofuel SC is directly related to the sustainable SCM in general and to the agri-food SCM in particular. Therefore, the biofuel industry can benefit from the transfer of technologies from these related areas. In particular, these areas may lend quantitative methods to enhance sustainability, deal with perishability, and plan harvesting. In addition, biofuel SCM may require new models for international operations, perhaps adapting features from available generic international SCM. The proximity of North American Free Trade Agreement- and Central America Free Trade Agreement-member countries and the favorable climates they offer to produce biomass can be expected to stimulate growth of biofuel supply chains in those countries. While few quantitative papers have addressed supply chains in these particular areas, some models (e.g., Wilhelm et al., 2005; and Robinson and Bookbinder, 2007) are available to serve as a starting point from which to formulate models to design biofuel

supply chains. To implement IT-driven SCM, the biofuel industry must be able to prescribe plans on a real-time basis. For example, real-time information about biomass inventory could help to manage harvesting and to make decisions about biomass transportation. Finally, because feedstock-supplier locations are dispersed, GIS is a necessary tool for planning the biofuel SC.

(8) Geographical area. The amount of biofuel needed to contribute meaningfully to energy independence is prodigious, so that feedstock sources must be found widely across the U.S. Thus, interest in, and need for, biofuel is ubiquitous and, even though each geographical area may emphasize certain features (farm size, distances, climate, length of growing season, availability of water) giving advantage (or disadvantage) to certain types of biomass, all share central features in common. This paper cites research conducted in a number of states, including Texas (Boske and Woodward, 2009), Oklahoma (Oklahoma Bioenergy Center, 2007), Louisiana (Sharma et al., 2010), Mississippi (Ekşioğlu et al., 2009; and Ekşioğlu et al., 2010), Tennessee (Zhu et al., 2010), Virginia (Cundiff et al., 1997), and California (Huang et al., 2010). The Midwest has emphasized corn for ethanol (Haddad et al., 2010); and the northwest, forest products (Clark et al., 2010). Cities may provide yet other types of biomass (Fehrs, 2011). Each of these studies has dealt with biomass types that are best suited to the features that are prevalent in a region; but, in each case, the resulting model represents these features parametrically so that input data can tailor it for application in a variety of geographical areas.

7.2 Conclusion and future research on formulation of BSCP with a case study

The model to design a cellulosic biomass/biofuel SC design considers strategic- and tactical-level decisions in both upstream and downstream echelons over multiple periods. In addition, our model deals with the unique features of the biofuel industry. Through a case study that represents the Central Texas region, we identify several important applications of our model and provide insights into the significance of system components and interactions among them.

Future research can extend this study in several ways. First, specialized solution algorithms must be developed to solve large-scale instances, which could cover a larger geographical area and prescribe detailed, tactical-level decisions that must deal with more time periods in the planning horizon. Second, the relationship between storage capacity and replenishment policy in the multi-echelon system must be determined so that both can be prescribed more accurately. Third, considering several alternative transportation modes (e.g., rail and truck) may allow costs to be reduced since large quantities of biomass must be transported. Fourth, dealing explicitly with uncertainty by using stochastic programming models can be expected to lead to robust SC designs. Lastly, models could be formulated to represent the interests of specific stake holders (e.g., biomass suppliers, refineries, distribution centers) so that ways in which they could co-operate to improve efficiency and profitability can be identified.

7.3 Conclusion and future research on an exact solution approach

This paper presents a new solution approach for BSCP. We deal with material flows based on the viewpoint of single-commodity, generalized flows, an alternative to the multi-commodity flow model of An et al. (2011b). We use a CG approach to solve the linear relaxation of model 2 at the root node. In this CG context, our BRA solves the subproblem, an uncapacitated, embedded generalized minimum cost circulation problem, generating improving flow-paths (i.e., columns) effectively in $O(m)$.

