

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](mailto:seksiog@clemson.edu)*

Kara G. Cafferty  
*Idaho National Laboratory*

Jacob J. Jacobson  
*Minds Eye Computing LLC*

Follow this and additional works at: [https://tigerprints.clemson.edu/industrialeng\\_pubs](https://tigerprints.clemson.edu/industrialeng_pubs)

 Part of the [Engineering Commons](#)

## Recommended Citation

Please use publisher's recommended citation: <https://link.springer.com/article/10.1007/s10479-015-2102-3/fulltext.html>

This Article is brought to you for free and open access by the Industrial Engineering at TigerPrints. It has been accepted for inclusion in Publications by an authorized administrator of TigerPrints. For more information, please contact [kokeefe@clemson.edu](mailto:kokeefe@clemson.edu).

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

*Mohammad. S. Roni<sup>b</sup>, Sandra D. Eksioglu<sup>a\*</sup>, Kara G. Cafferty<sup>b</sup>, Jacob J. Jacobson<sup>c</sup>*

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:**

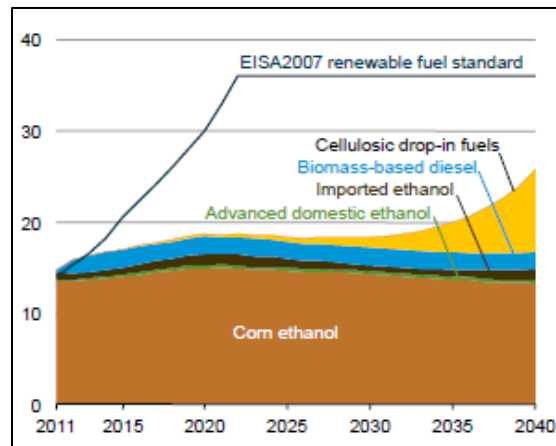
In this paper we propose a multi-objective, mixed integer linear programming model to design and manage the supply chain for biofuels. This model captures the trade-offs that exist between costs, environmental and social impacts of delivering biofuels. The in-bound supply chain for biofuel plants relies on a hub-and-spoke structure which optimizes transportation costs of biomass. The model proposed optimizes the CO<sub>2</sub> 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  $\epsilon$ -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  $\epsilon$ -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,

39 2010). The proposed 2014 production volume for cellulosic biofuel is 17 million gallons a year (MGY),  
40 and the proposed range is 8 – 30 MGY (EPA, 2014). Due to policies, such as RFS, it is expected that the  
41 share of biomass in the total renewable energy production will increase in the near future.  
42



43

44

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

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

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 2<sup>nd</sup> 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

64 makes sense since this is a new industry, thus, there is an opportunity here to do things right from the very  
65 beginning. Another contribution of this paper is providing a model that captures the environmental  
66 impacts of biofuels by estimating CO<sub>2</sub> emission due to transportation, biorefinery location, and  
67 biorefinery operations. The model also captures the social impacts of biofuels by estimating the number  
68 of jobs created due to biomass production, preprocessing, transportation, and biorefinery operating.

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

82 A contribution of this paper is the development of a case study which was developed using a  
83 number of reliable data sources (see Section 5). Thus, the results from the numerical analysis are very  
84 insightful. The results provide estimates of the delivery cost of cellulosic ethanol, unit emissions due to  
85 supply chain activities, and the number of new jobs created in this industry. The relationships revealed  
86 provide insights which help policy makers design policies that support renewable energy production.  
87 Finally, the mathematical model we propose is a challenging multi-objective linear mixed integer  
88 programming (MILP) model. We used an augmented  $\epsilon$ -constraint method to solve this multi-objective  
89 problem and generate a set of Pareto optimal solutions. We use lexicographic optimization to obtain the  
90 ranges of  $\epsilon_1$  and  $\epsilon_2$ . Doing this provides us with better estimates of the Pareto frontiers.

## 91 **2. Relevant literature**

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

96 The biomass supply chain optimization literature presents a number of deterministic and stochastic  
97 models. The deterministic models are extensions of the facility location model. These models are used to  
98 identify biorefinery sittings (Ekşioğlu et al., 2009; Parker et al., 210; Bai et al., 2011; Kim et al., 2011a;  
99 Papapostolou et al., 2011; Roni et al., 2014a; Marufuzzaman et al., 2014). Some deterministic models are  
100 used to identify the number, capacity and location of biofuel plants in order to make use of the available  
101 biomass in a particular region in a cost efficient manner. The stochastic research on biomass supply  
102 chains uses extensions of the two-stage, location-transportation stochastic programming model to identify  
103 biorefinery sittings (such as, Cundiff et al., 1997; Huang et al., 2010; Kim et al., 2011b; Chen and Fan,  
104 2012; Gebreslassie et al., 2012).

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

110 The hub-and-spoke design problem is conventionally called the hub location problem (Campbell,  
111 2012). A number of extensions of the hub location problem are found in the literature. These extensions  
112 are proposed in order to capture issues that arise when managing this supply chain, such as, non-linear  
113 economies of scale, traffic management, transportation mode selection, and congestion. The existing  
114 literature can be divided into two major groups, the single hub (SH) and the multiple hubs (MH) location  
115 problem. In a SH location model, the routing of the flow to/from a non-hub node is done through the hub.  
116 In a MH setting, the routing of the flow to/from a non-hub node is done through multiple hubs. Thus,  
117 flow initiated from a non-hub node traverses a number of hubs before reaching its final destination.  
118 Mixed integer programs (MIP) are used to model the problem to represent the fixed hub location costs,  
119 and nodes-to-hub allocations (Skorin-Kapov et al., 1996; Campbell, 2012). Due to computational  
120 challenges faced when solving these large sized MIP models, a number of different heuristic approaches  
121 have been design to solve the problems. For example, Chen (2007) developed a hybrid Simulated  
122 Annealing heuristics, Silva and Cunha (2009) developed a number of Tabu Search heuristics, Cunha and  
123 Silva (2007) developed a hybrid Genetic Algorithm and Simulated Annealing-based heuristics, Camargo  
124 et al. (2008) present a Benders Decomposition-based solution approach and Labbe and Yaman (2004)  
125 propose a Lagrangean Relaxation-based approach. For an extensive review of this problem see Alumur  
126 and Kara (2008), Tunc et al. (2011).

127 A limited number of papers in the literature propose multi-objective optimization models for the  
128 biofuel supply chain design and management. For example, Zamboni et al. (2009) present a MILP model  
129 that simultaneously minimizes the supply chain operating costs and GHG emissions due to supply chain

130 activities. Perimenis et al. (2011) provide a decision support tool to evaluate biofuel production pathways.  
131 This tool integrates technical, economic, environmental and social aspects along the entire value chain of  
132 biofuels starting from biomass production to biofuel end-use. Mele et al. (2009) address the problem of  
133 optimizing the supply chains for bioethanol and sugar production. Their bi-criteria MILP model addresses  
134 economic and environmental concerns. The model minimizes the total cost of managing the supply chain  
135 network, and minimizes the environmental impact over the entire product life cycle. El-Halwagi et al.  
136 (2013) incorporate safety concerns into the biorefinery location selection and capacity management  
137 problem. They establish tradeoffs between costs and safety issues using Pareto curves. You and Wang  
138 (2011) study the optimal design and planning of biomass-to-liquids (BTL) supply chains under economic  
139 and environmental criteria. You et al. (2012) address the optimal design and planning of cellulosic  
140 ethanol supply chains under economic, environmental, and social objectives.

