DECISION SCIENCES INSTITUTE

Data Envelopment Analysis Approach and Its Application for Biomass to Biofuel Supply Chain Design

Jae-Dong Hong
Industrial Engineering
South Carolina State University
Orangeburg, SC 29117
Email: jhong@scsu.edu

ABSTRACT

This paper proposes an innovative procedure of combining a goal programming (GP) model into Data Envelopment Analysis (DEA) approach for an efficient biomass to biofuel supply chain (*BBSC*) design under the risk of disruptions. Solving the GP model for various values of weights, which simultaneously takes four performance measures into consideration, generates various alternative options. A case study using the biomass data for South Carolina is conducted to evaluate the proposed procedure.

KEYWORDS: Data envelopment analysis, Biomass to biofuel supply chain, Goal programming model

INTRODUCTION

Biofuel production from biomass has been recognized as an important alternative that will reduce nation's dependency on limited fossil fuel source and have a positive impact on the economy, environment, and society. In this vein, Balaman and Selim (2015) emphasize that design, operation, and management issues in renewable energy supply chains are increasingly gaining importance in recent years in parallel with the rising interest in renewable energy sources.

Eksioglu et al. (2009) investigate the biofuel logistics network design, taking into account the optimization of facility locations (e.g., collection facilities, bio-refineries), transportation, and inventory control. Based upon the issues and challenges of designing and operating biomass chains that secure stable and competitively- priced feedstock supply for biofuel plants, Gold & Seuring (2011) classify those issues and challenges into the operations, harvesting and collection, storage, transport and pre-treatment techniques as well as into overall supply system design. Hong et al. (2014) consider the effect of the uncertainty in biomass yield and propose a simulation-based robust biofuel facility location model for solving a bio-energy logistics network problem with the objective of minimizing the total logistics cost (TLC). Poudel et al. (2016) present a pre-disaster planning model for designing a reliable biofuel supply chain network considering link failure probabilities.

In this paper, we assume that the locations of HS and BS are fixed, and the demand of each BS is constant throughout the planning period. We also assume that the two facilities in the inbound flows, CF and BRF, are under the risk of disruptions. The risk of disruptions implies that, as Cui et al. (2010) define, some of the constructed facilities may become unavailable due to disruptions caused by a disaster. Thus, we assuming that if a facility is disrupted, biomass would not be used to produce biofuel, we consider four goals as the major performance measures. The first goal is to minimize the total logistics cost (TLC). The second goal is to minimize the maximum demand-weighted coverage distance (MDWCD) between the demand

point and the supply point. The third goal is to maximize the expected amount of biomass feedstock (EABF) flowing into BRFs, whereas the fourth goal is to maximize the expected amount of biofuel production (EPAB) by BRFs. To accommodate these four goals in one objective function, we use goal programming (GP) approach as a tool for designing the effective *BBSC*. The typical GP model allows the decision maker to assign weights to the deviational variables in the objective function to better reflect the importance and desirability of deviations from the various goals. For a given set of weighting factors, we solve GP model to generate various supply chain network schemes which represent Decision Making Units (DMUs) used for Data Envelopment Analysis (DEA). We apply DEA to identify efficient supply chain network schemes among the alternatives generated by GP. Then we also apply the stratification/context-dependent DEA evaluation method to the efficient supply chain network schemes to find the best options.

As far as we know of, applying DEA technique to *BBSC* design problem has not been tried in the literature. This process would help practitioners as well as researchers to produce a finer evaluation of the efficiency of the supply chain network system and to provide a design and benchmarking framework for designing *BBSC* system to improve overall supply chain efficiency.