We devise POCs, inequalities based on a portion of the objective function and augment them to the linear relaxation of BSCP to cut off some portion of the B&B search tree, facilitating solution. Average ratios of runtime of our methods to that of CPLEX are 90% for POC1, 91% for POC2, 76% for POC1&2, respectively, implying that POC1&2 outperforms, on average, CPLEX, POC1, and POC2. POC1&2 improved best bound in the initial stage of B&B search. POC1&2 might be helpful to find integral solution earlier than others, so that it may contribute

to faster convergence than POC1 and POC2. However, the initial better bound does not guarantee less runtime, but, on average, performs better.

Several studies associated with B&B did not deal with each instance in detail for their analysis, but considered, in general, average performance. Providing better information to strong branching may not guarantee better solution because we can consider strong branching as a greedy heuristic.

Our CG approach solves the linear relaxation of BSCP at the root node of the B&B search tree faster than CPLEX. For one instance (F34T4), its runtime is just 11.3% of that of CPLEX. Our approach can be applied to many other important problems that involve an embedded GFP.

CPLEX cuts do not reduce runtimes to solve our BSCP instances. Our solution approach, which involves use of CG at the root node and CPLEX B&B supplemented with POC(s), solves most of our BSCP instances faster than CPLEX. Even though we had to maintain $Z_{inc} = 0$ or the incumbent solution from CPLEX heuristics because CPLEX does not allow a right-hand-side constant to be changed after branching begins, using POC(s) gives good results, implying that such inequalities may be helpful in accelerating B&B.

Finally, we suggest several fertile topics for future research. First, in addition to the embedded GFP sub-problem, it would be helpful to define a second type of sub-problem that is an integer problem to allow bounds at B&B nodes to be tightened to facilitate solution. Second, since our CG approach is not guaranteed to generate all alternative optimal shortest paths, it will be interesting to study the impact of making only alternative optimal shortest paths available to the master problem. Third, our approach incorporates several path-flows (columns) with positive reduced cost after solving the subproblem once at each CG iteration to accelerate CG convergence. However, future research could investigate avenues to improve our methods, better

managing the number of paths generated to enhance solution capability. Fourth, POC appears to offer promise and could be used to other mixed-integer programs that involve a similar objective function structure to ours. Finally, our solution approach offers promise in application to other problems that involve an embedded generalized flow problem.

7.4 Conclusion and future research on new IBSAL elements for the biomass module system

Several IBSAL elements were developed to evaluate a conceptual logistics system based on biomass modules. Those elements described the anticipated operation of a self-propelled forage harvester coupled to a module former and a module hauler capable of loading and unloading two modules to and from flat-bed semi-trailers. Simulation models based on the new and existing elements in IBSAL suggested the biomass logistics cost to range between $19.65 and $41.26/Mg with varying yield levels and transport distances. A sensitivity analysis indicated the factors with greatest influence on cost to be module density, forage harvester field efficiency and module hauler transport speed.

7.5 Conclusion and future research on evaluation of alternative biomass logistics systems

This study developed simulation models based on IBSAL to address several biomass logistics systems, including biomass module, silage, and bale systems. The biomass module system shows the best estimates for cost, energy, CO2 emissions and man-hour among biomass logistics systems considered. In particular, the transportation cost of the biomass module system was least affected by transportation distance compared to other systems. In addition, the doubling of crop yield reduced the CP cost significantly, ranging from 14 to 34%. The biomass module system was expected to generate the best total expected profit.

This paper suggests several future studies to improve the biomass logistics system further. The first is considering several weather conditions to investigate the impact of weather condition. The second is applying this study to other locations and various crops to obtain more

robust evaluation result. The third is devising a method to optimize the number of machines used in the simulation model to consider the restriction of available machines in the region studied. Finally, even though this study has not used any storage elements from IBSAL, it will be needed to improve the logic for storage in IBSAL to consider more realistic aspects of storage operation because current logic invokes some unrealistic assumptions that biomass materials will be stored steadily and the number of turnover is a constant value.

REFERENCES

Aden A., Ruth M., Ibsen K., Jechura J., Neeves K., Sheehan J. and Wallace B. (2002) Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. *Technical Report NREL/TP*-510-32438.

