1-2016

A multi-objective, hub-and-spoke model to design and manage biofuel supply chains

Mohammad S. Roni
Idaho National Laboratory

Sandra D. Eksioglu
Clemson University, seksiog@clemson.edu

Kara G. Cafferty
Idaho National Laboratory

Jacob J. Jacobson
Minds Eye Computing LLC

A Multi-Objective, Hub-and-Spoke Model to Design and Manage Biofuel Supply Chains

Mohammad. S. Roni[b], Sandra D. Eksioglu [a], Kara G. Cafferty[b], Jacob J. Jacobson[c]*

a: Department of Industrial Engineering, Clemson University
b: Biofuels & Renewable Energy Technologies, Idaho National Laboratory, Idaho Falls, ID
c: Minds Eye Computing LLC, Idaho Falls, ID

Abstract:
In this paper we propose a multi-objective, mixed integer linear programming model to design and manage the supply chain for biofuels. This model captures the trade-offs that exist between costs, environmental and social impacts of delivering biofuels. The in-bound supply chain for biofuel plants relies on a hub-and-spoke structure which optimizes transportation costs of biomass. The model proposed optimizes the CO_2 emissions due to transportation-related activities in the supply chain. The model also optimizes the social impact of biofuels. The social impacts are evaluated by the number of jobs created. The multi-objective optimization model is solved using an augmented ϵ-constraint method. The method provides a set of Pareto optimal solutions. We develop a case study using data from the Midwest region of the USA. The numerical analyses estimates the quantity and cost of cellulosic ethanol delivered under different scenarios generated. The insights we provide will help policy makers design policies which encourage and support renewable energy production.

Key Word: Multi-objective optimization, Hub-and-spoke supply chain, Densified biomass, Augmented ϵ-constraint method, Rail transportation

1. Introduction

Fossil fuels, such as oil, coal and natural gas currently represent the prime energy sources in the world. However, an increasing energy demand, coupled with increasing concerns over the environmental impact of fossil fuel consumption, have resulted in an increased interest in renewable energy. Some of the major sources of renewable energy are biomass, solar, and wind. The United States Department of Energy (2006) has identified biofuels as one of the future powers sources in the USA that will reduce nation's dependency on fossil fuels, thereby having a positive impact on the economy, environment, and society. A variety of biomass feedstocks are presently used to produce biofuel and electricity. According to EIA, biomass contributes nearly 3.9 quadrillion British thermal units (BTU) and accounts for more than 4% of total U.S. primary energy consumption (EIA, 2010). Over the last 30 years, the share of biomass in the total primary energy consumption has averaged less than 3.5% (EIA, 2010). The Energy Independence and Security Act of 2007 (EISA, 2007) set the Renewable Fuels Standard (RFS) in order to increase the share of biomass in the total energy production. RFS calls for an increase of cellulosic biofuel production to 16 billion gallons a year (BGY) by 2022 (USDA, 2008; Biomass Program Multi-Year Program Plan, 2010). The proposed 2014 production volume for cellulosic biofuel is 17 million gallons a year (MGY), and the proposed range is 8 – 30 MGY (EPA, 2014). Due to policies, such as RFS, it is expected that the share of biomass in the total renewable energy production will increase in the near future.

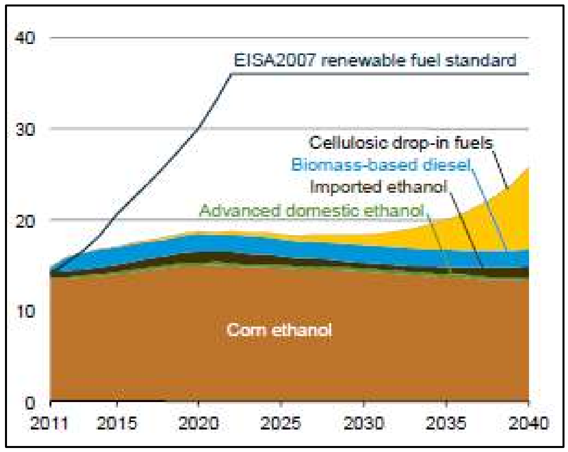


Figure 1: Increasing growth of biofuels consumption (US DOE, 2010)

Figure 1 presents the expected biofuels production for the period 2011 to 2040. The figure indicates that the production of cellulosic ethanol is expected to increase and will become a major contributor in meeting the RFS requirements. Consequently, the number of biofuel plants which produce cellulosic ethanol is expected to increase in the near future. These plants will need tools to aid their supply chain design and management decisions, such as, facility location, transportation mode selection, capacity expansion decisions, etc. One of the main contributions of this paper is the proposed optimization model which captures product and supply chain characteristics which are specific to biofuel industry. For example, a number of studies indicate that in order to reduce biomass transportation costs and make 2nd generation biofuels cost-competitive, we have to invest on large-capacity plants which gain from economies of scale in production (Hess et al., 2009). Large capacity plants would rely in a larger number of farms, most of which would be located further away. To decrease transportation costs plants would rely in using rail and barge for transportation. Additionally, biomass would be processed at the farm prior to delivery to increases its bulk density, and be transformed into a stable, dense, and flowable commodity, easier to load and unload, and cheaper to transport. These facts imply that the best design for the in-bound distribution network design is a hub-and-spoke network structure, which is indeed reflected in this model.

The main objective of many models developed and analyzed in the area of supply chain optimization, logistics management and transportation systems analysis has been minimizing costs. This is also the case with the literature related to biofuel supply chains. Most recently, there has been growing interest to incorporate environmental and social objectives to biomass supply chain models. This trend makes sense since this is a new industry, thus, there is an opportunity here to do things right from the very beginning. Another contribution of this paper is providing a model that captures the environmental impacts of biofuels by estimating CO_2 emission due to transportation, biorefinery location, and biorefinery operations. The model also captures the social impacts of biofuels by estimating the number of jobs created due to biomass production, preprocessing, transportation, and biorefinery operating.

Other papers in the literature use multi-objective optimization models to capture the economic, environmental, and social impacts of biofuels (You's et al., 2012). Different from the literature, this paper focusses on large-scale, regional biofuel supply chains. Thus, the model captures problem characteristics which become evident when you analyze large-scale supply chains. For example, based on current practices, the use of unit train to deliver biomass becomes cost competitive when transportation distances are longer than 100miles (Gonzales et al. 2013). The model we propose captures important details about rail transportation, such as, existing rail network structure and available capacities, non-linear railway cost function, and hub location costs. As a result, the model we propose can help policy makers evaluate the impacts of policies implemented at the Federal level. For example, the US Billion Ton Study led by the Oak Ridge National Laboratory indicates that there is enough biomass in the U.S. to meet the RFS goals set by EPA. The question is whether biomass can be collected and delivered to biofuel plants in a cost competitive manner. Studies like our can be used to evaluate the potential of meeting the RFS goals at the national level.

A contribution of this paper is the development of a case study which was developed using a number of reliable data sources (see Section 5). Thus, the results from the numerical analysis are very insightful. The results provide estimates of the delivery cost of cellulosic ethanol, unit emissions due to supply chain activities, and the number of new jobs created in this industry. The relationships revealed provide insights which help policy makers design policies that support renewable energy production.

Finally, the mathematical model we propose is a challenging multi-objective linear mixed integer programming (MILP) model. We used an augmented ϵ-constraint method to solve this multi-objective problem and generate a set of Pareto optimal solutions. We use lexicographic optimization to obtain the ranges of ε_1 and ε_2. Doing this provides us with better estimates of the Pareto frontiers.

2. Relevant literature

The model we propose is on-line with the following streams of research in the area of supply chain: biomass supply chain and logistics management, transportation cost analysis, hub-and-spoke network design problem, and multi-objective optimization. Next we provide a summary of these streams of research and identify our contributions.

The biomass supply chain optimization literature presents a number of deterministic and stochastic models. The deterministic models are extensions of the facility location model. These models are used to identify biorefinery sittings (Ekşioğlu et al., 2009; Parker et al., 210; Bai et al., 2011; Kim et al., 2011a; Papapostolou et al., 2011; Roni et al., 2014a; Marufuzzaman et al., 2014). Some deterministic models are used to identify the number, capacity and location of biofuel plants in order to make use of the available biomass in a particular region in a cost efficient manner. The stochastic research on biomass supply chains uses extensions of the two-stage, location-transportation stochastic programming model to identify biorefinery sittings (such as, Cundiff et al., 1997; Huang et al., 2010; Kim et al., 2011b; Chen and Fan, 2012; Gebreslassie et al., 2012).

The literature on biomass transportation cost analysis is focused on estimating truck; rail and barge transportation costs (Gonzales et al., 2013; Roni et al., 2014b). A study by Mahmudi and Flynn (2006) investigate biomass transportation by rail. A study by Ekşioğlu et al. (2011) investigate rail and barge transportation costs for biomass. Other works related to biomass logistics costs analysis are the ones by Kumar et al., 2007; Sokhansonj et al., 2006; Jacobson et al., 2014; Ren et al., 2015.

The hub-and-spoke design problem is conventionally called the hub location problem (Campbell, 2012). A number of extensions of the hub location problem are found in the literature. These extensions are proposed in order to capture issues that arise when managing this supply chain, such as, non-linear economies of scale, traffic management, transportation mode selection, and congestion. The existing literature can be divided into two major groups, the single hub (SH) and the multiple hubs (MH) location problem. In a SH location model, the routing of the flow to/from a non-hub node is done through the hub. In a MH setting, the routing of the flow to/from a non-hub node is done through multiple hubs. Thus, flow initiated from a non-hub node traverses a number of hubs before reaching its final destination. Mixed integer programs (MIP) are used to model the problem to represent the fixed hub location costs, and nodes-to-hub allocations (Skorin-Kapov et al., 1996; Campbell, 2012). Due to computational challenges faced when solving these large sized MIP models, a number of different heuristic approaches have been design to solve the problems. For example, Chen (2007) developed a hybrid Simulated Annealing heuristics, Silva and Cunga (2009) developed a number of Tabu Search heuristics, Cunha and Silva (2007) developed a hybrid Genetic Algorithm and Simulated Annealing-based heuristics, Camargo et al. (2008) present a Benders Decomposition-based solution approach and Labbe and Yaman (2004) propose a Lagrangean Relaxation-based approach. For an extensive review of this problem see Alumur and Kara (2008), Tunc et al. (2011).

A limited number of papers in the literature propose multi-objective optimization models for the biofuel supply chain design and management. For example, Zamboni et al. (2009) present a MILP model that simultaneously minimizes the supply chain operating costs and GHG emissions due to supply chain activities. Perimenis et al. (2011) provide a decision support tool to evaluate biofuel production pathways. This tool integrates technical, economic, environmental and social aspects along the entire value chain of biofuels starting from biomass production to biofuel end-use. Mele et al. (2009) address the problem of optimizing the supply chains for bioethanol and sugar production. Their bi-criteria MILP model addresses economic and environmental concerns. The model minimizes the total cost of managing the supply chain network, and minimizes the environmental impact over the entire product life cycle. El-Halwagi et al. (2013) incorporate safety concerns into the biorefinery location selection and capacity management problem. They establish tradeoffs between costs and safety issues using Pareto curves. You and Wang (2011) study the optimal design and planning of biomass-to-liquids (BTL) supply chains under economic and environmental criteria. You et al. (2012) address the optimal design and planning of cellulosic ethanol supply chains under economic, environmental, and social objectives.