141 Multi-objective integer linear programs have been solved using exact and heuristics solution  
142 approaches. An exact algorithm identifies the whole set of non-dominated solutions for the problem.  
143 Heuristics approximate, identify bounds for the set of non-dominated solutions. For example, Abounacer  
144 et al. (2014) propose an  $\varepsilon$ -constraint method to generate an exact Pareto frontier of a complex three  
145 objective location-transportation problem. The following is a list of exact methods. Zhang and Reimann  
146 (2013) provide a simple augmented  $\varepsilon$ -constraint method to generate all non-dominated solutions for a  
147 multi-objective integer programming problem. Kirlik and Sayin (2014) propose an algorithm to generate  
148 all non-dominated solutions for multi-objective discrete optimization problems with any number of  
149 objective functions. Jozefowicz et al. (2012) provide a generic branch-and-cut algorithm. Mavrotas  
150 (2009) and Mavrotas and Florios (2013) propose enhancements of the augmented  $\varepsilon$ -constraint method.  
151 The non-exact methods use metaheuristics (Yuan and Wang, 2009; Laumanns et al., 2006),  
152 approximations (see Köksalan and Lokman, 2009), greedy search algorithms (Özdamar and Wei, 2008;  
153 Chang et al., 2014), goal programming (Vitoriano et al., 2011; Li et al. 2012), and fuzzy multi-objective  
154 programming (Sheu, 2010) in order to find non-dominated solutions.

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

### 159 **3. Problem Description and Formulation**

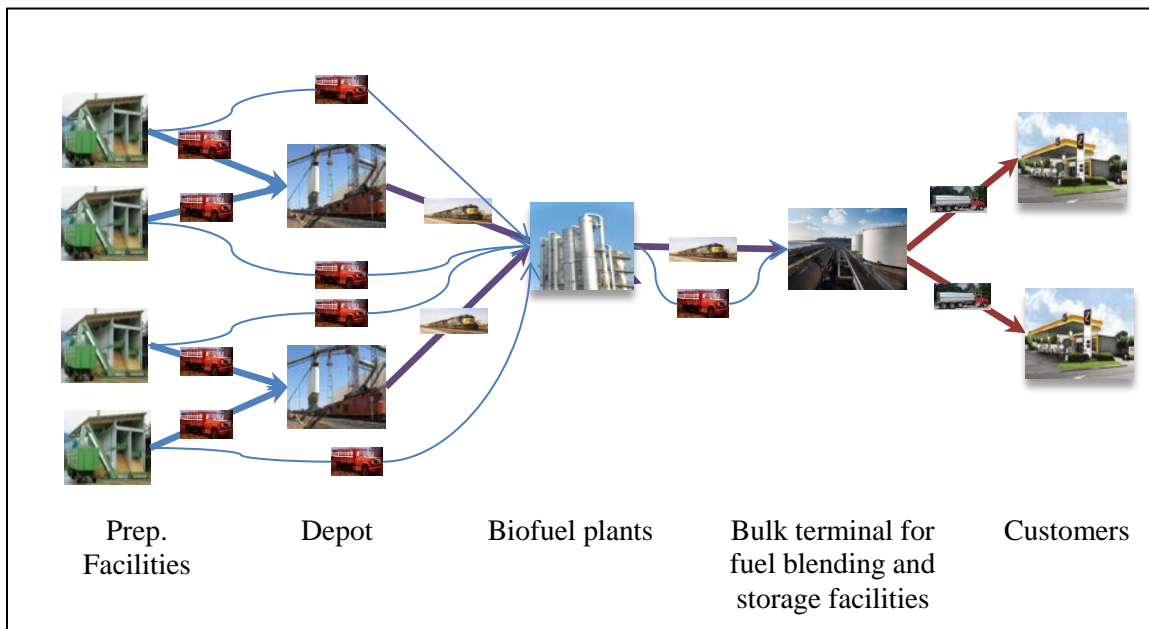
#### 160 **3.1 Supply Chain Structure for Biofuel Delivery**

161 The proposed structure of the supply chain follows the Advanced Supply System concept proposed by the  
162 Idaho National Laboratory (INL) (2014). This system uses preprocessing of biomass to mitigate density

163 and stability issues that prevent biomass from being handled in high-efficiency bulk dry solid or liquid  
164 distribution systems. Advanced supply system relies on densifying biomass at local preprocessing  
165 facilities before delivering to a biorefinery and before long distance transportation.

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

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



185

**Figure 2:** Supply chain network structure.186 **3.2 Model Formulation**

187 We propose a mixed integer linear program (MILP) to model this supply chain design and  
 188 management problem. This model is an extension of the facility location model since it identifies  
 189 locations for depots, and biofuel plants based on information about investment costs, transportation costs,  
 190 etc. Let  $G(N, A)$  denote the supply chain network, where,  $N$  represents the set of nodes and  $A$  represents  
 191 the set of arcs. Set  $N$  consists of subset  $P$  which represents the set of preprocessing facilities, subset  $D$   
 192 which represents the set of depot, subset  $B$  which represents the set of potential biofuel plant locations,  
 193 subset  $L$  which represents set of bulk terminal locations and subset  $C$  which represents set of customers.  
 194 Set  $A$  consists of subset  $T_1$  which represents the set of arcs that connect preprocessing facilities to depot,  
 195  $T_2$  which represents the set of arcs that connect preprocessing facilities to biofuel plant, subset  $T_3$  which  
 196 represents the set of arcs that connect biofuel plant to the bulk terminal, subset  $T_4$  which represents the set  
 197 of arcs that connect bulk terminal to the customer, subset  $R_1$  which represents the set of arcs that connect  
 198 depots to biofuel plants and subset  $R_2$  which represents the set of arcs that connect biofuel plants to the  
 199 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  
 200 arcs in  $T$  and  $R$  are truck and rail respectively.

201 **Cost Objective:**

202 The costs along arcs in  $T$  are linear, and there are no upper bounds on the amount shipped using  
 203 these arcs. For truck transportation, we consider that a fixed cost ( $\theta^T$ ) occurs per mile and per ton shipped  
 204 due to fuel consumption. Additionally, a fixed cost ( $\vartheta^T$ ) occurs per ton loaded/unloaded in the truck. Let  
 205  $d_{ij}$  denote the distance traveled along arc  $(i, j) \in T$ , then, transportation cost per ton shipped along this arc  
 206 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  
 207 transportation cost along this arc is  $f(X_{ij}) = c_{ij}X_{ij}$  (Searcy et al., 2007).