Ahuja R.K., Magnanti T.L. and Orlin J.B. (1993) Network flows theory, algorithms, and applications, Prentice-Hall, Inc.

Ahumada O. and Villalobos J.R. (2009) Application of planning models in the agri-food supply chain: A review. *European Journal of Operational Research*, **195**, 1-20.

Al-Qahtani K. and Elkamel A. (2008) Multisite facility network integration design and coordination: An application to the refining industry. *Computers and Chemical Engineering*, **32**, 2189-2202.

Alvelos F. and Valerio de Carvalho J.M. (2007) An extended model and a column generation algorithm for the planar multicommodity flow problem. *Networks*, **50**(1), 3-16.

An H., Wilhelm W.E. and Searcy S.W. (2011a) Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass & Bioenergy*, **35** (9), 3763-3774.

An H., Wilhelm W.E. and Searcy S.W. (2011b) A mathematical model to design a lignocellulosic biofuel supply chain system with a case study based on a region in Central Texas. *Bioresource Technology*, **102** (17), 7860-7870.

Anderson K.B. and Noyes R.T. (2010) Grain storage costs in Oklahoma. OSU Fact Sheets, 210, 2-4.

Aronofsky J.S. and Williams A.C. (1962) The use of linear programming and mathematical models in underground oil production. *Management Science*, **8**(4), 394-407.

Bilgen B. and Ozkarahan I. (2004) Strategic tactical and operational production-distribution

models: a review. *International Journal of Technology Management*, **28**(2), 151-171.

Birge J.R. and Louveaux F. (1997) Introduction to Stochastic Programming. Springer Series in Operations Research, New York, NY.

Boske L.B. and Woodward J.T. (2009) Analysis of Texas Biofuel Supply Chains Originating in the United States and Brazil. *National Technical Information Service*, Report Numbers:SWUTC/09/169201-1.

Brimberg J., Hansen P., Lih K.W., Mladenovic N. and Breton M. (2003) An oil pipeline design problem. *Operations Research*, **51**(2), 228-239.

Brunner J.O., Bard J.F. and Kolisch R. (2011) Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Transaction*, **43**, 84-109.

Catchpole A.R. (1962) The application of linear-programming to integrated supply problems in the oil industry. *Operations Research Quarterly*, **13**(2), 161-169.

Chenga L. and Duran M.A. (2004) Logistics for world-wide crude oil transportation using discrete event simulation and optimal control. *Computers and Chemical Engineering*, **28**, 897-911.

Clark J., Sessions J., Wing M. and Schmidt C. (2010) Forest Biomass Assessment and Transportation Analysis for Evaluating the Feasibility of Establishing a Biomass Power Plant in Northwest Oregon. *In Proceedings of the 33rd Annual Meeting of the Council on Forest Engineering: Fueling the Future*, 1-8.

Constantino M., Martins I. and Borges J.G. (2008) A new mixed-integer programming model for harvest scheduling subject to maximum area restrictions. *Operations Research*, **56**(3), 542-551.

Cundiff J.S., Dias N. and Sherali H.D. (1997) A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource Technology*, **59**, 47-55.

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

Dempster M.A.H., Hicks Pedron N., Medova E.A., Scott J.E. and Sembos A. (2000) Planning logistics operations in the oil industry. *Journal of the Operational Research Society*, **51**, 1271-1288.

Desrochers M. and Soumis F.A. (1988) Generalized permanent labeling algorithm for the shortest path problem with time windows. *INFORMATION*, **26**,193-214.

Dumitrescu I. and Boland N. (2003) Improved preprocessing, labeling and scaling algorithms for the weight-constrained shortest path problem. *Networks*, **42**(3), 135-53.

Dunnett A., Adjiman C. and Shah N. (2007) Biomass to heat supply chains applications of process optimization. *Transactions of IChemE, Part B, Process Safety and Environmental Protection*, **85** (B5), 419-429.

Ekşioğlu S.D., Acharya A., Leightley L.E. and Arora S. (2009) Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers & Industrial Engineering*, **57**, 1342-1352.