Multi-objective integer linear programs have been solved using exact and heuristics solution approaches. An exact algorithm identifies the whole set of non-dominated solutions for the problem. Heuristics approximate, identify bounds for the set of non-dominated solutions. For example, Abounacer et al. (2014) propose an ε-constraint method to generate an exact Pareto frontier of a complex three objective location-transportation problem. The following is a list of exact methods. Zhang and Reimann (2013) provide a simple augmented ε-constraint method to generate all non-dominated solutions for a multi-objective integer programming problem. Kirlik and Sayın (2014) propose an algorithm to generate all non-dominated solutions for multi-objective discrete optimization problems with any number of objective functions. Jozefowiez et al. (2012) provide a generic branch-and-cut algorithm. Mavrotas (2009) and Mavrotas and Florios (2013) propose enhancements of the augmented ε-constraint method. The non-exact methods use metaheuristics (Yuan and Wang, 2009; Laumanns et al., 2006), approximations (see Köksalan and Lokman, 2009), greedy search algorithms (Özdamar and Wei, 2008; Chang et al., 2014), goal programming (Vitoriano et al., 2011; Li et al. 2012), and fuzzy multi-objective programming (Sheu, 2010) in order to find non-dominated solutions.

The work by You et al. (2012) is closely related our study. Different from You et al. (2012) who focus on analyzing the state of Illinois, this work focusses on large-scale (region-based) supply chain modeling and captures problem characteristics which become evident when one analyzes large-scale supply chains. Our modeling approach and solution methodology are substantially different.

3. Problem Description and Formulation

3.1 Supply Chain Structure for Biofuel Delivery

The proposed structure of the supply chain follows the Advanced Supply System concept proposed by the Idaho National Laboratory (INL) (2014). This system uses preprocessing of biomass to mitigate density and stability issues that prevent biomass from being handled in high-efficiency bulk dry solid or liquid distribution systems. Advanced supply system relies on densifying biomass at local preprocessing facilities before delivering to a biorefinery and before long distance transportation.

Figure 2 presents a supply chain consisting of four local preprocessing facilities, two depots, one biofuel plant, one terminal for biofuel blending and storage, and two customers. Preprocessing facilities are located at farms. These facilities deliver biomass to depots through truck shipments. If a preprocessing facility is located within 75 miles of a biofuel plant, it is assumed that the facility has the option of shipping directly to the biofuel plant bypassing the depots. This assumption is supported by studies that find truck transportation of biomass is not cost efficient beyond 50 miles (Brower, 2010). This transportation option is not made available to facilities located further away from a plant in order to reduce the problem size.

Depots are rail ramps (or ports) where truck shipments of biomass are consolidated. High-volume, long-haul shipments are delivered from depots to biofuel plants by rail (or barge). It is expected that a biofuel plant will have railway access to handle the large amount of biomass required to operate at high capacity. Thus, depots represent the first hubs and biofuel plants represent the second hubs in this supply chain. The final product, cellulosic ethanol, is shipped to a bulk terminal or a redistribution bulk terminal from where it is then delivered to customers. Bulk terminals are typically blending facilities where cellulosic ethanol is stored until it is blended with gasoline. Depending on the volume shipped and transportation distance either truck or rail is used for cellulosic ethanol delivery. Typically, rail is used for distances longer than 75 miles. From the bulk terminal, shipments of cellulosic ethanol are delivered by truck and in smaller quantities to gas stations.

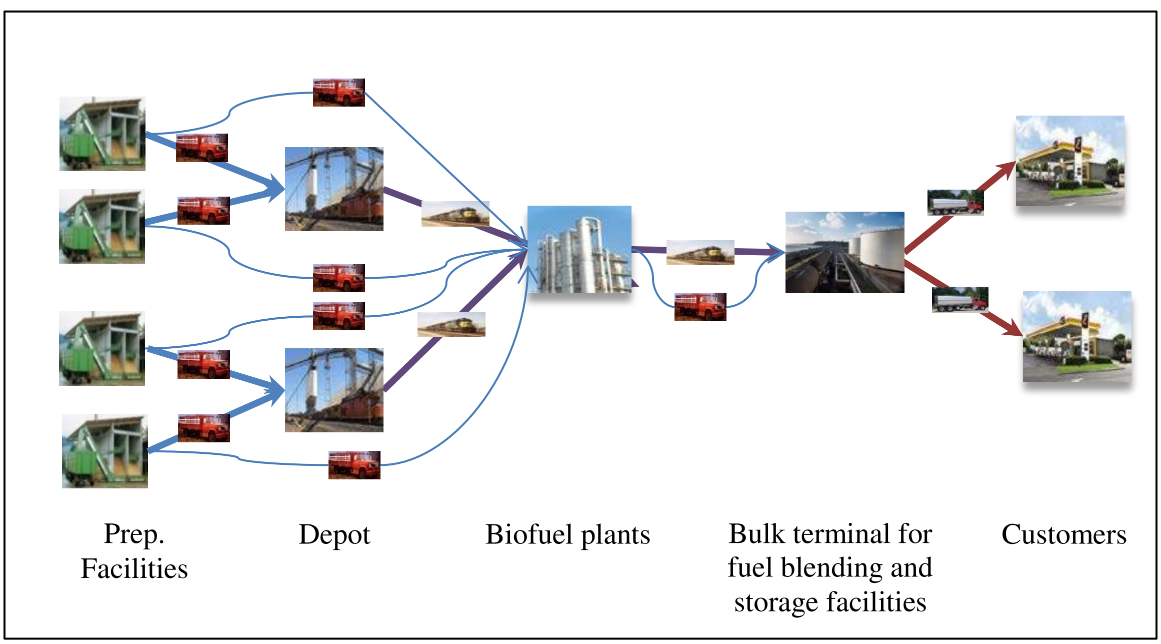


Figure 2: Supply chain network structure.

3.2 Model Formulation

We propose a mixed integer linear program (MILP) to model this supply chain design and management problem. This model is an extension of the facility location model since it identifies locations for depots, and biofuel plants based on information about investment costs, transportation costs, etc. Let $G(N, A)$ denote the supply chain network, where, N represents the set of nodes and A represents the set of arcs. Set N consists of subset P which represents the set of preprocessing facilities, subset D which represents the set of depot, subset B which represents the set of potential biofuel plant locations, subset L which represents set of bulk terminal locations and subset C which represents set of customers. Set A consists of subset T_1 which represents the set of arcs that connect preprocessing facilities to depot, T_2 which represents the set of arcs that connect preprocessing facilities to biofuel plant, subset T_3 which represents the set of arcs that connect biofuel plant to the bulk terminal, subset T_4 which represents the set of arcs that connect bulk terminal to the customer, subset R_1 which represents the set of arcs that connect depots to biofuel plants and subset R_2 which represents the set of arcs that connect biofuel plants to the bulk terminals. Let $T = \{T_1 \cup T_2 \cup T_3 \cup T_4\}$ and $R = \{R_1 \cup R_2\}$. The transportation mode used along arcs in T and R are truck and rail respectively.

Cost Objective:

The costs along arcs in T are linear, and there are no upper bounds on the amount shipped using these arcs. For truck transportation, we consider that a fixed cost (θ^T) occurs per mile and per ton shipped due to fuel consumption. Additionally, a fixed cost (ϑ^T) occurs per ton loaded/unloaded in the truck. Let d_{ij} denote the distance traveled along arc $(i, j) \in T$, then, transportation cost per ton shipped along this arc are equal to $c_{ij} = \vartheta^T + \theta^T * d_{ij}$. Let X_{ij} be the amount shipped along arc (i, j), then the total transportation cost along this arc is $f(X_{ij}) = c_{ij}X_{ij}$ (Searcy et al., 2007).

Total transportation cost along an arc in R is of a multiple-setup structure as described by Equation (1). In this equation, Ψ_{ij} is the fixed cost for loading/unloading a unit train, c_{ij} is the unit transportation cost per ton shipped along (i, j), v_{ij} is the capacity of a unit train (i, j), and n is the number of unit trains used (Roni, 2014b).

$$f(X_{ij}) = \begin{cases} 0 & if \quad X_{ij} = 0 \\ \Psi_{ij} + c_{ij}X_{ij} & if \quad 0 < X_{ij} \leq v_{ij} \\ 2 * \Psi_{ij} + c_{ij}X_{ij} & if \quad v_{ij} < X_{ij} \leq 2 * v_{ij} \\ \vdots & \\ n * \Psi_{ij} + c_{ij}X_{ij} & if \quad (n-1) * v_{ij} < X_{ij} \leq n * v_{ij} \end{cases} \qquad (1)$$

Equation (1) presents a piecewise linear cost function. In order to incorporate this function within the objective function of the MILP model presented below, we introduce integer variables Z_{ij}. These variables represent the number of unit trains moving along arc (i, j). Thus, $f_{ij}(X_{ij}) = \Psi_{ij}Z_{ij} + c_{ij}X_{ij}$. Total transportation costs in this supply chain are:

$$\mathrm{T_RC} = \sum_{(i,j)\in T} c_{ij}X_{ij} + \sum_{(i,j)\in R}(c_{ij}X_{ij} + \Psi_{ij}Z_{ij}) \quad (2)$$

Hub location costs represent the investment costs necessary to build the infrastructure in support of loading/unloading unit trains at a depot. Let W_i be a binary variable which takes the value 1 when node $i \in D$ is used as a depot, and takes the value 0 otherwise. Let ς_i be the fixed investment cost at node $i \in D$. Total hub location costs are HC = $\sum_{i\in D} \varsigma_i W_i$. Let ϱ_{ik} be the fixed investment costs to build a biofuel plant of capacity k ($k \in K$) at node $i \in B$. Let β_{ik} be a binary variable which takes the value 1 if node i is selected as biofuel plant location, and takes the value 0 otherwise. Total biofuel plant location costs are BC = $\sum_{k\epsilon K}\sum_{i\epsilon B} \varrho_{ik}\beta_{ik}$.

In this formulation we consider that the system is penalized for not meeting demand. Let π_i represent demand shortage and let α_i represent the corresponding penalty cost at customer i. Then, expression $\sum_{i\in C} \alpha_i \Pi_i$ represents the penalty for not meeting demand.