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

$$212 \quad f(X_{ij}) = \begin{cases} 0 & \text{if } X_{ij} = 0 \\ \Psi_{ij} + c_{ij}X_{ij} & \text{if } 0 < X_{ij} \leq v_{ij} \\ 2 * \Psi_{ij} + c_{ij}X_{ij} & \text{if } v_{ij} < X_{ij} \leq 2 * v_{ij} \\ \vdots & \vdots \\ n * \Psi_{ij} + c_{ij}X_{ij} & \text{if } (n - 1) * v_{ij} < X_{ij} \leq n * v_{ij} \end{cases} \quad (1)$$



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

$$217 \quad \text{TRC} = \sum_{(i,j) \in T} c_{ij}X_{ij} + \sum_{(i,j) \in R} (c_{ij}X_{ij} + \Psi_{ij}Z_{ij}) \quad (2)$$

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

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

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

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

## 230 Environmental Objective

231 The model captures CO<sub>2</sub> emissions which result from fuel combustion due to transportation in the  
 232 supply chain. The model also captures CO<sub>2</sub> emissions due to constructing and operating biofuel plants,  
 233 and operating the hubs. We consider that the emission function is linear with respect to quantities shipped  
 234 and quantities processed in facilities (Argo et al., 2013). Let  $e_{ij}$  represent CO<sub>2</sub> emission per ton per mile  
 235 shipped along arc  $(i, j) \in A$ . Let  $\epsilon_{ik}$  represents CO<sub>2</sub> emission per ton processed at the biofuel plant located  
 236 in  $i \in B$ . Let  $\mu_i$  represents CO<sub>2</sub> emission for establishing a hub in  $i \in D$ . The following environmental  
 237 objective minimizes total emissions in the supply chain.

$$238 \quad \text{Minimize: TE} = \sum_{(i,j) \in T, R} e_{ij}X_{ij} + \sum_{i \in D} \mu_i W_i + \sum_{k \in K} \sum_{i \in B} \epsilon_{ik} \beta_{ik} \quad (3)$$

239

240 **Social Objective**

241 The social benefits of this supply chain are measured by the number of accrued local jobs. Jobs are  
 242 created to support biomass and biofuel transportation, biofuel plant construction and operation and hub  
 243 operation. The number of transportation jobs created is linear and depends on the transportation distance,  
 244 and quantity shipped. The number of job created due to biofuel plant construction and operation depends  
 245 on the production capacity of the plant. The number of jobs created due to hub operation is fixed (NREL,  
 246 2013). Let  $p_{ij}^T$  represent the number of transportation jobs created, let  $p_i^D$  represent the number of job  
 247 created due to hub operations, and let  $p_{ik}^B$  represent the number of job created due to construction and  
 248 support operations of biofuel plant  $i$ . Then, the social objective function is defined as follows:

249 
$$\max SB = \sum_{(i,j) \in T} p_{ij}^T X_{ij} + \sum_{(i,j) \in R_1} p_{ij}^T Z_{ij} + \sum_{(i,j) \in R_2} p_{ij}^T Y_{ij} + \sum_{i \in D} p_i^D W_i + \sum_{k \in K} \sum_{i \in D} p_{ik}^B \beta_{ik} \quad (4)$$

250

251 **The MILP Model**

252

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

$$\text{Minimize: } (TC(X, Z, Y, \beta, W, \Pi), TE(X, \beta, W, \Pi)) \quad (P)$$

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

Subject to:

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

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

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

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

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

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

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

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

$$\begin{array}{llll}
\sum_{j \in P \cup D} X_{ji} - \sum_{k \in K} q_{ik} \beta_{ik} \leq 0 & \forall i \in B & (5) & \\
\sum_{k \in K} \beta_{ik} \leq 1 & \forall i \in B & (14) & \\
X_{ij} \in R^n & \forall (i, j) \in A & (15) & \\
\pi_i \in R^n & \forall i \in C & (16) & \\
W_i \in \{0,1\}, & \forall i \in D & (17) & \\
\beta_{ik} \in \{0,1\}, & \forall i \in B, k \in K & (18) & \\
Z_{ij} \in Z^+ & \forall (i, j) \in R_1 & (19) & \\
Y_{ij} \in Z^+ & \forall (i, j) \in R_2 & (20) & 
\end{array}$$

Constraints (5) indicate

265 that the amount of biomass shipped from a preprocessing facility is limited by its availability. Constraints  
266 (6)-(8) are the flow balance constraints at depots, biofuel plants, and bulk terminals respectively.  
267 Constraints (9) indicate that customer demand could be satisfied through shipments from terminals or the  
268 market. These equations also measure demand shortage. Constraints (10) and (11) set an upper limit on  
269 the amount of biomass shipped using rail cars. Constraints (12) set a limit on the storage capacity of a  
270 hub. Constraints (13) set a limit on the capacity of a biorefinery. Constraints (14) set a limit on the  
271 number of biofuel plants at a particular location. Constraints (15) and (16) are the non-negativity  
272 constraints. Constraints (17) and (18) are binary constraints. Constraints (19) and (20) are the integrity  
273 constraints.

#### 274 4. Solution Approach

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

$$\begin{array}{ll}
\min TC(X, Z, Y, \beta, W, \Pi) & (Q) \\
\text{Subject to:} & (5)-(20)
\end{array}$$

285 
$$TE(X, \beta, W, \Pi) \leq \varepsilon_1 \quad (21)$$

286 
$$SB(X, \beta, Z, Y, W, \Pi) \geq \varepsilon_2 \quad (22)$$

287  
288  
289

290 The values of  $\varepsilon_1$  and  $\varepsilon_2$  are bounds set on the value of the environmental and social benefit  
291 objectives. Traditionally, the  $\varepsilon$ - constraint method requires identifying upper and lower bounds – in other  
292 words, defining a range - for each objective incorporated in the constraint set. Calculating these ranges for  
293  $TE$  and  $SB$  is not a trivial task (Isermann and Steuer, 1987; Reeves and Reid, 1988; Steuer, 1997).  
294 Moreover, the optimal solution of formulation (Q) is guaranteed to be an efficient solution for (P) only if  
295 both constraints (21) and (22) are binding (Miettinen, 1998; Ehrgott and Wiecek, 2005). Otherwise, there  
296 is an alternative optimal solution to this problem, and the solution obtained from solving formulation (Q)  
297 is not efficient. Such a solution is a weakly efficient solution.

298 In this paper we apply a novel version of  $\varepsilon$ - constraint method known as the augmented  $\varepsilon$ -  
299 constraint method (Mavrotas and Florios, 2013; Mavrotas, 2009) in order to find the Pareto optimal  
300 solutions. In this method the ranges of  $\varepsilon_1$  and  $\varepsilon_2$  are calculated using the Lexicographic optimization  
301 method. The efficiency of the solution found is guaranteed since the reformulated  $\varepsilon$ - constraint model  
302 uses appropriate slack or surplus variables.