Ekşioğlu S.D., Li S., Zhang S., Sokhansanj S. and Petrolia D. (2010) Analyzing impact of intermodal facilities on design and management of biofuel supply chain. Transportation Research Record: *Journal of the Transportation Research Board*, **2191**, 144-151.

Escudero L.F., Quintana F.J. and Salmeon J. (1999) CORO, a modeling and an algorithmic framework for oil supply, transformation and distribution optimization under uncertainty. *European Journal of Operational Research*, **114**, 638-656.

Fehrs J.E. (1999) Secondary Mill Residues and Urban Wood Waste Quantities in the United States. Report in Northeast Regional Biomass Program;

(http://www.nrbp.org/publications/, last accessed on September 28, 2011).

Fiedler P., Lange M. and Schultze M. (2007) Supply logistics for the industrialized use of biomass principles and planning approach. *In Proceeding of the International Symposium on Logistics and Industrial Informatics*, September, Wildau, Germany, 41-46.

Gallis C.T. (1996) Activity oriented stochastic computer simulation of forest biomass logistics in Greece. *Biomass and Bioenergy*, **10**, 377-382.

Gemtos T.A. and Tsiricoglou T. (1999) Harvesting of cotton residue for energy production. *Biomass and Bioenergy*, **16**, 51-59.

Gigler J.K., Hendrix E.M.T., Heesen R.A., van den Hazelkamp V.G.W. and Meerdink G. (2002) On optimisation of agri chains by dynamic programming. *European Journal of Operational Research,* **139**, 613-625.

Glassner D.A., Hettenhaus J.R. and Schechinger T.M. (1998) Corn stover collection project. *In Proceedings of BioEnergy '98: Expanding BioEnergy Partnerships*, **2**, 1100-1110.

Goetschalckx M., Vidal C.J. and Dogan K. (2002) Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research,* **143**, 1-18.

Goldfarb D. and Lin Y. (2002) Combinatorial interior point methods for generalized network flow problems. *Mathematical Programming Series A*, **93**, 227-246.

Goyal S.K. and Giri B.C. (2001) Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, 134, 1-16.

Goycoolea M., Murray A.T., Barahona F., Epstein R. and Weintraub A. (2005) Harvest scheduling subject to maximum area restrictions-exploring exact approaches. *Operations Research*, **53**(3), 490-500.

Graham R.L., English B.C. and Noon C.E. (2000) A geographic information system based modeling system for evaluating the cost of delivered energy crop feedstock. *Biomass and Bioenergy*, **18**, 309-329.

Gronalt M. and Rauch P. (2007) Designing a regional forest fuel supply network. *Biomass and Bioenergy*, **31**, 393-402.

Grønhaug R., Christiansen M., Desaulniers G. and Desrosiers J. (2010) A branch-and-price method for a liquefied natural gas inventory routing problem. *Transportation Science*, **44**(3), 400-415.

Gunn E.A. and Richards E.W. (2005) Solving the adjacency problem with stand-centred constraints. *Canadian Journal of Forest Research*, **35**, 832-842.

Gunnarsson H., Ronnqvist M. and Lundgren J.T. (2004) Supply chain modeling of forest fuel. *European Journal of Operational Research*, **158**, 103-123.

Gutierrez-Jarpa G., Desaulniers G., Laporte G. and Marianov V. (2010) A branch-and-price algorithm for the vehicle routing problem with deliveries, selective pickups and time windows. *European Journal of Operations Research*, **206**(2), 341-349.

Haddad M.A., Taylor G. and Owusu F. (2010) Locational Choices of the Ethanol Industry in the Midwest Corn Belt. *Economic Development Quarterly*; **24**, DOI: 10.1177/089124240934772.

Hamelinck C.N., Hooijdonk G.V. and Faaij A.P.C. (2005) Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. *Biomass and Bioenergy*, **28**, 384-410.

Hamelinck C.N., Suurs R.A.A. and Faaij A.P.C. (2005) International bioenergy transport costs and energy balance. *Biomass and Bioenergy*, **29**, 114-134.