The cost objective function minimizes the total of transportation cost, hub location costs, and a penalty costs for unmet demand, and it is defined as follows:

$$minimize\text{: TC} = \sum_{(i,j)\in T} c_{ij}X_{ij} + \sum_{(i,j)\in R_1}(c_{ij}X_{ij} + \Psi_{ij}Z_{ij}) + \sum_{(i,j)\in R_2}(c_{ij}X_{ij} + \lambda_{ij}Y_{ij})$$
$$+ \sum_{i\in D} \varsigma_i W_i \; + \sum_{k\in k}\sum_{i\in B} \varrho_{ik}\beta_{ik} + \sum_{i\in C} \alpha_i \Pi_i$$

Environmental Objective

The model captures CO_2 emissions which result from fuel combustion due to transportation in the supply chain. The model also captures CO_2 emissions due to constructing and operating biofuel plants, and operating the hubs. We consider that the emission function is linear with respect to quantities shipped and quantities processed in facilities (Argo et al., 2013). Let e_{ij} represent CO_2 emission per ton per mile shipped along arc $(i,j) \in A$. Let ϵ_{ik} represents CO_2 emission per ton processed at the biofuel plant located in $i \in B$. Let μ_i represents CO_2 emission for establishing a hub in $i \in D$. The following environmental objective minimizes total emissions in the supply chain.

$$Minimize\text{: } TE = \sum_{(i,j)\in T,R} e_{ij}X_{ij} + \sum_{i\in D} \mu_i W_i + \sum_{k\in K}\sum_{i\in D} \epsilon_{ik}\,\beta_{ik} \quad (3)$$

Social Objective

The social benefits of this supply chain are measured by the number of accrued local jobs. Jobs are created to support biomass and biofuel transportation, biofuel plant construction and operation and hub operation. The number of transportation jobs created is linear and depends on the transportation distance, and quantity shipped. The number of job created due to biofuel plant construction and operation depends on the production capacity of the plant. The number of jobs created due to hub operation is fixed (NREL, 2013). Let p_{ij}^{T} represent the number of transportation jobs created, let p_i^D represent the number of job created due to hub operations, and let p_{ik}^{B} represent the number of job created due to construction and support operations of biofuel plant *i*. Then, the social objective function is defined as follows:

$$\max SB = \sum_{(i,j)\in T} p_{ij}^{T} X_{ij} + \sum_{(i,j)\in R_1} p_{ij}^{T} Z_{ij} + \sum_{(i,j)\in R_2} p_{ij}^{T} Y_{ij} + \sum_{i\in D} p_i^D W_i + \sum_{k\in K}\sum_{i\in D} p_{ik}^{B}\, \beta_{ik} \qquad (4)$$

The MILP Model

Table A.1 in Appendix A summarizes the parameters, and decision variables declared in this model. Next, we present the multi-objective MILP problem formulation. We refer to this as formulation (P).

$$Minimize: (TC(X, Z, Y, \beta, W, \Pi), TE(X, \beta, W, \Pi)) \qquad (P)$$

$$Maximize: (SB(X, Z, Y, \beta, W, \Pi))$$

Subject to:

$$\sum_{j\in D \cup B} X_{ij} \leq s_i \qquad \forall\, i \in P \qquad (5)$$

$$\sum_{i\in P} X_{ij} - \sum_{i\in B} X_{ji} = 0 \qquad \forall\, j \in D \qquad (6)$$

$$\sum_{i\in P U D} X_{ij} - \sum_{i\in L} X_{ji} = 0 \qquad \forall\, j \in B \qquad (7)$$

$$\sum_{i\in B} X_{ij} - \sum_{i\in C} X_{ji} = 0 \qquad \forall\, j \in L \qquad (8)$$

$$\sum_{i\in L} X_{ij} + \Pi_j = g_j \qquad \forall\, j \in C \qquad (9)$$

$$X_{ij} - v_{ij} Z_{ij} \leq 0 \qquad \forall\, (i,j) \in R_1 \qquad (10)$$

$$X_{ij} - \tau_{ij} Y_{ij} \leq 0 \qquad \forall\, (i,j) \in R_2 \qquad (11)$$

$$\sum_{i\in P} X_{ij} - u_j W_j \leq 0 \qquad \forall\, j \in D \qquad (12)$$

$$\sum_{j \in P \cup D} X_{ji} - \sum_{k \in K} q_{ik} \beta_{ik} \leq 0 \qquad \forall\, i \in B \tag{13}$$

$$\sum_{k \in K} \beta_{ik} \leq 1 \qquad \forall\, i \in B \tag{14}$$

$$X_{ij} \in R^n \qquad \forall\, (i,j) \in A \tag{15}$$

$$\pi_i \in R^n \qquad \forall\, i \in C \tag{16}$$

$$W_i \in \{0,1\}, \qquad \forall\, i \in D \tag{17}$$

$$\beta_{ik} \in \{0,1\}, \qquad \forall\, i \in B, k \in K \tag{18}$$

$$Z_{ij} \in Z^+ \qquad \forall\, (i,j) \in R_1 \tag{19}$$

$$Y_{ij} \in Z^+ \qquad \forall\, (i,j) \in R_2 \tag{20}$$

Constraints (5) indicate that the amount of biomass shipped from a preprocessing facility is limited by its availability. Constraints (6)-(8) are the flow balance constraints at depots, biofuel plants, and bulk terminals respectively. Constraints (9) indicate that customer demand could be satisfied through shipments from terminals or the market. These equations also measure demand shortage. Constraints (10) and (11) set an upper limit on the amount of biomass shipped using rail cars. Constraints (12) set a limit on the storage capacity of a hub. Constraints (13) set a limit on the capacity of a biorefinery. Constraints (14) set a limit on the number of biofuel plants at a particular location. Constraints (15) and (16) are the non-negativity constraints. Constraints (17) and (18) are binary constraints. Constraints (19) and (20) are the integrity constraints.

4. Solution Approach

In this section we describe the approach used in order to generate the set of Pareto optimal solution for our MILP problem. The set of Pareto optimal solutions is also known as the set of efficient, non-dominated, non-inferior solutions. These are solutions for which we cannot improve the value of one of the functions without deteriorating the performance of the rest of the objective functions. The two main approaches used in the literature to solve a multi objective problem are the weighted sum method and the ε- constraint method. Works (Mavrotas, 2009; Steuer, 1986; Miettinen, 1998) point out that the ε - constraint method is advantageous over the weighting sum method. This is mainly due to the fact that the ε-constrained method is computationally efficient. The ε- constraint method optimizes one of the objective functions. The remaining objectives are incorporated in the constraint set as shown below. We refer to this as formulation (Q).

$$\min TC(X, Z, Y, \beta, W, \Pi) \qquad \text{(Q)}$$

Subject to: (5)-(20)

$$TE(X, \beta, W, \Pi) \leq \varepsilon_1 \quad (21)$$

$$SB(X, \beta, Z, Y, W, \Pi) \geq \varepsilon_2 \quad (22)$$

The values of ε_1 and ε_2 are bounds set on the value of the environmental and social benefit objectives. Traditionally, the ε- constraint method requires identifying upper and lower bounds – in other words, defining a range - for each objective incorporated in the constraint set. Calculating these ranges for *TE* and *SB* is not a trivial task (Isermann and Steuer, 1987; Reeves and Reid, 1988; Steuer, 1997). Moreover, the optimal solution of formulation (Q) is guaranteed to be an efficient solution for (P) only if both constraints (21) and (22) are binding (Miettinen, 1998; Ehrgott and Wiecek, 2005). Otherwise, there is an alternative optimal solution to this problem, and the solution obtained from solving formulation (Q) is not efficient. Such a solution is a weakly efficient solution.

In this paper we apply a novel version of ε- constraint method known as the augmented ε-constraint method (Mavrotas and Florios, 2013; Mavrotas, 2009) in order to find the Pareto optimal solutions. In this method the ranges of ε_1 and ε_2 are calculated using the Lexicographic optimization method. The efficiency of the solution found is guaranteed since the reformulated ε- constraint model uses appropriate slack or surplus variables.

4.1 Lexicographic optimization to obtain the ranges of ε_1 and ε_2

The Lexicographic optimization method starts by ranking the objective functions based on their priority level. The function with highest priority makes the top of the list. In our problem, the total cost function has the highest priority, followed by the total emission and the social benefit functions. Next, based on the Lexicographic optimization method, we optimize the following 3 problems, and calculate corresponding objective function values. The 1st problem to optimize is: $minimize$: TC s.t. (5)-(20). The solution to this problem is $(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$, and the corresponding objective function value is $\boldsymbol{f_1^1} = TC(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$. The solution found is then used to evaluate the objective function values for the total emission ($\boldsymbol{f_2^1}$) and the social benefit ($\boldsymbol{f_3^1}$) functions. The 2nd problem optimized is: $minimize$: TE s.t (5)-(20) and the additional constraint $TC(X, Z, Y, \beta, W, \Pi) = \boldsymbol{f_1^1} + \delta_1$. Where δ_1 is a very small number. We increase the value of δ_1 from 0 to some small positive number in order to obtain a feasible solution to this problem. Adding this constraint guarantees that the new solution found optimizes *TE* while maintaining the value of the cost function (*TC*) at its lowest possible value. We denote this new solution by $(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \widetilde{\Pi})$. The corresponding objective function value is $\boldsymbol{f_2^2} = TE(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \widetilde{\Pi})$. The solution found is then used to calculate the objective function values for the total cost function ($\boldsymbol{f_1^2}$) and the social benefit function ($\boldsymbol{f_3^2}$). Finally, the 3rd problem optimized is: $Max\ SB(X, Z, Y, \beta, W, \Pi)$ s.t (5)-(20) and two additional constraints: $TC(X, Z, Y, \beta, W, \Pi) = \boldsymbol{f_1^2} + \delta_1$, and $TE(X, Z, Y, \beta, W, \Pi) = \boldsymbol{f_2^2} - \delta_2$. Where δ_2 is a very small positive number. We increase the values of δ_1 and δ_2 from 0 to some small positive numbers to obtain a feasible solution to this problem. We denote this new solution by $(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$. The corresponding objective function value is $\boldsymbol{f_3^3} = SB(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$. The solution found is then used to calculate the objective function values for the total cost function ($\boldsymbol{f_1^3}$) and the emission function ($\boldsymbol{f_2^3}$). At the end of implementing the Lexicographic optimization method we construct the payoff table shown in Table 1.

Let $s_1^{max} = max\,(\boldsymbol{f_2^1}, \boldsymbol{f_2^2}, \boldsymbol{f_2^3})$, $s_2^{max} = max\,(\boldsymbol{f_3^1}, \boldsymbol{f_3^2}, \boldsymbol{f_3^3})$, $s_1^{\min} = min\,(\boldsymbol{f_2^1}, \boldsymbol{f_2^2}, \boldsymbol{f_2^3})$, $s_2^{\min} = min\,(\boldsymbol{f_3^1}, \boldsymbol{f_3^2}, \boldsymbol{f_3^3})$. We use these values to create a range for the values that ε_1 and ε_2 can take during the optimization. We divide this interval into *k* equal subintervals in order to obtain good estimates on the values of ε_1 and ε_2. The benefit of using the Lexicographic optimization method is to identify a range of values that ε_1 and ε_2 can take. These values provide a dense representation of the efficient set.

Table 1: Payoff table generated by Lexicographic optimization method

Optimization Problems	**Objective function values for**		
	TC function	**TE function**	**SB function**
Problem 1	$\boldsymbol{min}: TC$ $s.to.$ (5)-(20) **Find**: $(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$ $\boldsymbol{f_1^1} = TC(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$	$\boldsymbol{f_2^1} = TE(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$	$\boldsymbol{f_3^1} = SB(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$.
Problem 2	$\boldsymbol{f_1^2} = TC(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \tilde{\Pi})$	$\boldsymbol{min}: TE$ $s.to.$ (5)-(20) $TC = \boldsymbol{f_1^1} + \delta_1$. **Find**: $(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \tilde{\Pi})$ $\boldsymbol{f_2^2} = TE(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \tilde{\Pi})$	$\boldsymbol{f_3^2} = SB(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \widetilde{W}, \tilde{\Pi})$.