#### 303 **4.1 Lexicographic optimization to obtain the ranges of $\varepsilon_1$ and $\varepsilon_2$**

304 The Lexicographic optimization method starts by ranking the objective functions based on their  
305 priority level. The function with highest priority makes the top of the list. In our problem, the total cost  
306 function has the highest priority, followed by the total emission and the social benefit functions. Next,  
307 based on the Lexicographic optimization method, we optimize the following 3 problems, and calculate  
308 corresponding objective function values. The 1<sup>st</sup> problem to optimize is: *minimize: TC* s.t. (5)-(20). The  
309 solution to this problem is  $(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$ , and the corresponding objective function value is  
310  $f_1^1 = TC(X^*, Z^*, Y^*, \beta^*, W^*, \Pi^*)$ . The solution found is then used to evaluate the objective function values  
311 for the total emission ( $f_2^1$ ) and the social benefit ( $f_3^1$ ) functions. The 2<sup>nd</sup> problem optimized is:  
312 *minimize: TE* s.t (5)-(20) and the additional constraint  $TC(X, Z, Y, \beta, W, \Pi) = f_1^1 + \delta_1$ . Where  $\delta_1$  is a  
313 very small number. We increase the value of  $\delta_1$  from 0 to some small positive number in order to obtain a  
314 feasible solution to this problem. Adding this constraint guarantees that the new solution found optimizes  
315  $TE$  while maintaining the value of the cost function ( $TC$ ) at its lowest possible value. We denote this new  
316 solution by  $(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \tilde{W}, \tilde{\Pi})$ . The corresponding objective function value is  
317  $f_2^2 = TE(\tilde{X}, \tilde{Z}, \tilde{Y}, \tilde{\beta}, \tilde{W}, \tilde{\Pi})$ . The solution found is then used to calculate the objective function values for the

318 total cost function ( $f_1^2$ ) and the social benefit function ( $f_3^2$ ). Finally, the 3<sup>rd</sup> problem optimized is:  
 319  $Max SB(X, Z, Y, \beta, W, \Pi)$  s.t (5)-(20) and two additional constraints:  $TC(X, Z, Y, \beta, W, \Pi) = f_1^2 + \delta_1$ , and  
 320  $TE(X, Z, Y, \beta, W, \Pi) = f_2^2 - \delta_2$ . Where  $\delta_2$  is a very small positive number. We increase the values of  $\delta_1$   
 321 and  $\delta_2$  from 0 to some small positive numbers to obtain a feasible solution to this problem. We denote  
 322 this new solution by  $(\bar{X}, \bar{Z}, \bar{Y}, \bar{\beta}, \bar{W}, \bar{\Pi})$ . The corresponding objective function value is  
 323  $f_3^3 = SB(\bar{X}, \bar{Z}, \bar{Y}, \bar{\beta}, \bar{W}, \bar{\Pi})$ . The solution found is then used to calculate the objective function values for the  
 324 total cost function ( $f_1^3$ ) and the emission function ( $f_2^3$ ). At the end of implementing the Lexicographic  
 325 optimization method we construct the payoff table shown in Table 1.

326 Let  $s_1^{max} = \max (f_2^1, f_2^2, f_2^3)$ ,  $s_2^{max} = \max (f_3^1, f_3^2, f_3^3)$ ,  $s_1^{min} = \min (f_2^1, f_2^2, f_2^3)$ ,  $s_2^{min} =$   
 327  $\min (f_3^1, f_3^2, f_3^3)$ . We use these values to create a range for the values that  $\varepsilon_1$  and  $\varepsilon_2$  can take during the  
 328 optimization. We divide this interval into  $k$  equal subintervals in order to obtain good estimates on the  
 329 values of  $\varepsilon_1$  and  $\varepsilon_2$ . The benefit of using the Lexicographic optimization method is to identify a range of  
 330 values that  $\varepsilon_1$  and  $\varepsilon_2$  can take. These values provide a dense representation of the efficient set.

331

332 **Table 1:** Payoff table generated by Lexicographic optimization method

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

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

333

#### 334 4.2 Reformulating the $\varepsilon$ - constraint method with appropriate slack or surplus variable

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

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

Subject to: (5)-(20)

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

$$SB(X, Z, Y, \beta, W, \Pi) - S_2 = \varepsilon_2 \quad (24)$$

$$S_1, S_2 \in R^+ \quad (25)$$

339

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

345

<b>Step 1</b>	Build the payoff table (Table 1) using the Lexicographic optimization method Calculate the range of values for $\varepsilon_1$ and $\varepsilon_2$ 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 $\Lambda = \emptyset$
<b>Step 2</b>	<b>For</b> $i = 0$ to $k$ <b>do</b> $\varepsilon_2 = \varepsilon_2^{min}$
	<b>For</b> $j = 0$ to $k$ <b>do</b> Update the values of $\varepsilon_1$ and $\varepsilon_2$ in RMMILP

	Solve RMMILP <b>If</b> RMMILP feasible <b>Then</b> Add solution to $\Lambda$ <b>Else</b> <b>Break</b> <b>End If</b> $\varepsilon_2 = \varepsilon_2 + \Delta \varepsilon_2$ <b>Next j</b>
	$\varepsilon_1 = \varepsilon_1 - \Delta \varepsilon_1$ <b>Next i</b>

346 **Figure 3:** A procedure for the augmented  $\varepsilon$ - constraint method

## 347 **5. Data Collection for the Case Study**

### 348 **5.1 Biomass supply**

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

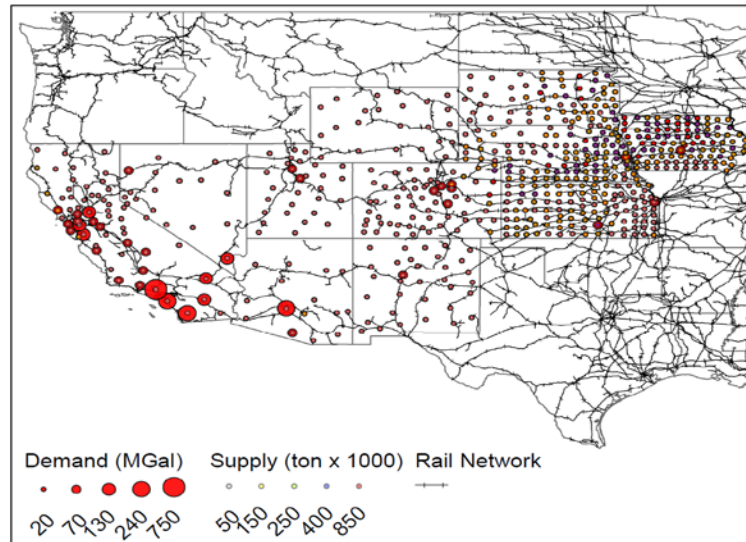
### 360 **5.2 Biofuel demand**

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

### 365 **5.3 Rail network data**

366 The data about the US railway network structure was provided by Oak Ridge National Laboratory  
 367 (2009). This database consists of 80,486 rail links, and 36,393 unique origin and destination nodes. Of the  
 368 36,393 nodes, only 20,686 are rail stations. The data set provides the following information for each rail

369 link: origin, destination, length, ownership, terrain, number of main line tracks, main track authority  
370 (signal system), interval of passing sidings, speed limit, federal information processing standard state  
371 code (FIPS), and standard point location code (SPLC). Figure 4 summarizes the input data used. The  
372 figure lays out the distribution of available biomass and biofuel demand in the states we are investigating,  
373 and the corresponding rail network. Our model considers this network structure as given and does not  
374 suggest modifications to its structure.  
375



376  
377

**Figure 4.** A summary of the input data

#### 378 **5.4 Transportation cost**

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

##### 383 **5.4.1. Truck**

384 In order to estimate the costs of biomass transportation using trucks we use data provided by Searcy  
385 et al. (2007). Searcy et al. (2007) provide two cost components, a distance variable cost (DVC) and a  
386 distance fixed cost (DFC). The distance variable cost includes fuel and labor costs. The distance fixed  
387 cost includes the cost of loading and unloading a truck. These costs were provided for different types of  
388 biomass, such as, woodchips, straw and stover. We used the data provided for woodchips since the  
389 physical properties of densified biomass are similar to woodchips. The DVC of woodchips is estimated  
390 \$0.112/ton-mile and DFC is estimated \$3.01/tons. Woodchips are shipped using truck with a capacity of



391 40 tons. Truck transportation costs of biofuel are estimated based on Searcy et al. (2007). Biofuel  
 392 transportation is evaluated based on a tandem tanker carrying 40 tons of ethanol. The DVC of ethanol is  
 393 estimated \$0.08/ton-mile and DFC is estimated \$3.86 /tons. This data is used as follows in order to  
 394 calculate  $c_{ij}$  (in \$/ton) for  $(i,j) \in T$ :  $c_{ij} = DFC + DVC * d_{ij}$ . In this equation,  $d_{ij}$  represents the distance  
 395 between locations  $i$  and  $j$ .