Higgins A.J. (1999) Optimizing cane supply decisions within a sugar mill region. *Journal of*

Scheduling, **2**, 229-244.

Higgins A.J. (2002) Australian sugar mills optimize harvester rosters to improve production. *Interfaces*, **32**(3), 15-25.

Higgins A.J. and Postma S. (2004) Australian sugar mills optimize siding rosters to increase profitability. *Annals of Operations Research*, **128**, 235-249.

Huang Y., Chen C.W. and Fan Y. (2010) Multistage optimization of the supply chains of biofuels. *Transportation Research Part E*, **46**, 820-830.

Imagine That, Inc. Extendsim User manual. Version 8.0.1. San Jose, CA: Imagine That Inc; 2010.

Jenkins B.M., Arthur J.F., Miller G.E. and Parsons P.S. (1984) Logistics and economics of biomass utilization. *Transactions of the ASAE*, **27**(6), 1984-1904.

Jones K.L., Lustig I.J., Farvolden J.M. and Powell W.B. (1993) Multicommodity network flows: The impact of formulation on decomposition. *Mathematical Programming*, **62**, 95-117.

Kamath A. and Palmon O. (1995) Improved interior point algorithms for exact and approximate solution of multicommodity flow problems. *In Proceedings of 6th annual ACM-SIAM Symposium on Discrete Algorithms*, 502-511.

Khor C.S., Elkamel A., Ponnambalam K. and Douglas P.L. (2008) Two-stage stochastic programming with fixed recourse via scenario planning with economic and operational risk management for petroleum refinery planning under uncertainty. *Chemical Engineering and Processing*, **47**, 1744-1764.

Kim Y., Yun C., Park S.B., Park S. and Fan L.T. (2008) An integrated model of supply network and production planning for multiple fuel products of multi-site refineries. *Computers and Chemical Engineering*, **32**, 2529-2535.

Klingman D., Phillips N., Steiger D., Wirth R., Padman R. and Krishnan R. (1987) An

optimization based integrated short-term refined petroleum product planning system. *Management Science*, **33**(7), 813-830.

Klose A. and Drexl A. Facility location models for distribution system design. *European Journal of Operational Research 2005*, **162**, 4-29.

Kofman P.D. (2006) Quality wood chip fuel. *COFORD Connects Notes: Harvesting / Transportation*. No. 6, 1-4.

Kouvelis P., Chambers C. and Wang H. (2006) Supply chain management research and production and operations management: review, trends, and opportunities. *Production and Operations Management*, **15**(3), 449-469.

Kumar A. and Sokhansanj S. (2007) Switchgrass (Panicum vigratum, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresource Technology*, **98**, 1033-1044.

Lababidi H.M.S., Ahmed M.A., Alatiqi I.M. and Al-Enzi A.F. (2004) Optimizing the supply chain of a petrochemical company under uncertain operating and economic conditions. I*ndustrial Engineering Chemistry Research Journal*, **43**(1), 63-73.

Lejars C., Gal P.Y.L. and Auzoux S. (2008) A decision support approach for cane supply management within a sugar mill area. *Computers and Electronics in Agriculture*, **60**, 239-249.

Levary R.R. and Dean B.V. (1980) A natural gas flow model under uncertainty in demand. *Operations Research*, **28**(6), 1360-1374.

Li W., Hui C.W. and Li A. (2005) Integrating CDU, FCC and product blending models into refinery planning. *Computers and Chemical Engineering*, **29**, 2010-2028.

Liang D. and Wilhelm W.E. (2010) A generalization of column generation to accelerate convergence. *Mathematical Programming Series A*, **122**(2), 349-378

Mantovani B. and Gibson H. (1992) A simulation model for analysis of harvesting and transport costs for biomass based on geography, density and plant location. In Analysis of Agricultural Energy Systems. *Energy in World Agriculture*: R.M. Peart & R.C. Brook (Eds.), Vol. 5. Elsevier, Amsterdam, 253-280.