Problem 3	$\boldsymbol{f_1^3} = TC(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$	$\boldsymbol{f_2^3} = TE(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$	$\boldsymbol{max}: SB$ $s.to.$ (5)-(20) $TC = \boldsymbol{f_1^2} + \delta_1.$ $TE = \boldsymbol{f_2^2} + \delta_2.$ **Find**: $(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$ $\boldsymbol{f_3^3} = SB(\bar{\bar{X}}, \bar{\bar{Z}}, \bar{\bar{Y}}, \bar{\bar{\beta}}, \bar{\bar{W}}, \bar{\bar{\Pi}})$

4.2 Reformulating the ε- constraint method with appropriate slack or surplus variable

We overcome the problem of generating weakly efficient solutions when using the ε- constraint method by incorporating the appropriate slack or surplus variables in the constraint set and in the objective function. Introducing these variables forces the algorithm to produce only efficient solutions. The new problem, which we call RMMILP is the following:

$$\min TC(X, Z, Y, \beta, W, \Pi) + \delta(S_1 + S_2)$$

$$\text{Subject to:} \qquad (5)\text{-}(20)$$

$$TE(X, Z, Y, \beta, W, \Pi) + S_1 = \varepsilon_1 \qquad (23)$$

$$\text{SB}(X, Z, Y, \beta, W, \Pi) - S_2 = \varepsilon_2 \qquad (24)$$

$$S_1, S_1 \in R^+ \qquad (25)$$

In the objective function, δ is an adequately small number. Typically, δ takes values between 10^-3 and 10^-6. This reformulation of the ε- constraint method avoids the generation of weakly efficient solutions (Mavrotas, 2009). We are now ready to present the procedure we develop to solve our multi-objective optimization problem using the augmented ε- constraint method. The procedure is shown in Figure 3.

Step 1	Build the payoff table (Table 1) using the Lexicographic optimization method Calculate the range of values for ε_1 and ε_2 using the payoff table Set number of intervals to *k* and compute step size by $\Delta\, \varepsilon_1 = \frac{\varepsilon_1^{max} - \varepsilon_1^{min}}{k}$, $\Delta\, \varepsilon_2 = \frac{\varepsilon_2^{max} - \varepsilon_2^{min}}{k}$ $\varepsilon_1 = \varepsilon_1^{max} - \Delta\, \varepsilon_1$ Set the Pareto optimal set $\Lambda = \emptyset$	
Step 2	**For** $i = 0$ to k **do** $\varepsilon_2 = \varepsilon_2^{min}$	
		For $j = 0$ k **do** Update the values of ε_1 and ε_2 in RMMILP

Solve RMMILP
If RMMILP feasible **Then**
Add solution to Λ
Else
Break
End If
$\varepsilon_2 = \varepsilon_2 + \Delta\, \varepsilon_2$
Next *j*
$\varepsilon_1 = \varepsilon_1 - \Delta\, \varepsilon_1$
Next *i*

Figure 3: A procedure for the augmented ε- constraint method

5. Data Collection for the Case Study

5.1 Biomass supply

Biomass availability data at the county level was extracted from the Knowledge Discovery Framework (KDF) database (2012), an outcome of the US Billion Ton Study led by the Oak Ridge National Laboratory. This data was further processed by INL to identify potential locations for preprocessing facilities and the corresponding amount of densified biomass available. This paper considers the biomass available on the following nine states, some located in the Midwest and some in the West of USA. The selected states are: Iowa, Nebraska, Kansas, South Dakota, California, New Mexico, Nevada, and Arizona. We focus our analysis in these states because they have substantial amounts of biomass available for biofuel production (such as, Iowa, Nebraska, Kansas and South Dakota) or are major users of biofuel (such as, California). The total number of counties considered in this study is 602. The primary biomass sources considered in this study are agricultural residue originated from primary crop such as corn, wheat, sorghum, oats, and barley.

5.2 Biofuel demand

We estimate the demand for biofuel at the county level. In order to estimate demand we investigated the size of population and gasoline consumption in each county. The data about population size is collected from the 2010 US Census (2010). The data about gasoline consumption is obtained from the Energy Information Administration (EIA) (2013).

5.3 Rail network data

The data about the US railway network structure was provided by Oak Ridge National Laboratory (2009). This database consists of 80,486 rail links, and 36,393 unique origin and destination nodes. Of the 36,393 nodes, only 20,686 are rail stations. The data set provides the following information for each rail link: origin, destination, length, ownership, terrain, number of main line tracks, main track authority (signal system), interval of passing sidings, speed limit, federal information processing standard state code (FIPS), and standard point location code (SPLC). Figure 4 summarizes the input data used. The figure lays out the distribution of available biomass and biofuel demand in the states we are investigating, and the corresponding rail network. Our model considers this network structure as given and does not suggest modifications to its structure.

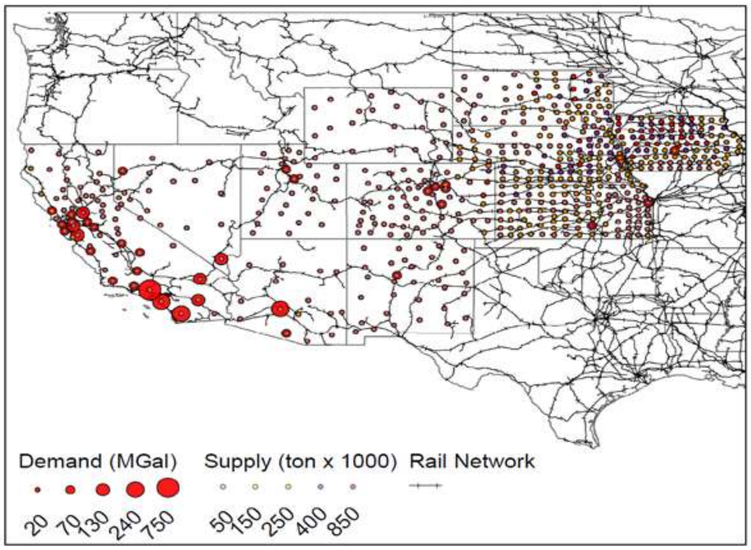


Figure 4. A summary of the input data

5.4 Transportation cost

Next we provide details about the structure of truck and rail transportation cost functions. Note that, we assume full-truck-load (FTL) shipments via truck or rail mainly because of the nature of the products delivered. Biofuel is a liquid and biomass is bulky, thus, we expect that a truck/rail car will be used for single-customer deliveries. To minimize the transportation costs, one would deliver FLT shipments.

5.4.1. Truck

In order to estimate the costs of biomass transportation using trucks we use data provided by Searcy et al. (2007). Searcy et al. (2007) provide two cost components, a distance variable cost (DVC) and a distance fixed cost (DFC). The distance variable cost includes fuel and labor costs. The distance fixed cost includes the cost of loading and unloading a truck. These costs were provided for different types of biomass, such as, woodchips, straw and stover. We used the data provided for woodchips since the physical properties of densified biomass are similar to woodchips. The DVC of woodchips is estimated \$0.112/ton-mile and DFC is estimated \$3.01/tons. Woodchips are shipped using truck with a capacity of

40 tons. Truck transportation costs of biofuel are estimated based on Searcy et al. (2007). Biofuel transportation is evaluated based on a tandem tanker carrying 40 tons of ethanol. The DVC of ethanol is estimated \$0.08/ton-mile and DFC is estimated \$3.86 /tons. This data is used as follows in order to calculate c_{ij} (in \$/ton) for $(i,j) \in T$: $c_{ij} = DFC + DVC * d_{ij}$. In this equation, d_{ij} represents the distance between locations *i* and *j*.

5.4.2 Unit train and single car shipment

The majority of freight transportation in the US is handled by four Class I railway companies. The two Class I railways that span the West USA are Burlington Northern Santa Fe Corporation (BNSF) and Union Pacific Railroad Company (UP) (CBO, 2006). Roni (2013) presents a regression analysis of rail transportation costs using rail waybill data; and uses this data to estimate the variable cost of transporting densified biomass and biofuel. The regression equations quantify the relationship between variable transportation unit cost (\$/ton) and car type, shipment size, rail movement type, commodity type, etc. Equations (26) and (27) are extracted from Roni (2013). These equations represent the relationship among variable unit cost (y) (in \$/ton), railway distance (x_1 given in miles) and car ownership (x_2) for received moves by BNSF and UP. Note that, x_2 is an indicator variable, which takes the value 1 if the railcar used is owned by the railway company, and takes the value 0 otherwise. The adjusted R^2 value for these regression equations is greater than 95% and *p*-values for the independent variables are less than 0.01%.

$$y_{BNSF} = -0.65 + 0.015x_1 + 1.96x_2 \quad (26)$$

$$y_{UP} = 0.78 + 0.0138x_1 + 3.78x_2 \quad (27)$$

Equations (26) and (27) assume that the type of rail car used is covered hopper and a single railway moves a shipment from its origin to its destination. The capacity of each rail car is 100 ton. The size of a unit train operated by BNSF is typically 100 cars. Since it is mainly BNSF that serves the states we consider in this analysis, we assume that a unit train is 100 cars long.

Equations (28) and (29) are used to estimate the variable unit cost for cellulosic ethanol for single car shipments. These equations assume that the type of rail car used is tank car with capacity over 22,000 gallons; the rail car is owned by the customer; and a single railway company moves the rail car from its origin to its destination.

$$y_{BNSF} = 6.40 + 0.0276x_1 \quad (28)$$

$$y_{UP} = 6.7174 + 0.0239x_1 \quad (29)$$

5.5 Hub investment costs

Only a few rail ramps are equipped to handle the loading and unloading of unit trains. In addition to equipment, there are certain infrastructural requirements necessary to handle unit trains. The infrastructure necessary is typically built by corn elevators, blenders, coal plants, or third-party logistics service providers.

In this study we consider that unit trains are loaded at rail ramps in case that the facilities exist. Otherwise, investments are required to build additional sidings. These investments are what we consider as hub location costs. Table 2 summarizes the typical costs which occur when building a railroad siding. We consider that one turnout and additional tracks are required. Since in this study we calculate annual costs of the supply chain, the annual equivalent for these investments is calculated and used. We assume the lifetime of such an investment is 30 years, and the discount factor is 10%.

Table 2: Costs related to railway sidings

Items	Costs
Track - rail and ties	$717.80/yard
Turnout - allows rail cars to switch tracks	$110,000.00

5.6 Biofuel plant investment costs

You et al. (2011) provide investment and operating cost for a 45 MGY ethanol productions plant that uses simultaneously scarification and fermentation technologies. They estimate the investment costs for build a biorefinery that produces 45 MGY of cellulosic ethanol are $159,400,000. Wallace et al. (2006) in his study estimates that doubling the size of a biofuel plant increases the investment costs by a factor of 1.6. We used this factor and interpolate investments costs in order to estimate investment costs for biofuel plants of different sizes. We use a 20 years project life and a 15% interest rate. The project life and interest rate is used to calculate the equivalent annual investment costs. In order to be consistent with the literature, and due to the availability of data, we consider 3 different biorefinery sizes: 60 MGY, 90 MGY and 120 MGY (Searcy and Flynn, 2008; Jacobson et al., 2014).