#### 396 5.4.2 Unit train and single car shipment

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

$$408 \quad y_{BNSF} = -0.65 + 0.015x_1 + 1.96x_2 \quad (26)$$

$$409 \quad y_{UP} = 0.78 + 0.0138x_1 + 3.78x_2 \quad (27)$$

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

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

$$418 \quad y_{BNSF} = 6.40 + 0.0276x_1 \quad (28)$$

$$419 \quad y_{UP} = 6.7174 + 0.0239x_1 \quad (29)$$

420 **5.5 Hub investment costs**

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

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

431

432

**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

433

434 **5.6 Biofuel plant investment costs**

435 You et al. (2011) provide investment and operating cost for a 45 MGY ethanol productions plant  
436 that uses simultaneously scarification and fermentation technologies. They estimate the investment costs  
437 for build a biorefinery that produces 45 MGY of cellulosic ethanol are \$159,400,000. Wallace et al.  
438 (2006) in his study estimates that doubling the size of a biofuel plant increases the investment costs by a  
439 factor of 1.6. We used this factor and interpolate investments costs in order to estimate investment costs  
440 for biofuel plants of different sizes. We use a 20 years project life and a 15% interest rate. The project life  
441 and interest rate is used to calculate the equivalent annual investment costs. In order to be consistent with  
442 the literature, and due to the availability of data, we consider 3 different biorefinery sizes: 60 MGY, 90  
443 MGY and 120 MGY (Searcy and Flynn, 2008; Jacobson et al., 2014).

444 **5.7 CO<sub>2</sub> emissions**

445 Emissions due to rail and truck transportations are calculated using the following equation: CO<sub>2</sub>  
446 emissions (in kg) = (Transport volume by transport mode) \* (Average transport distance by transport  
447 mode) \*(Average CO<sub>2</sub>-emission factor per ton-mile by transport mode). The average CO<sub>2</sub> emission factor  
448 recommended by the World Resource Institute and World Business Council for Sustainable Development  
449 for road transport operations is 0.297 kg/ton-mile. The average CO<sub>2</sub> emission factor recommended by the

450 same organizations for rail transport operations is 0.0252 kg/ton-mile. The unit CO<sub>2</sub> emission from  
451 biofuel plant operations is provided by a study from Argo et al. (2013). This study shows that the average  
452 CO<sub>2</sub> emissions, due to the use of chemicals and enzymes in a biofuel plant, are 2.2 kg/gallon.

## 453 **5.8 Social impact data**

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

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

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

## 473 **5.9 Data pre-processing**

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

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

480 The data about the US railway network consists of 80,486 rail links, and 36,393 unique origin and  
481 destination nodes. Of the 36,393 nodes, only 20,686 are rail stations. Of the rails stations listed, 11,301

482 are operated by BNSF, CSXT, NS and UP. Of the 80,486 rail links, 72% of are shorter than 5 miles.  
483 Since a unit train is a dedicated train, it will follow a single path from shipment origin to its destination  
484 without being regrouped in rail ramps along the way. This is why the network structure between depots  
485 and biofuel plant is represented by a bipartite network (see Figure 2). Each arc of this bipartite network  
486 represents the shortest path between a depot and a biofuel plant. We calculated the shortest paths using  
487 the Dijkstra's algorithm (Ahuja et al., 1993).

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

## 491 **6. Experimental Results**

492 The augmented  $\epsilon$ -constraint algorithm is implemented using C++. The IBM CPLEX 12.5.1 Concert  
493 Technology is used to solve the MILP models. All tests were conducted on a desktop computer with Intel  
494  $\text{\textcircled{R}}$  Core i7 3.1 GHz CPU and 32 GB memory limit, on a windows operating system.

### 495 **6.1 Comparing the cost minimization and the multi-objective optimization models**

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

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

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

514 solutions generated we selected the one with highest number of jobs created, and then, among those  
 515 solutions, we selected the one with the lowest emission levels.

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

521 **Table 3: Model comparisons based on biomass delivery**

Scenario	Cost Minimization Model			Multi-objective Model		
	Costs	Emissions	Jobs	Costs	Emissions	Jobs
	(\$/gal)	(lbs/gal)	(nr)	(\$/gal)	(lbs/gal)	(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

522  
 523 Table 4 compares the two models based on the amount of biomass delivered by truck and rail. Hubs  
 524 are used to facilitate rail transportation. The multi-objective model relies more on rail transportation.  
 525 Emissions are smaller for this transportation mode due to the fact that in each trip, higher volumes of  
 526 biomass and biofuel are delivered. To facilitate rail transportation more hubs are utilized.

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

532  
 533 **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

534  
 535 **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

536

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

547 Note that, the RFS goals set by EPA were reduced in 2014 below the volumes originally set by  
548 Congress (EPA, 2014). Based on the new goals, in 2014, only 33 MGY of cellulosic biofuel is expected  
549 to be produced. This number increases to 206 MGY by 2016. In 2015, the total RFS requirements are  
550 15.93BGY. The percentages presented in Tables 7(a) and 7(b) are with respect to overall RFS  
551 requirements. Clearly, the requirements set on cellulosic biomass can be met at a unit cost between \$3.5-4  
552 per gallon.

553

**Table 6:** Model comparisons based on network design

554

555

556

Scenario	Nr. of biofuel plants open				Total capacity (MGY)	Utilization (%)	Biofuel production (MGY)	% of RFS goals met
	60 (MGY)	90 (MGY)	120 (MGY)	120 (MGY)				
1	2	2	7	3	1,140	95.69%	1,091	6.82%
2	1	2	8	3	1,200	85.81%	1,004	6.27%
3	1	3	7	4	1,110	92.86%	1,114	6.96%
3	1	3	7	3	1,170	94.42%	1,048	6.55%
3	5	5	3	3	1,110	90.21%	1,161	7.75%
3	5	5	3	3	1,110	90.91%	1,009	6.31%

557

558

(a) Cost minimization model

559

560

561

562

(b) Multi-objective optimization model

563

564

565

566

567

568

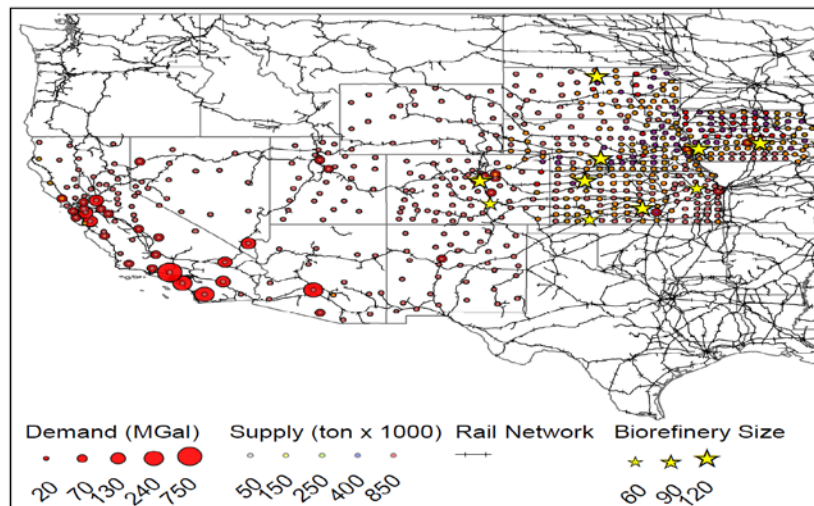
569

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.