Martins I., Constantino M. and Borges J.G. (2005) A column generation approach for solving a non-temporal forest harvest model with spatial structure constraints. *European Journal of Operational Research*, **161**, 478-498.

Meixell M.J. and Gargeya V.B. (2005) Global supply chain design: A literature review and critique. *Transportation Research Part E*, **41**, 531-550.

Melo M.T., Nickel S. and Saldanha-da-Gama F. (2009) Facility location and supply chain management - A review. *European Journal of Operational Research*, **196**, 401-412.

Mestry S., Damodaran P. and Chen C. (2011). A branch and price solution approach for order acceptance and capacity planning in make-to-order operations. *European Journal of Operational Research*, **211**(3), 480-495.

Milbrandt A. (2005) A geographic perspective on the current biomass resource availability in the United States. Technical Report NREL/TP-560-39181.

Min H. and Zhou G. (2002) Supply chain modeling- past, present and future. *Computers & Industrial Engineering*, **43**, 231-249.

MirHassani S.A. (2008) An operational planning model for petroleum products logistics under uncertainty. *Applied Mathematics and Computation*, **196**, 744-751.

Morey R.V., Kaliyan N., Tiffany D.G. and Schmidt D.R. (2010) A corn stover supply logistics system. *American Society of Agricultural and Biological Engineers*. **26**(3), 455-461.

Murray A.T. (1999) Spatial restrictions in harvest scheduling. *Forest Science*, **1**, 45-52.

Murray S.M. (1992) An interior point approach to the generalized flow problem with costs and

related problems, Ph.D. Thesis, Stanford University, Stanford CA.

National Priorities Project Database (http://database.nationalpriorities.org., last accessed on August 1 2010).

Neiro S.M.S. and Pinto J.M. (2004) A general modeling framework for the operational planning of petroleum supply chains. *Computers and Chemical Engineering*, **28**, 871-896.

Nguyen M.H. and Prince R.G.H. (1996) A simple rule for bioenergy conversion plant size optimisation: bioethanol from sugar cane and sweet sorghum. *Biomass and Bioenergy*, **10**, 561-465.

Oklahoma Bioenergy Center 2007 Annual Report (http://okbioenergycenter.org/downloads-links/, last accessed on September 28, 2011).

Perrin R., Vogel K., Schmer M. and Mitchell R. (2008) Farm-Scale Production Cost of Switchgrass for Biomass. *Bioenergy Research*, **1**, 91-97.

Petrou E. and Mihiotis A. (2007) Design of a factories' supply system with biomass in order to be used as an alternative fuel-a case study. *Energy Fuels*, **21**(6), 3718-3722.

Pinto J.M. and Moro L.F.L. (2000) A planning model for petroleum refineries. *Brazilian Journal of Chemical Engineering*, **17**, 575-586.

Ravula P.P., Grisso R.D. and Cundiff J.S. (2008) Cotton logistics as a model for a biomass transportation system. *Biomass and Bioenergy*, **32**, 314-325.

Robinson A.G. and Bookbinder J.H. (2007) NAFTA supply chains: Facilities location and logistics. *International Transactions in Operational Research*, **14**, 179-199.

Ronen D. (1995) Dispatching petroleum-products. *Operations Research*, **43**(3), 379-387.

Santoso T., Ahmed S., Goetschalckx M. and Shapiro A. (2005) A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, **167**(1), 96-115.

Sear T.N. (1993) Logistics planning in the downstream oil industry. *Journal of Operational Research Society*, **44**(1), 9-17.

Senne E.L.F., Lorena L.A.N. and Pereira M.A. (2005) A branch-and-price approach to p-median location problems. *Computers & Operations Research*, **32**, 1655-1664.

Seuring S. and Muller M. (2008) From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, **16**, 1699-1710.

Sharma P., Sarker B.R. and Romagnoli J.A. (2010) Integrated framework for enterprise management-a synergistic approach towards sustainable biorefineries. *Computer Aided Chemical Engineering*, **28**, 1009-1014

Shinners K.J. and Binversie B.N. (2004) Harvest and storage of wet corn stover biomass. An ASAE/CSAE Meeting Presentation, Paper Number: 041159.