5.7 CO_2 emissions

Emissions due to rail and truck transportations are calculated using the following equation: CO_2 emissions (in kg) = (Transport volume by transport mode) * (Average transport distance by transport mode) *(Average CO_2-emission factor per ton-mile by transport mode). The average CO_2 emission factor recommended by the World Resource Institute and World Business Council for Sustainable Development for road transport operations is 0.297 kg/ton-mile. The average CO_2 emission factor recommended by the same organizations for rail transport operations is 0.0252 kg/ton-mile. The unit CO_2 emission from biofuel plant operations is provided by a study from Argo et al. (2013). This study shows that the average CO_2 emissions, due to the use of chemicals and enzymes in a biofuel plant, are 2.2 kg/gallon.

5.8 Social impact data

The number of accrued local jobs for biorefinery construction and operations is extracted from the Jobs and Economic Development Impact (JEDI) model developed by National Renewable Energy Laboratory (NREL, 2013). JEDI is a tool that estimates the economic impacts of constructing and operating power generation and biofuel plants at the local and state levels. Table B2 presents the number of jobs created due to biorefinery construction and operations as extracted from JEDI. Note that, the number of jobs created is a function of the plant size.

The number of job created in the trucking industry is estimated based on the travel distance and amount of biomass shipped annually. We assume that a truck can carry a maximum load of 40 tons of bulk solids, and 8,000 gallons of liquids. The average travel speed is assumed 40 miles/hour. Additionally, we assume a truck has 2 drivers; there are 40 working hours per week; and 50 weeks per year. Based on these assumptions, the number of miles traveled by one truck is (40 hours/week)*(50 weeks/year)*(40 miles/hour) = 80,000 miles/year. The number of ton-miles per truck is (80,000 miles/year)*(40 tons) = 3,200,000 tons-miles/year. Thus, the number of jobs created for ton-mile is (2 drivers)/(3,200,000 tons-miles/year). To calculate the number of trucking jobs per ton along arcs $(i,j) \in T_1 \cup T_2$ (p_{ij}^T) we multiply (2/3,200,000) with the distance of arc (i,j). We follow a similar approach to calculate p_{ij}^T for $(i,j) \in T_3 \cup T_4$.

We assume that each unit train requires two crews. The number of job openings in the railway industry is calculated based on the distance traveled in each route and the number of unit trains operating annually. We assume that two jobs per hub will be created in order to operate the hub.

5.9 Data pre-processing

In this section we describe three approaches we follow in order to reduce the size of the problem investigated without compromising the quality of the solutions found.

Typically, trucks would deliver biomass directly to the biofuel plant when travel distances are short. For this reason, we did add an arc between a preprocessing facility and a biofuel plant only when the distance between the two is 75 miles or less. Doing this reduced the number of arcs in the network, and consequently the problem size.

The data about the US railway network consists of 80,486 rail links, and 36,393 unique origin and destination nodes. Of the 36,393 nodes, only 20,686 are rail stations. Of the rails stations listed, 11,301 are operated by BNSF, CSXT, NS and UP. Of the 80,486 rail links, 72% of are shorter than 5 miles. Since a unit train is a dedicated train, it will follow a single path from shipment origin to its destination without being regrouped in rail ramps along the way. This is why the network structure between depots and biofuel plant is represented by a bipartite network (see Figure 2). Each arc of this bipartite network represents the shortest path between a depot and a biofuel plant. We calculated the shortest paths using the Dijkstra's algorithm (Ahuja et al., 1993).

Finally, when creating arcs between a biofuel plant and bulk terminals we examine the length of a path. If the length is less than 75 miles, then we create an arc $(i, j) \in T_3$; otherwise, we create an arc $(i, j) \in R_2$.

6. Experimental Results

The augmented ϵ-constraint algorithm is implemented using C++. The IBM CPLEX 12.5.1 Concert Technology is used to solve the MILP models. All tests were conducted on a desktop computer with Intel ® Core i7 3.1 GHz CPU and 32 GB memory limit, on a windows operating system.

6.1 Comparing the cost minimization and the multi-objective optimization models

In order to evaluate the performance of the models proposed in this paper we create three scenarios. Each scenario is generated based on the maximum allowable travel distance between a preprocessing facility and a depot (Table 3). In Scenarios 1, 2 and 3, the travel distance is 10, 30 and 50 miles respectively. That means, in Scenario 1, an arc is added between a particular preprocessing facility and a depot if the corresponding travel distance is less than or equal to 10 miles. Therefore, as we go from Scenario 1 to 3 the amount of biomass available to be shipped through the network increases. The motivation for creating these scenarios is the fact that deliveries to depots will be completed by trucks, and it is not economical to ship biomass to a depot if the transportation distance is longer than 30 miles.

Clearly, the number of integer variables and number of constraints varies with the three scenarios described. The largest problem we solved has a total of 212,320 continuous variables, 2,849 binary variables, 153,466 integer variables and 160,491 constraints. The running time to solve one problem was anywhere between 10-20min.

A set of metrics are used in order to compare the cost minimization model with the multi-objective model. On addition to the unit delivery cost of biofuel, emissions and number of jobs created, other important metrics are: amount of biomass delivered and total amount of biofuel produced; transportation mode used and transportation cost, number of biofuel plants built and hubs used. A summary of these metrics is provided in Tables 4-8. In order to identify which of the Pareto optimal solutions of the multiple-objective model to select for these tables, we followed this logic. Among the Pareto-optimal solutions generated we selected the one with highest number of jobs created, and then, among those solutions, we selected the one with the lowest emission levels.

Table 3 compares the two models based on cost, emissions, and number of jobs created. While the minimum cost model focuses on minimizing costs, the multi-objective model provides solutions which have a greater positive impact on the environment and create more jobs. The minimum cost model provides solutions that are 2.31% to 12.66% cheaper. The multiple-objective model provides solutions that create 449 – 1,186 more jobs, and reduce emissions by 13.78% to 25.48%.

Table 3: Model comparisons based on biomass delivery

Scenario	Cost Minimization Model			Multi-objective Model		
	Costs ($/gal)	Emissions (lbs/gal)	Jobs (nr)	Costs ($/gal)	Emissions (lbs/gal)	Jobs (nr)
1	3.38	7.28	4,068	3.87	6.57	5,000
2	3.39	7.68	4,322	3.47	6.51	5,508
3	3.28	6.54	3,751	3.55	6.25	4,200

Table 4 compares the two models based on the amount of biomass delivered by truck and rail. Hubs are used to facilitate rail transportation. The multi-objective model relies more on rail transportation. Emissions are smaller for this transportation mode due to the fact that in each trip, higher volumes of biomass and biofuel are delivered. To facilitate rail transportation more hubs are utilized.

Table 5 compares the two models based on the total delivery cost of biofuel. This cost consists of transportation, labor, and investment costs. The unit transportation costs are smaller for the multi-objective transportation since the model heavily relies on rail transportation. More hubs are utilized in order to minimize truck deliveries and increase access to rail. For this reason, labor and investment costs are higher, and consequently the total unit cost is higher.

Table 4: Model comparisons based on biomass delivery

Scenario	Available Biomass (in MT)	Cost Minimization Model			Multi-objective Model		
		Biomass Delivered (in MT)		Number of Hubs	Biomass Delivered (in MT)		Number of Hubs
		Truck	Rail		Truck	Rail	
1	52.99	18.11	3.71	20	4.99	15.10	80
2	62.92	19.04	3.24	13	3.82	17.14	135
3	63.45	16.42	6.79	18	4.16	16.03	101

Table 5: Model comparisons based on the delivery cost of cellulosic ethanol

	Cost Minimization Model	Multi-objective Model

Scenario	Transportation cost ($/gal)	Other costs ($/gal)	Total unit cost ($/gal)	Transportation cost ($/gal)	Other costs ($/gal)	Total unit cost ($/gal)
1	0.60	2.78	3.38	0.41	3.46	3.87
2	0.56	2.83	3.39	0.40	3.06	3.47
3	0.61	2.67	3.28	0.42	3.13	3.55

Table 6 summarizes the number of biofuel plants open and corresponding sizes, the total production capacity, the utilization rate of these plants, the biofuel production, and the percentage of RFS goals met under each scenario. These results are provided separately for each model (Tables 7(a) and 7(b)). The minimum cost model in order to minimize the total biofuel plant investment costs, and gain from the economies of scale that come with large production facilities, opens fewer biofuel plants, but of larger capacity. Consequently, transportation costs to these plants are higher. The multi-objective model opens smaller sized plants. This mode also invests in utilizing more hubs, therefore, investment costs are higher, more people are employed; however, transportation costs and emission levels are lower. Since maximizing biofuel production and meeting RFS goals was not an objective, the multi-objective model does not try to maximize utilization rates of plants.

Note that, the RFS goals set by EPA were reduced in 2014 below the volumes originally set by Congress (EPA, 2014). Based on the new goals, in 2014, only 33 MGY of cellulosic biofuel is expected to be produced. This number increases to 206 MGY by 2016. In 2015, the total RFS requirements are 15.93BGY. The percentages presented in Tables 7(a) and 7(b) are with respect to overall RFS requirements. Clearly, the requirements set on cellulosic biomass can be met at a unit cost between $3.5-4 per gallon.

Table 6: Model comparisons based on network design

Scenario	Nr. of biofuel plants open			Total capacity (MGY)	Utilization (%)	Biofuel production (MGY)	% of RFS goals met
	60 (MGY)	**90** (MGY)	**120** (MGY)				
1	2	2	7	1,140	95.69%	1,091	6.82%
2	1	2	8	1,200	92.86%	1,114	6.96%
3	1	3	7	1,170	99.21%	1,161	7.25%

Scenario	Nr. of biofuel plants open			Total capacity (MGY)	Utilization (%)	Biofuel production (MGY)	% of RFS goals met
	60 (MGY)	**90** (MGY)	**120** (MGY)				
1	12	1	3	1,170	85.81%	1,004	6.27%
2	9	1	4	1,110	94.42%	1,048	6.55%
3	5	5	3	1,110	90.91%	1,009	6.31%

(a) Cost minimization model

(b) Multi-objective optimization model

Table B1 in the Appendix lists the location of biorefineries for the cost minimization and multi-objective problems. Figures 5 and 6 present the network structure for the cost minimization and multi-objective models. These are the results from solving Scenario 3. Based on these results, biofuel plants are located closer to the supply, and therefore, in Iowa, Kansas, Nebraska, and South Dakota. Two biofuel plants are located in Colorado to be close to customers. Tables B1 and B2 in the appendix present the specific locations of biofuel plants and the number of jobs created in each state.