570

571

572

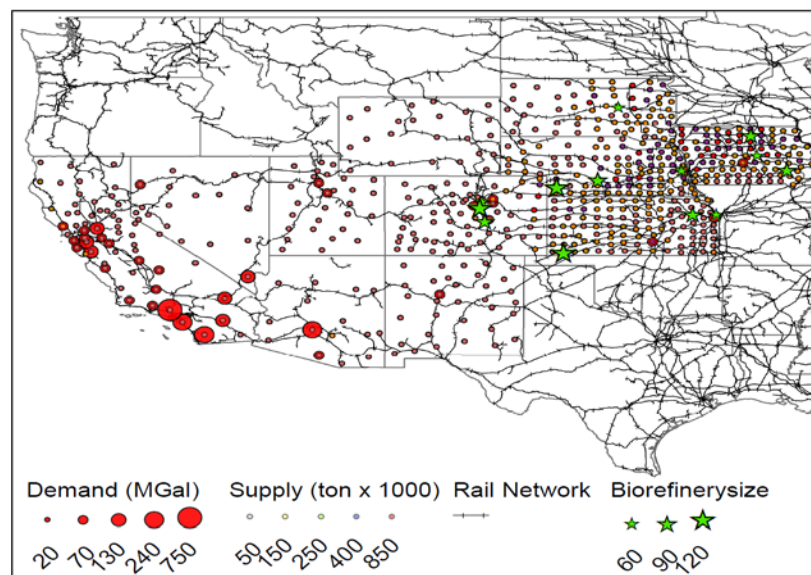


**Figure 5:** Network structure for Scenario 3 of cost minimization model

573

574

575



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

576 **6.2 Pareto curve**

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

584 Figure 7(a) plots the relationship between the unit delivery for cost and CO<sub>2</sub> emissions for different  
585 levels of targeted number of jobs created under Scenario 1. Figure 7(b) plots the relationship between the  
586 unit delivery for cost and number of jobs created for different levels of targeted CO<sub>2</sub> emissions under  
587 Scenario 1. Similar plots for Scenario 2 are presented in Figures 8(a) and 8(b), and for Scenario 3, results  
588 are presented in Figures 9(a) and 9(b).

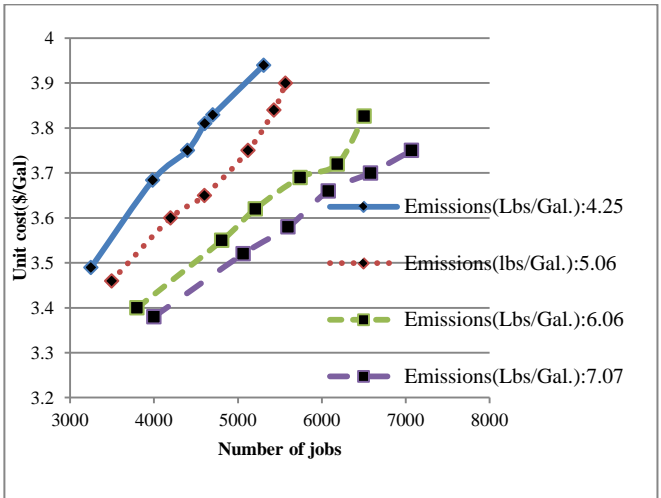
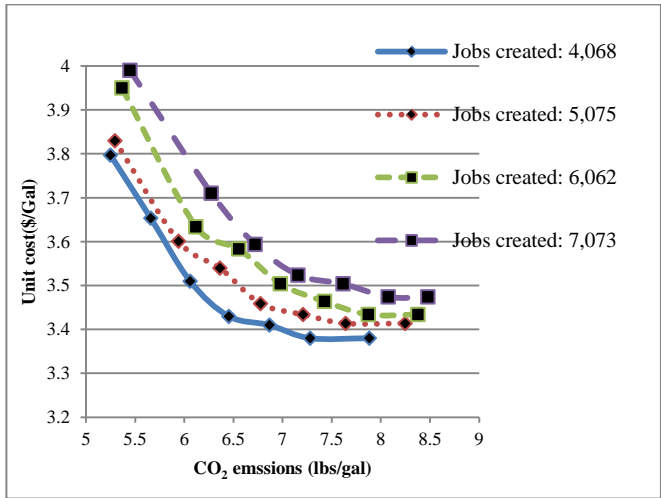
589 Results from these figures indicate that, for a given job target as the emission level decreases,  
590 delivery cost increases. These relationships are intuitive. To decrease emission levels, biofuel plants  
591 should reduce shipment volumes by truck. This requires investments to increase the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

595 The shape of the curves presented in Figures 7(a), 8(a) and 9(a) is similar and indicates a negative  
596 relationship between unit costs and unit emissions. That means, reducing CO<sub>2</sub> emissions from supply  
597 chain activities increases the cost of delivering biomass. However, the shape of the Pareto curve becomes  
598 flatter when emission levels are between 6 and 8 lbs/gal. That means, reducing CO<sub>2</sub> emissions from 8 to 6  
599 lbs/gal (Figure 8a) increases the unit cost by 10 cents. The marginal increase in costs increases as  
600 emission reductions approach 4 lbs/gal. Reductions in emissions could be achieved via imposing an  
601 emission tax, setting an emission cap, etc. Clearly these policies would impact costs in the supply chain.  
602 However, it is often possible to have a great impact on emission reductions with only marginal increases  
603 in costs.