Sokhansanj S., Kumar A. and Turhollow A.F. (2006a) Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy*, **30**, 838-847.

Sokhansanj S., Turhollow A.F. and Wilkerson E.G. (2006b) Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). Technical Memorandum ORNL/TM-2006/57. Oak Ridge, TN: Oak Ridge National Laboratory.

Suh K. and Suh S. (2010) Economic and environmental implications of corn stover densification options for biofuel in Minnesota. *American Society of Agricultural and Biological Engineers* ISSN 2151-0032, **53**(4), 1183-1192.

Tatsiopoulos I.P. and Tolis A.J. (2003) Economic aspects of the cotton-stalk biomass logistics and comparison of supply chain methods. *Biomass and Bioenergy*, **24**, 199-214.

The Wall Street Journal (2009a) U.S. Biofuel Boom Running on Empty, Thursday, August 27, A1.

The Wall Street Journal (2009b) Wood Pellets Catch Fire as Renewable Energy Source, Tuesday, July 7, A4.

U.S. DOE Theoretical Ethanol Yield Calculator (www1.eere.energy.gov/biomass/ethanol_yield_calculator.html, last accessed on August 1 2010).

Troncosoa J.J. and Garrido R.A. (2005) Forestry production and logistics planning: an analysis using mixed-integer programming. *Forest Policy and Economics*, **7**, 625-633.

Turhollow A., Downing M. and Butler J. (1996) The cost of silage harvest and transport systems for herbaceous crops. ORNL 37831-6205. Oak Ridge, TN: Oak Ridge National Laboratory.

U.S. Department of Agriculture and U.S. Department of Energy (2008) National biofuels action plan, October (http://www1.eere.energy.gov/biomass/pdfs/nbap.pdf, last accessed on October 1 2011).

U.S. Department of Energy (2010) Biomass multi-year program plan, November (www1.eere.energy.gov/biomass/pdfs/mypp_april_2011.pdf, last accessed on October 1 2011).

U.S. Environmental Protection Agency (2010) Renewable fuel standard program (rfs2). EPA, USA (http://www.epa.gov/otaq/fuels/renewablefuels/index.htm, last accessed on October 1 2011).

Vaidya P.M. (1989) Speeding up linear programming using fast matrix multiplication. *In Proceedings of 30th IEEE Annual symposium on Foundations of Computer Science*, 332-337.

Villa C. and Hoffman K. (2006) A column-generation and branch-and-cut approach to the bandwidth-packing problem. *Journal of Research of the National Institute of Standards and Technology*, **111**, 161-185.

Wayne K.D. (2002) A polynomial combinatorial algorithm for generalized minimum cost flow. *Mathematics of Operations Research*, **27**(3), 445-459.

Wilhelm W., Liang D., Rao B., Warrier D., Zhu X. and Bulusu S. (2005) Design of international assembly systems and their supply chains under NAFTA. *Transportation Research E*, **41**, 467-493.

Wilhelm W.E. (2001) A technical review of column generation in integer programming. *Optimization and Engineering*, **2**, 159-200.

Zhu X., Li X., Yao Q. and Chen Y. (2010) Challenges and models in supporting logistics system design for dedicated-biomass-based bioenery industry. *Bioresour Technol*, **102**(2), 1344-1351.

Zhu X., and Wilhelm W. E. (2012) A three-stage approach for the resource-constrained shortest path sub-problem in a column-generation context. *Computers in Operations Research*, **39**, 164-178.

VITA

Name: Heungjo An

Address: Department of Industrial and Systems Engineering
Texas A&M University
College Station, TX 77843-3131

Email Address: Heungjo.An@gmail.com

Education: B.S., Material and Science Engineering, Hanyang University, Korea, 1997
M.S., Material and Science Engineering, Hanyang University, Korea, 1999
Ph.D., Industrial and Systems Engineering, Texas A&M University, 2011