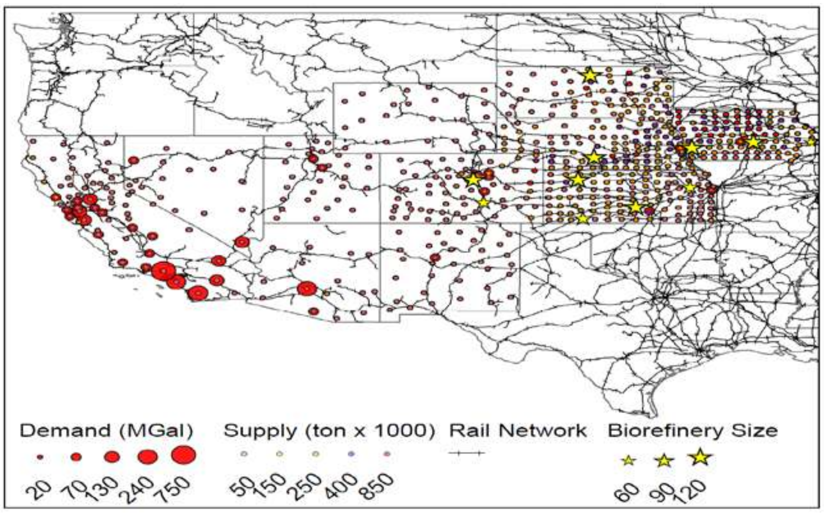


Figure 5: Network structure for Scenario 3 of cost minimization model

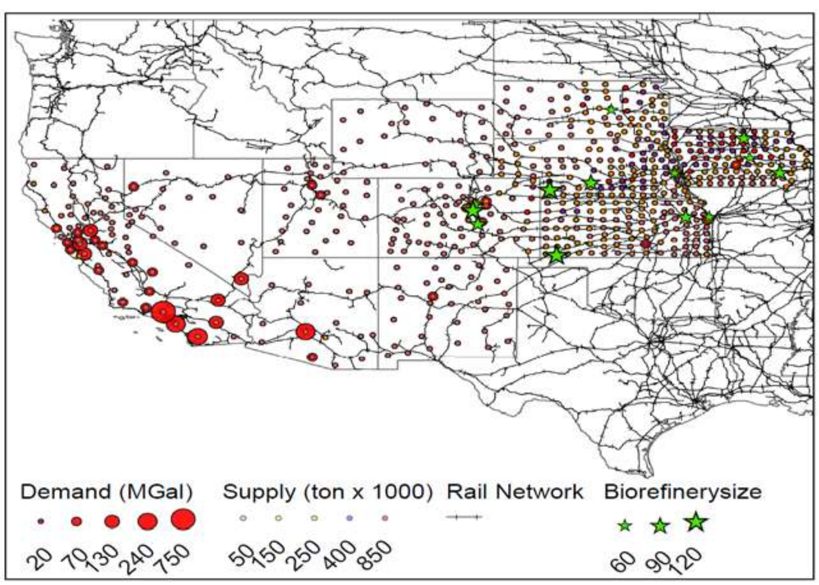


Figure 6: Network structure for Scenario 3 of multi-objective model

6.2 Pareto curve

The Pareto curves in Figures 7-9 present the tradeoffs that exist among economic, environmental and social objectives. It would be interesting to show the three-dimensional plots for the three objectives considered. However, creating three dimensional plots requires many points for the vectorization. As these three objectives are interrelated, we had to identify many weakly efficient solutions to create the three-dimensional plot. Therefore, we are presenting instead a number of two-dimensional Pareto optimal solutions. These two dimensional charts represent the tradeoffs between two of the three objectives which satisfy a threshold level set on the third objective.

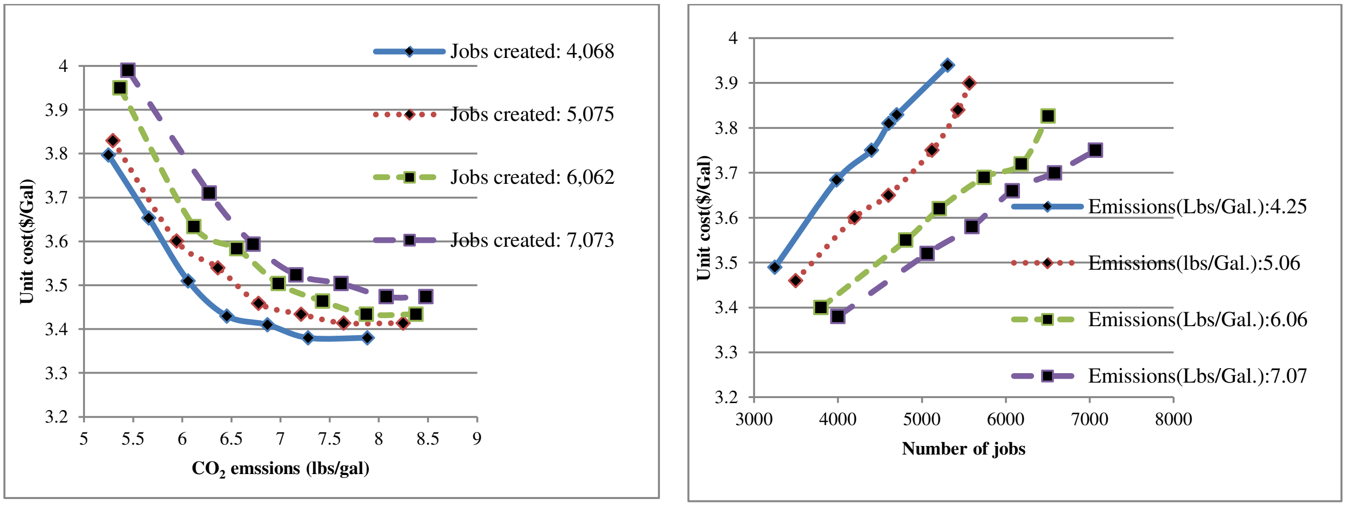
Figure 7(a) plots the relationship between the unit delivery for cost and CO_2 emissions for different levels of targeted number of jobs created under Scenario 1. Figure 7(b) plots the relationship between the unit delivery for cost and number of jobs created for different levels of targeted CO_2 emissions under Scenario 1. Similar plots for Scenario 2 are presented in Figures 8(a) and 8(b), and for Scenario 3, results are presented in Figures 9(a) and 9(b).

Results from these figures indicate that, for a given job target as the emission level decreases, delivery cost increases. These relationships are intuitive. To decrease emission levels, biofuel plants should reduce shipment volumes by truck. This requires investments to increase the number of hubs used and consequently improve accessibility to railway lines. Another observation is that: as the number of jobs increases, delivery cost increases as well. Increasing the number of jobs in this system affects labor costs and consequently the unit delivery cost of cellulosic ethanol.

The shape of the curves presented in Figures 7(a), 8(a) and 9(a) is similar and indicates a negative relationship between unit costs and unit emissions. That means, reducing CO_2 emissions from supply chain activities increases the cost of delivering biomass. However, the shape of the Pareto curve becomes flatter when emission levels are between 6 and 8 lbs/gal. That means, reducing CO_2 emissions from 8 to 6 lbs/gal (Figure 8a) increases the unit cost by 10 cents. The marginal increase in costs increases as emission reductions approach 4 lbs/gal. Reductions in emissions could be achieved via imposing an emission tax, setting an emission cap, etc. Clearly these policies would impact costs in the supply chain. However, it is often possible to have a great impact on emission reductions with only marginal increases in costs.

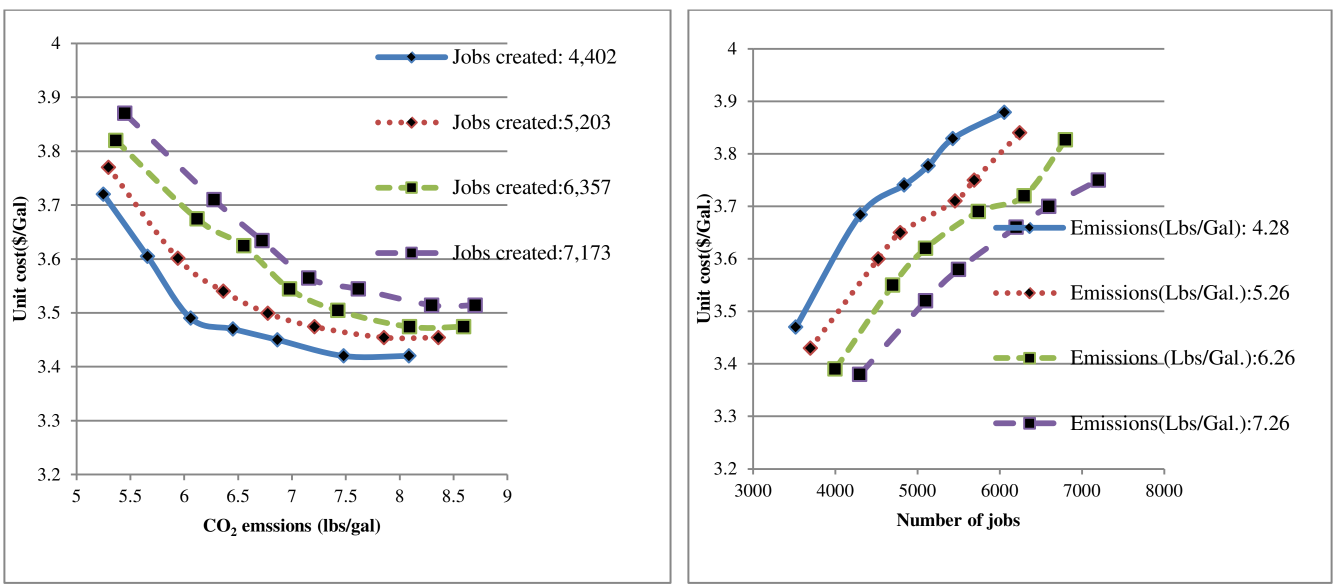
The results in Figures 7(a), 8(a) and 9(a) indicate that, in order to comply with increased restrictions on CO_2 emissions, plants need to rely on rail shipments. For this reason, at low emission levels more hubs are utilized and the investments on the infrastructure are higher. As emission levels increase, the restriction on emissions become redundant and do not have an effect on costs anymore. This is the reason why at high emission levels, increasing emissions does not affect the unit cost.

The results from Figures 7(b), 8(b), and 9(b) indicate a positive relationship between the number of jobs created and the unit cost. More jobs are created when truck - rather than rail - is used to deliver biomass. This is mainly because to ship the same amount of biomass, less railroad crew members are required as compared to truck drivers.



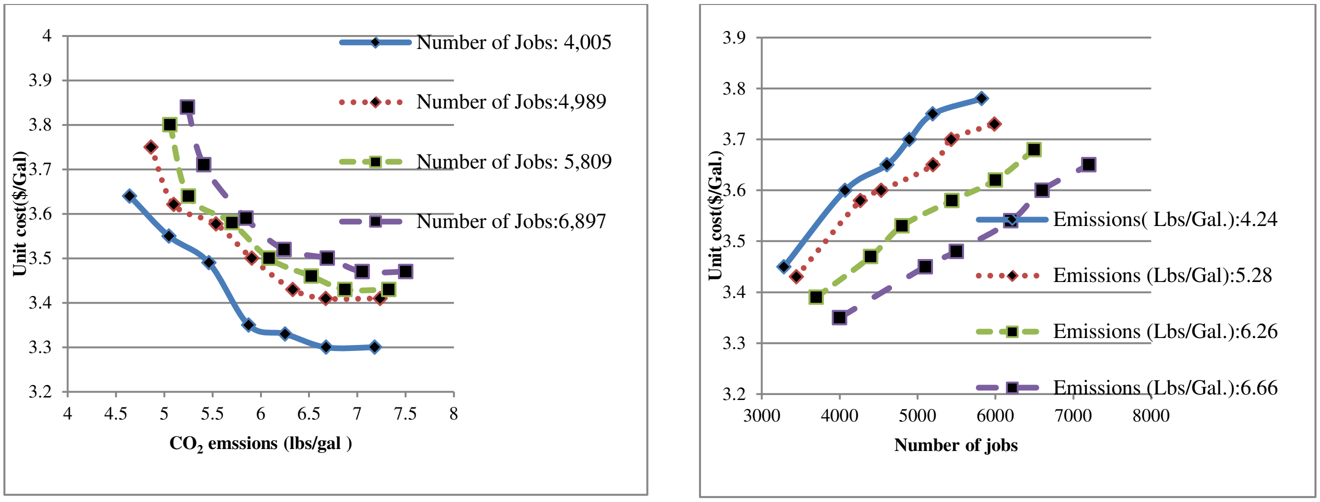
(a) (b)

Figure 7: Pareto curves for Scenario 1: (a) Unit delivery cost versus CO_2 emissions for different targeted number of job created; (b) Unit delivery cost versus number of job for particular emission target.