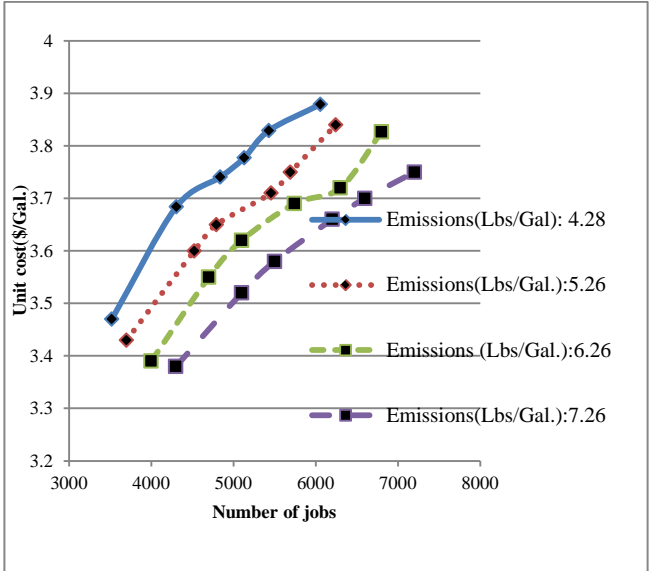
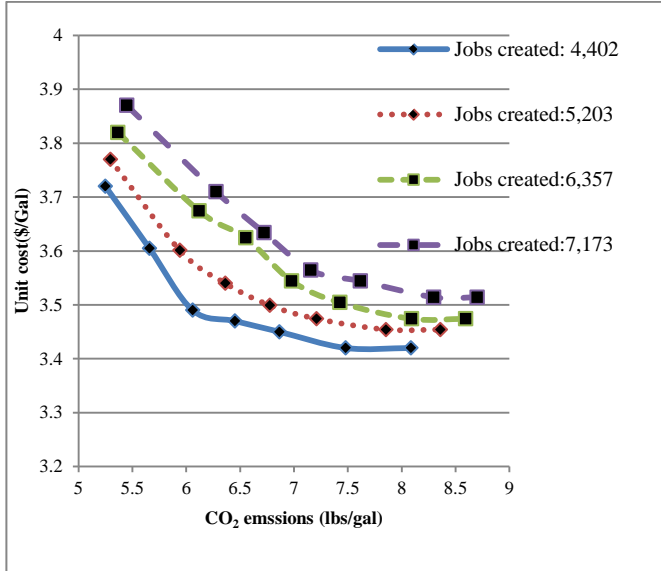
604 The results in Figures 7(a), 8(a) and 9(a) indicate that, in order to comply with increased restrictions  
605 on CO<sub>2</sub> emissions, plants need to rely on rail shipments. For this reason, at low emission levels more hubs  
606 are utilized and the investments on the infrastructure are higher. As emission levels increase, the  
607 restriction on emissions become redundant and do not have an effect on costs anymore. This is the reason  
608 why at high emission levels, increasing emissions does not affect the unit cost.



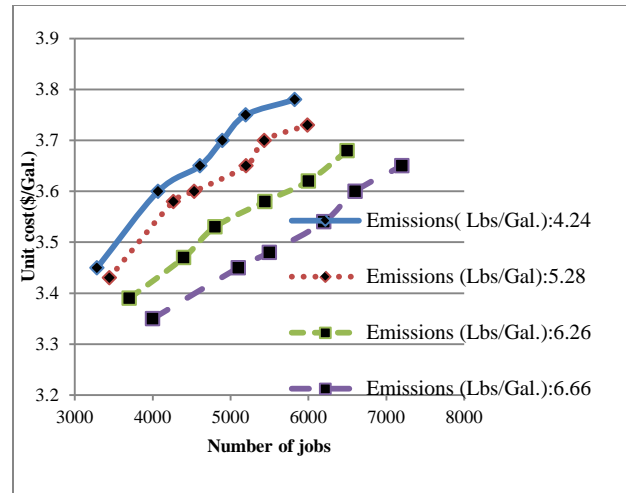
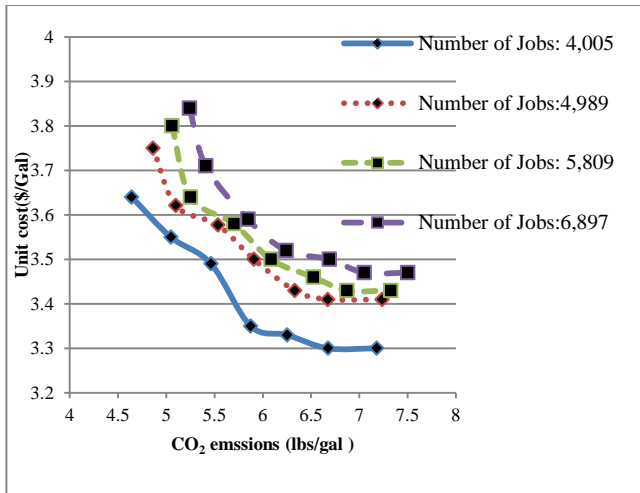
609 The results from Figures 7(b), 8(b), and 9(b) indicate a positive relationship between the number  
 610 of jobs created and the unit cost. More jobs are created when truck - rather than rail - is used to deliver  
 611 biomass. This is mainly because to ship the same amount of biomass, less railroad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



613 (a) (b)  
 614 **Figure 7:** Pareto curves for Scenario 1: (a) Unit delivery cost versus CO<sub>2</sub> emissions for different targeted  
 615 number of job created; (b) Unit delivery cost versus number of job for particular emission target.  
 616  
 617



618 (a) (b)  
 619 **Figure 8:** Pareto curves for Scenario 2: (a) Unit delivery cost versus CO<sub>2</sub> emissions for different targeted  
 620 number of job created; (b) Unit delivery cost versus number of job for particular emission target.



621

(a)

(b)

622

623 **Figure 9:** Pareto curves for Scenario 3: (a) Unit delivery cost versus CO<sub>2</sub> emissions for different targeted  
 624 number of job created; (b) Unit delivery cost versus number of job for particular emission target.

625

## 626 7. Conclusion

627 In this paper, we present a multi-objective optimization model for the cellulosic ethanol supply  
 628 chain. The model optimizes costs, environmental, and social impacts of this supply chain. The cost  
 629 objective represents transportation, facility location, and operations costs. The environmental objective  
 630 represents CO<sub>2</sub> emissions due to transportation, facility construction, and operations. The social objective  
 631 represents the number of new jobs created in order to handle transportation, hub operations, biofuel plant  
 632 construction and operations. The multi-objective model is solved using an augmented  $\epsilon$ -constraint  
 633 method. This method identifies a set of Pareto optimal solutions. The relationship among the  
 634 corresponding objectives is depicted through a number of graphs presented in the paper.

635 The underlying supply chain has a hub-and-spoke network structure. Such a network structure is  
 636 appropriate for the delivery of bulk products, such as biomass, or cellulosic ethanol. In this network,  
 637 depots serve as shipment consolidation points where small shipments of biomass from preprocessing  
 638 facilities are consolidated into high-volume shipments. High-volume shipments of biomass are then  
 639 delivered to biofuel plants by rail. Such a system positively impacts transportation costs, and  
 640 consequently, the delivery cost of cellulosic ethanol, and CO<sub>2</sub> emissions. Using rail transportation, rather  
 641 than truck, for high-volume and long-haul shipments reduces emissions.

642 The numerical analyses indicate that the goals set by the 2014 RFS for production of cellulosic  
 643 biofuel can be met. The minimum cost model does minimize the delivery cost of cellulosic biofuel, but  
 644 the multi-objective model has a greater positive impact on the environment and society. The minimum

645 cost model invests on building large sized production plants to take advantage of the economies of scale  
646 that come with producing in large quantities. This model does not invest as much in building rail hubs,  
647 and relies on truck transportation. The multi-objective model proposes investments in building more  
648 small sized plants that employ additional workforce. The corresponding supply chain relies on rail  
649 transportation to reduce CO<sub>2</sub> emissions, and uses a larger number of hubs to enable the delivery of  
650 biomass.