(a) (b)

Figure 8: Pareto curves for Scenario 2: (a) Unit delivery cost versus CO_2 emissions for different targeted number of job created; (b) Unit delivery cost versus number of job for particular emission target.



(a) (b)

Figure 9: Pareto curves for Scenario 3: (a) Unit delivery cost versus CO_2 emissions for different targeted number of job created; (b) Unit delivery cost versus number of job for particular emission target.

7. Conclusion

In this paper, we present a multi-objective optimization model for the cellulosic ethanol supply chain. The model optimizes costs, environmental, and social impacts of this supply chain. The cost objective represents transportation, facility location, and operations costs. The environmental objective represents CO_2 emissions due to transportation, facility construction, and operations. The social objective represents the number of new jobs created in order to handle transportation, hub operations, biofuel plant construction and operations. The multi-objective model is solved using an augmented ϵ-constraint method. This method identifies a set of Pareto optimal solutions. The relationship among the corresponding objectives is depicted through a number of graphs presented in the paper.

The underlying supply chain has a hub-and-spoke network structure. Such a network structure is appropriate for the delivery of bulk products, such as biomass, or cellulosic ethanol. In this network, depots serve as shipment consolidation points where small shipments of biomass from preprocessing facilities are consolidated into high-volume shipments. High-volume shipments of biomass are then delivered to biofuel plants by rail. Such a system positively impacts transportation costs, and consequently, the delivery cost of cellulosic ethanol, and CO_2 emissions. Using rail transportation, rather than truck, for high-volume and long-haul shipments reduces emissions.

The numerical analyses indicate that the goals set by the 2014 RFS for production of cellulosic biofuel can be met. The minimum cost model does minimize the delivery cost of cellulosic biofuel, but the multi-objective model has a greater positive impact on the environment and society. The minimum cost model invests on building large sized production plants to take advantage of the economies of scale that come with producing in large quantities. This model does not invest as much in building rail hubs, and relies on truck transportation. The multi-objective model proposes investments in building more small sized plants that employ additional workforce. The corresponding supply chain relies on rail transportation to reduce CO_2 emissions, and uses a larger number of hubs to enable the delivery of biomass.

We plan on extending the work presented in this paper. We are currently extending the scope of the case study by investigating the whole USA. Extending the scope of the case study will impact the problem size. We are developing decomposition-based algorithms to solve efficiently each single-objective optimization models within the algorithm scheme proposed here.

8. Acknowledgments

This work is partially supported by National Science Foundation, grant CMMI 1462420. This support is gratefully acknowledged.

9. Reference:

Abounacer,R., Rekik, M., Renaud, J. (2014). An exact solution approach for multi-objective location–transportation problem for disaster response. *Computers & Operations Research*, 41(1), 83-93.

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A. and Lukas, J. (2002). Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. *NREL/TP-510-32438*. Golden Co: National Renewable Energy Laboratory.

Ahuja, R.K., Magnanti, T.L., Orlin, J.B. (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall Publishers, Upper Saddle River, NJ, pp. 115–121.

Alumur, S., Kara, B.Y. (2008). Network hub location problems: the state of the art. *European Journal of Operational Research*, 190, 1–21.

Argo, M.R., Tan, E.C.D., Inman, D, Langholtz, M.H., Eaton, L.M., Jacobson, J.J, Wright, C.T., Muth, D.J., Wu, M.W., Chiu, Y.W., Graham, R.L. (2013). Investigation of biochemical biorefinery sizing and environmental sustainability impacts for conventional bale system and advanced uniform biomass logistics designs. *Journal of Biofuels, Bioproducts, and Biorefining*, 7(3), 282–302.

Bai, Y., Hwang, T., Kang, S., Ouyang, Y. (2011). Biofuel refinery location and supply chain planning under traffic congestion. *Transportation Research Part B: Methodological*, 45 (1), 162–175.

Brower, M., (2010). Woody biomass economics. Available on-line at: http://www.mosaicllc.com/documents/34.pdf

Camargo, R.S., Miranda Jr., G., Ferreira, R., Luna, H.P. (2009a). Multiple allocation hub and spoke network design under hub congestion. *Computers and Operations Research,* 36 (12), 3097–3106.

Campbell, J.F., O'Kelly, M.E., (2012). Twenty-five years of hub location research. *Transportation Science,* 46 (2), 153–169.

Congress of the United States, Congressional Budget Office (CBO) (2006). Freight Rail Transportation: Long-term Issues.

Chen, J. (2007). A hybrid heuristic for the uncapacitated single allocation hub location problem. *Omega*, 35 (2), 211–220.

Chen, C., Fan, Y. (2012). Bioethanol supply chain system planning under supply and demand uncertainties. *Transportation Research Part E: Logistics and Transportation Review*, 48 (1), 150–164.

Cundiff, J.S., Dias, N., Sherali, H.D. (1997). A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource Technology*, 59 (1), 47–55.

Cunha, C.B., Silva, M.R. (2007). A genetic algorithm for the problem of configuring a hub-and-spoke network for a LTL trucking company in Brazil. *European Journal of Operational Research*, 179 (3), 747–758.

Ehrgott M.,Wiecek M. (2005). Multiobjective programming. In: J. Figueira, S. Greco, M. Ehrgott (Eds.), *Multiple Criteria Decision Analysis. State of the Art Surveys*. Springer.

El-Halwagi, A.M., Rosas, C., Ponce-Ortega, J.M., Jiménez-Gutiérrez, A., Mannan, M. S., El-Halwagi, M. M. (2013). Multi-objective optimization of biorefineries with economic and safety objectives. *AIChE Journal*, 59: 2427–2434.

Eksioglu, S.D., Acharya, A., Leightley, L.E., Arora, S., (2009). Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers and Industrial Engineering*, 57 (4), 1342–1352.

Eksioglu, S.D., Zhang, S., Li, S., Sokhansanj, S., Petrolia, D. (2011). Analyzing the impact of intermodal facilities to the design of the supply chains for biorefineries. *Transportation Research Record, Journal of the Transportation Research Board*, 2191 (1), 144–151.

Gebreslassie, B.H., Yao, Y., You, F. (2012). Design under uncertainty of hydrocarbon biorefinery supply chains: multiobjective stochastic programming models, decomposition algorithm, and a comparison between CVaR and downside risk. *AIChE Journal*, 58 (7), 1547–5905.

Greenhouse Gas Protocol (2012). Emission Factors from Cross-Sector Tools. Available on line at: http://www.ghgprotocol.Org/calculation-tools/all-tools.

Gonzales, D., Searcy, E.M., Ekşioğlu, S.D. (2013). Cost analysis for high-volume and long-haul transportation of densified biomass feedstock. *Transportation Research Part A: Policy and Practice*, 49, 48-61

Hess, J.R., Wright, T.C., Kenney L.K., Searcy M.E. (2009). Uniform-Format solid feedstock supply system: A commodity-scale design to produce an infrastructure-compatible bulk solid from lignocellulosic biomass". INL/EXT-09-15423. Available on-line at: http://www.inl.gov/technicalpublications/Documents/ 4408280.pdf

Huang, Y., Chen, C.W., Fan, Y. (2010). Multistage optimization of the supply chains of biofuels. *Transportation Research Part E: Logistics and Transportation Review*, 46 (6), 820–830.

Idaho National Laboratory (2014). *Feedstock and Conversion Supply System Design and Analysis, The Feedstock Logistics Design Case*. Available on-line at: file:///C:/Users/seksiog/Downloads/Feedstock%20and%20Conversion%20Supply%20System%20Design%20and%20Analysis.pdf

Isermann, H, Steuer, R.E. (1987). Computational experience concerning payoff tables and minimum criterion values over the efficient set. *European Journal of Operational Research*, 33, 91–97.

Jacobson, J.J., Roni, M.S., Cafferty, K., Kenney, K., Searcy, E., Hansen, J.K. (2014). Biomass feedstock supply system design and analysis. *Report INL/EXT-14-33227*.

Jozefowiez, N., Laporte, G., Semet. F. (2012). A generic branch-and-cut algorithm for multi-objective optimization problems: Application to the multi-label traveling salesman problem. *INFORMS Journal on Computing*, 24, 554-564.

Kim, J., Realff, M.J., Lee, J.H., Whittaker, C., Furtner, L. (2011a). Design of biomass processing network for biofuel production using an MILP model. *Biomass and Bioenergy*, 35 (2), 853–871.

Kim, J., Realff, M.J., Lee, J.H. (2011b). Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty. *Computers & Chemical Engineering*, 35 (9), 1738–1751.

Kirlik, G., Sayın, S. (2014). A new algorithm for generating all nondominated solutions of multiobjective discrete optimization problems. *European Journal of Operational Research*, 232(3), 479-488.

Kumar, A.; Sokhansanj, S. (2007). Switchgrass (Panicum vigratum, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresource Technology*, 98(5), 1033-1044.

Labbe, M., Yaman, H. (2004). Projecting flow variables for hub location problems. *Networks*, 44 (2), 84–93.

Li, J., Burke, E. K., Curtois, T., Petrovic, S., Qu, R. (2012). The falling tide algorithm: A new multi-objective approach for complex workforce scheduling. *Omega*, 40(3), 283-293.

Mahmudi, H., Flynn, P. (2006). Rail vs. truck transport of biomass. *Applied Biochemistry and Biotechnology*, 129 (1), 88–103.

Marufuzzaman M., Eksioglu S.D., Hernandez R. (2014). Environmentally Friendly Supply Chain Planning and Design for Biodiesel Production via Wastewater Sludge. *Transportation Science*, 48(4), 555-574.

Mavrotas, G. (2009). Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. *Applied Mathematics and Computation*, 213(2), 455-465.

Mavrotas, G., Florios, K. (2013). An improved version of the augmented ε-constraint method (AUGMECON2) for finding the exact pareto set in multi-objective integer programming problems. *Applied Mathematics and Computation*, 219(18), 9652-9669.

Mele, F.D, Guille´n-Gosa´lbez, G., Jime´nez, L. (2009). Optimal planning of supply chains for bioethanol and sugar production with economic and environmental concerns. *Comput Aided Chem Eng*, 26, 997–1002.

Miettinen, K.M. (1998). *Nonlinear multiobjective optimization*. Kluwer Academic Publishers, Boston.

National Renewable Energy Laboratory (2013). Job and economic development impact model. Available on-line at: http://www.nrel.gov/analysis/jedi/.