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

## 655 **8. Acknowledgments**

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

## 658 **9. Reference:**

- 659 Abounacer,R., Rekik, M., Renaud, J. (2014). An exact solution approach for multi-objective location–  
660 transportation problem for disaster response. *Computers & Operations Research*, 41(1), 83-93.
- 661 Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton,  
662 A. and Lukas, J. (2002). Lignocellulosic biomass to ethanol process design and economics utilizing  
663 co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. *NREL/TP-510-32438*.  
664 Golden Co: National Renewable Energy Laboratory.
- 665 Ahuja, R.K., Magnanti, T.L., Orlin, J.B. (1993). *Network Flows: Theory, Algorithms, and Applications*.  
666 Prentice Hall Publishers, Upper Saddle River, NJ, pp. 115–121.
- 667 Alumur, S., Kara, B.Y. (2008). Network hub location problems: the state of the art. *European Journal of*  
668 *Operational Research*, 190, 1–21.
- 669 Argo, M.R., Tan, E.C.D., Inman, D, Langholtz, M.H., Eaton, L.M., Jacobson, J.J, Wright, C.T., Muth,  
670 D.J., Wu, M.W., Chiu, Y.W., Graham, R.L. (2013). Investigation of biochemical biorefinery sizing  
671 and environmental sustainability impacts for conventional bale system and advanced uniform biomass  
672 logistics designs. *Journal of Biofuels, Bioproducts, and Biorefining*, 7(3), 282–302.
- 673 Bai, Y., Hwang, T., Kang, S., Ouyang, Y. (2011). Biofuel refinery location and supply chain planning  
674 under traffic congestion. *Transportation Research Part B: Methodological*, 45 (1), 162–175.
- 675 Brower, M., (2010). Woody biomass economics. Available on-line at: [http://www.mosaicllc.com/  
676 documents/34.pdf](http://www.mosaicllc.com/documents/34.pdf)

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

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

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

683 Chen, J. (2007). A hybrid heuristic for the uncapacitated single allocation hub location problem. *Omega*,  
684 35 (2), 211–220.

685 Chen, C., Fan, Y. (2012). Bioethanol supply chain system planning under supply and demand  
686 uncertainties. *Transportation Research Part E: Logistics and Transportation Review*, 48 (1), 150–  
687 164.

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

690 Cunha, C.B., Silva, M.R. (2007). A genetic algorithm for the problem of configuring a hub-and-spoke  
691 network for a LTL trucking company in Brazil. *European Journal of Operational Research*, 179 (3),  
692 747–758.

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

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

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

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

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

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

708 Gonzales, D., Searcy, E.M., Ekşioğlu, S.D. (2013). Cost analysis for high-volume and long-haul  
709 transportation of densified biomass feedstock. *Transportation Research Part A: Policy and Practice*,  
710 49, 48-61

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

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

717 Idaho National Laboratory (2014). *Feedstock and Conversion Supply System Design and Analysis, The  
718 Feedstock Logistics Design Case*. Available on-line at: [file:///C:/Users/seksiog/Downloads/Feedstock  
719 %20and%20Conversion%20Supply%20System%20Design%20and%20Analysis.pdf](file:///C:/Users/seksiog/Downloads/Feedstock%20and%20Conversion%20Supply%20System%20Design%20and%20Analysis.pdf)

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

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

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

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

729 Kim, J., Realff, M.J., Lee, J.H. (2011b). Optimal design and global sensitivity analysis of biomass supply  
730 chain networks for biofuels under uncertainty. *Computers & Chemical Engineering*, 35 (9), 1738–  
731 1751.

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

734 Kumar, A.; Sokhansanj, S. (2007). Switchgrass (*Panicum virgatum*, L.) delivery to a biorefinery using  
735 integrated biomass supply analysis and logistics (IBSAL) model. *Bioresource Technology*, 98(5),  
736 1033-1044.

737 Labbe, M., Yaman, H. (2004). Projecting flow variables for hub location problems. *Networks*, 44  
738 (2), 84–93.

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

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

743 Marufuzzaman M., Eksioğlu S.D., Hernandez R. (2014). Environmentally Friendly Supply Chain  
744 Planning and Design for Biodiesel Production via Wastewater Sludge. *Transportation Science*, 48(4),  
745 555-574.

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

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

751 Mele, F.D, Guille'n-Gosa'lbez, G., Jimé'nez, L. (2009). Optimal planning of supply chains for bioethanol  
752 and sugar production with economic and environmental concerns. *Comput Aided Chem Eng*, 26, 997–  
753 1002.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

805 United States Department of Energy (2011). U.S. Billion-Ton Update: Biomass Supply for a  
806 Bioenergy and Bioproducts Industry. R.D. Perlack and B.J. Stokes (Leads), ORNL/TM-  
807 2011/224. Oak Ridge National Laboratory, Oak Ridge, TN. 227p.

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

810 United States Energy Information Administration (2010). Monthly Energy Review. Available on-line at:  
811 <http://tonto.eia.doe.gov/FTP/ROOT/multifuel/mer/00351005.pdf>

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

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

817 United States Environmental Protection Agency (EPA) (2007). Energy Independence and Security Act of  
818 2007. In: *RL34294, P.L.110–140 ed. 2007*.

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

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

825 You, F., Wang, B. (2011). Life-cycle optimization of biomass-to-liquid supply chains with distributed–  
826 centralized processing networks. *Industrial & Engineering Chemistry Research*, 50(17), 10102-  
827 10127.

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

831 Zamboni, A., Bezzo, F., Shah, N. (2009). Spatially explicit static model for the strategic design of future  
832 bioethanol production systems. 2. Multi-objective environmental optimization. *Energy & Fuels*, 23  
833 (10), 5134-5143.

834 Zhang, W., Reimann, M. (2014). A simple augmented  $\epsilon$ -constraint method for multi-objective  
835 mathematical integer programming problems. *European Journal of Operational Research*, 234(1),  
836 15-24.

837



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 $(i, j) \in A$
$\Psi_{ij}$	reflects a fixed cost for loading/unloading a unit train $(i, j) \in R_1$
$\lambda_{ij}$	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 $(i, j) \in R_1$
$\tau_{ij}$	represents the maximum capacity of a rail car along arc $(i, j) \in R_2$
$\zeta_i$	fixed investment cost at node $i \in D$
$u_i$	Capacity of node $i \in D$

$q_{ik}$	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 $i \in C$
$\alpha_i$	shortage cost at customer location $i \in C$
$q_{ik}$	capacity of biorefinery node $i \in B$ is $k \in K$
<i>Emission parameters</i>	
$e_{ij}$	CO <sub>2</sub> emission per ton per mile in arc set $(i, j) \in T_1, T_2, T_3, R$
$\epsilon_{ik}$	CO <sub>2</sub> emission from biorefinery $i \in B$ with capacity $k \in K$
$o_i$	CO <sub>2</sub> emission for establishing a hub at node $i \in D$
<i>Social factors</i>	
$p_{ij}^T$	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 $i \in D$
<i>Decision variables</i>	
$X_{ij}$	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
$\beta_{ik}$	a binary variable which takes the value 1 if $i$ is used as a biorefinery, with capacity $k$ and 0 O/W
$\Pi_i$	demand shortage at customer location $i \in C$

845

## APPENDIX B

846 **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

847

848 **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

	120	139	286				
AZ	60	98	191		120	128	207
	90	123	248				
	120	157	286				

849