Oak Ridge National Laboratory, Center for Transportation Analysis (2009). Oak Ridge National Laboratory Rail Network Analysis. Available on-line at: http://www-cta.ornl.gov/transnet/RailRoads.html

Papapostolou, C., Kondili, E., Kaldellis, J.K. (2011). Development and implementation of an optimization model for biofuels supply chain. *Energy*, 36 (10), 6019–6026.

Perimenis, A., Walimwipi, H., Zinoviev, S., Müller-Langer, F., Miertus, S. (2011). Development of a decision support tool for the assessment of biofuels. *Energy Policy*, 39(3), 1782-1793

Parker, N., Tittmann, P., Hart, Q., Nelson, R., Skog, K., Schmidt, Edward A.G., Jenkins, B., (2010). Development of a biorefinery optimized biofuel supply curve for the Western United States. *Biomass and Bioenergy*, 34 (11), 1597–1607.

Reeves, G.R., Reid, R.C.(1988). Minimum values over the efficient set in multiple objective decision making. *European Journal of Operational Research*, 36, 334–338.

Roni, M, Eksioglu, S., Searcy, E., Jha, K. (2014a). A supply chain network design model for biomass cofiring in coal-fired power plants. *Transportation Research Part E: Logistics and Transportation Review*, 61,115-134.

Roni, M. (2013). Analyzing the impact of a hub and spoke supply chain design for long-haul, high-volume transportation of densified biomass. Mississippi State University. Ph.D. Dissertation.

Roni, M. S., Eksioglu, S. D., Searcy, E., Jacobson, J. J. (2014b). Estimating the variable cost for high-volume and long-haul transportation of densified biomass and biofuel. *Transportation Research Part D: Transport and Environment*, *29*, 40-55.

Ren, L., Cafferty, K., Roni, Jacobson, J., Xie, G., Ovard, L., Wright, C. (2015). Analyzing and comparing biomass feedstock supply systems in China: Corn Stover and Sweet Sorghum Case Studies. *Energies* (Accepted).

Searcy, E., Flynn, P. (2008). The impact of biomass availability and processing cost on optimum size and processing technology selection. *Applied Biochemistry and Biotechnology*, 154, 271–286.

Searcy, E., Flynn, P., Ghafoori, E., Kumar A. (2007) The Relative Cost of Biomass Energy Transport. *Applied Biochemistry and Biotechnology*, 136–140.

Sheu, J. B. (2010). Dynamic relief-demand management for emergency logistics operations under large-scale disasters. *Transportation research part E: logistics and transportation review*, 46(1), 1-17.

Silva, M.R., Cunga, C.B., 2009. New simple and efficient heuristics for the uncapacitated single allocation hub location problem. *Computers and Operations Research* 36 (12), 3152–3165

Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M. E. (1996). Tight linear programming relaxations of uncapacitated p-hub median problems. *European Journal of Operational Research*, 94(3), 582-593.

Sokhansonj, S., Kumar, A., Turhollow, A. (2006). Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy*, 30(10), 838-847.

Steuer R.E. (1986). *Multiple Criteria Optimization. Theory, Computation and Application*. 2nd edition, John Wiley and Sons Inc.

Steuer, R.E. (1997). Non-fully resolved questions about the efficient/nondominated set. In: J. Climaco (Ed.), *Multicriteria Analysis*, Springer.

Tunc, H., Eksioglu, B., Eksioglu, S.D., Jin, M. (2011). Hub-based network design: a review. *International Journal of Networking*, 1 (2), 17–24.

United States Census Data. Available on line at: http://www.census.gov/2010census/.

United States Department of Agriculture (USDA) and United States Department of Energy (2008). National Biofuels Action Plan. Biomass Research and Development Board.

United States Department of Energy (2006). Office of the biomass program: Multi-year analysis plan FY04-FY08. Available on-line at: http://www.nrel.gov/docs/fy05osti/36999.pdf.

United States Department of Energy: Office of Energy Efficiency and Renewable Energy (2007). Biomass Program Multi-Year Program Plan 2010.

United States Department of Energy (2011). U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry. R.D. Perlack and B.J. Stokes (Leads), ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN. 227p.

United States Department of Energy (2012). Bioenergy: Knowledge Discovery Framework. Available on-line at: https://www.bioenergykdf.net/.

United States Energy Information Administration (2010). Monthly Energy Review. Available on-line at: http://tonto.eia.doe.gov/FTPROOT/multifuel/mer/00351005.pdf

United States Energy Information Administration (2013). Annual Energy Outlook: Energy Markets Summary. Available on-line at: http://www.eia.gov/forecasts/steo/tables/pdf/1tab.pdf.

United States Environmental Protection Agency (EPA) (2014). EPA Proposes 2014 Renewable Fuel Standards, 2015 Biomass-Based Diesel Volume. Available on-line at: http://www.epa.gov/otaq/fuels/renewablefuels/documents/420f13048.pdf

United States Environmental Protection Agency (EPA) (2007). Energy Independence and Security Act of 2007. In: *RL34294, P.L.110–140 ed. 2007*.

United States Environmental Protection Agency (EPA) (2015). Renewable Fuel Standard: Proposed Renewable Fuel Volumes for Calendar Years 2014, 2015, and 2016. Available on-line at: http://www.epa.gov/otaq/fuels/renewablefuels/documents/420f15028.pdf

You, F., Tao, L., Graziano, D. J., Snyder, S. W. (2012). Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input–output analysis. *AIChE Journal*, 58(4), 1547-5905.

You, F., Wang, B. (2011). Life-cycle optimization of biomass-to-liquid supply chains with distributed–centralized processing networks. *Industrial & Engineering Chemistry Research*, 50(17), 10102-10127.

Wallace, R., Ibsen, K., McAloon, A., Yee, W. (2005). Feasibility study for co-locating and integrating ethanol production plants from corn starch and lignocellulose feedstocks. NREL/TP-510-37092. USDA/USDOE/NREL, Revised Edition. Golden, Colorado.

Zamboni, A., Bezzo, F., Shah, N. (2009). Spatially explicit static model for the strategic design of future bioethanol production systems. 2. Multi-objective environmental optimization. *Energy & Fuels*, 23 (10), 5134-5143.

Zhang, W., Reimann, M. (2014). A simple augmented< i>∊</i>-constraint method for multi-objective mathematical integer programming problems. *European Journal of Operational Research*, 234(1), 15-24.

10. Appendix

APPENDIX A

Table A1: The definitions of sets, parameters and decision variables used

Set definitions	
N	set of nodes in supply chain network $G(N, A)$
P	set of preprocessing facilities
D	set of hub
B	Set of biorefinery locations
L	Set of bulk terminal
C	set of customers
A	set of arcs in G(N, A)
T_1	set of arcs that connect preprocessing facilities to hub
T_2	set of arcs that connect preprocessing facilities to the biorefinery
T_3	set of arcs that connect biorefinery facilities to the blending facilities
T_4	set of arcs that connect blending facilities to the customer
R_1	set of rail arcs that connect depots to biofuel plants
R_2	set of rail arcs that connect biofuel plants to the bulk terminals
K	set of biofuel plant capacity level indexed by k
Problem Parameters	
c_{ij}	unit cost charged per ton shipped along $(i, j) \in A$
d_{ij}	distance of $(i, j) \in A$
Ψ_{ij}	reflects a fixed cost for loading/unloading a unit train $(i, j) \in R_1$
λ_{ij}	reflects a fixed cost for loading/unloading a unit train $(i, j) \in R_2$
v_{ij}	represents the maximum capacity of a unit train along arc $(i, j) \in R_1$
τ_{ij}	represents the maximum capacity of a rail car along arc $(i, j) \in R_2$
ς_i	fixed investment cost at node $i \in D$
u_i	Capacity of node $i \in D$

Table A1 (Continued)

ϱ_{ik}	the fixed investment cost at node $i \in B$ with capacity $k \in K$
s_i	supply of biomass at a pre-processing facility $i \in P$
g_i	demand of biomass at a customer location $i \in C$
α_i	shortage cost at customer location $i \in C$
q_{ik}	capacity of biorefinery node $i \in B$ is $k \in K$
Emission parameters	
e_{ij}	CO_2 emission per ton per mile in arc set$(i, j) \in T_1, T_2, T_3, R$
ϵ_{ik}	CO_2 emission from biorefinery $i \in B$ with capacity $k \in K$
o_i	CO_2 emission for establishing a hub at node $i \in D$
Social factors	
p_{ij}^T	Number of jobs created per ton due to transportation activities in arc $(i, j) \in A$
p_{ik}^B	Number of job created for biorefinery $i \in B$ with capacity $k \in K$
p_i^D	Number of job created due to locating depot $i \in D$
Decision variables	
X_{ij}	flow along arc $(i, j) \in A$
Z_{ij}	number of unit trains moving from hub *i* to biorefinery *j*
Y_{ij}	number of single care moving from biorefinery *i* to bulk terminal *j*
W_i	a binary variable which takes the value 1 if *i* is used as a hub, and 0 O/W
β_{ik}	a binary variable which takes the value 1 if *i* is used as a biorefinery, with capacity *k* and 0 O/W
Π_i	demand shortage at customer location $i \in C$

APPENDIX B

Table B1: Biorefinery locations

Cost Minimization Model				Multi-objective Model			
State	**SPLC**	**City**	**Capacity (MGY)**	**State**	**SPLC**	**City**	**Capacity (MGY)**
CO	746413	Blakeland	120	CO	744149	Roydale	60
CO	748538	Southern JCT	90	CO	746453	Sedalia	120
IA	536640	Newton	120	CO	746689	Crews	90
IA	534553	Eldridge	60	IA	533370	Burchinal	90
IA	549256	McClelland	120	IA	536244	Minerva JCT	60
KS	592634	Selden	120	IA	537370	Washington	90
KS	584261	Menoken	90	KS	581577	Muncie	60
KS	589156	Partridge	120	KS	584261	Menoken	90
KS	598754	Meade	90	KS	599754	Hugoton	120
NE	555973	Darr	120	NE	553346	Elkhorn	60
SD	522530	Selby	120	NE	555973	Darr	90
				NE	559550	Imperial	120
				SD	525160	Miller	60

Table B2: Number of Job Created due to Construction and Operations of a Biorefinery

State	**Plant size (MGY)**	**Nr. of construction jobs**	**Nr. of operation jobs**	**State**	**Plant size (MGY)**	**Nr. of construction jobs**	**Nr. of operation jobs**
KS	60	89	137	CO	60	92	171
	90	112	170		90	116	220
	120	143	186		120	148	250
NE	60	93	150	UT	60	106	172
	90	118	187		90	133	217
	120	150	207		120	170	242
IA	60	91	148	NM	60	93	160
	90	115	186		90	117	214
	120	147	205		120	149	230
SD	60	98	157	WY	60	76	134
	90	124	197		90	96	168
	120	158	218		120	123	186
CA	60	86	188	NV	60	79	148
	90	109	246		90	100	186

<table>
<tr><td></td><td>120</td><td>139</td><td>286</td><td rowspan="4"></td><td rowspan="4">120</td><td rowspan="4">128</td><td rowspan="4">207</td></tr>
<tr><td rowspan="3">AZ</td><td>60</td><td>98</td><td>191</td></tr>
<tr><td>90</td><td>123</td><td>248</td></tr>
<tr><td>120</td><td>157</td><td>286</td></tr>
</table>

849