GOAL PROGRAMMING MODEL

Let F be the set of all harvesting sites (HSs) and potential collection facility (CF) locations, indexed by f. Now, let J, I, and K respectively be the set of CFs, BRFs, and BSs, indexed by j, i, and k. Also, let L and G respectively be the set of capacities of BRF and CF, indexed by l and g. The parameters used in this formulation are the following: ψ_{il}^{b} is amortized annual cost of constructing and operating a BRF_i with the l^{th} size; ψ_{jg}^{c} is amortized annual cost of constructing and operating a CF_j with the g^{th} size; C_l^b and C_g^c denote the actual capacity of l^{th} and g^{th} size of BRF and CF, respectively; β_f and γ_f are conversion rates to bio-energy of biomass feedstock shipped from CF to BRF and from HS to BRF, respectively; S_f denotes the yield of biomass feedstock from HS_f; D_k is the demand of biofuel for BS_k; δ_i is the maximum number of HSs that ship biomass directly to BRF_i; d_{fj}^1, d_{fi}^2, d_{ji}^3, and d_{ik}^4 are unit transportation cost (UTC) from HS_f to CF_j, from HS_f to BRF_i, from CF_j to BRF_i, and from BRF_i to BS_k, respectively. In this study, we set $d_{fi}^2 \geq d_{fj}^1$, to denote a higher unit transportation cost for shipping biomass from HS_f directly to BRF_i.

The decision variables are the following: x_{il}^{b} is a binary variable that equals 1 if a biorefinery of size l is located in site i, and 0 otherwise; x_{jg}^{c} is a binary variable that equals 1 if a collection facility of size g is located in site j, and 0 otherwise; y_{fj}^1 is a binary variable that equals 1 if HS_f's yielded biomass shipped to CF_j and 0 otherwise; y_{fi}^2 is a binary variable that equals 1 if HS_f ships biomass directly to BRF_i, and 0 otherwise; y_{ji}^3 is a binary variable that equals to 1 if CF_j is assigned to BRF_i, and 0 otherwise; y_{ik}^4 is the fraction of BRF_i' produced biofuel shipped to BS_k. Letting N_b and N_c denote the maximum number of BRFs and CFs to be built, we formulate the following MIQP model that minimizes the total logistics cost (*TLC*), which is the sum of the annualized construction and operation cost for CFs and BRFs and the transportation costs from HSs to CFs, HSs to BRFs, CFs to BRFs, and BRFs to BSs:

$$TLC = \left[\sum_{i\in I}\sum_{l\in L}\psi_{il}^{b}\,x_{il}^{b} + \sum_{j\in J}\sum_{g\in G}\psi_{jg}^{c}\,x_{jg}^{c}\right] + \left[\sum_{j\in J}\sum_{f\in F}S_f\,d_{fj}^{1}y_{fj}^{1} + \sum_{i\in I}\sum_{f\in F}S_f\,d_{fi}^{2}y_{fi}^{2}\right]$$

$$+\left[\sum_{i\in I}\sum_{j\in J}\left(\sum_{f\in F}S_f y_{fj}^1\right)d_{ji}^3 y_{ji}^3\right]+\left[\sum_{i\in I}\sum_{k\in K}D_k\, d_{ik}^4 y_{ik}^4\right] \quad (1)$$

Since the long transportation distance/time might cause the poor quality of biomass feedstock, our next goal is to minimize the maximum demand-weighted coverage (transportation) distance (*MDWCD*), which is given by

$$MDWCD = Max\{S_f d_{fj}^1 y_{fj}^1, S_f d_{fi}^2 y_{fi}^2,\ S_f d_{ji}^3 y_{fj}^1\}\ \ \forall f, i, and\ j. \quad (2)$$

Now, let p_j^c and p_i^b denote the probability that CF_j and BRF_i are disrupted, respectively. Then the expected amount of biomass feedstock (*EABF*) flowing into BRFs would be express as

$$EABF = \sum_{i\in I}\sum_{j\in J}\left[\sum_{f\in F}S_f\, y_{fj}^1\, y_{ji}^3(1-p_j^c)\left(1-p_i^b\right)\right]+\left[\sum_{i\in I}\sum_{f\in F}S_f\, y_{fi}^2\left(1-p_i^b\right)\right]. \quad (3)$$

To find the expected production amount of biofuel (*EPAB*), let β_f denote the conversion rate to biofuel of biomass feedstock shipped from CF to BRF and γ_f the conversion rate to biofuel for biomass feedstock shipped from HS to BRF. Then, *EPAB* is expressed as

$$EPAB = \sum_{i\in I}\sum_{j\in J}\left[\sum_{f\in F}S_f\, y_{fj}^1\, y_{ji}^3(1-p_j^c)\left(1-p_i^b\right)\right]\beta_f+\left[\sum_{i\in I}\sum_{f\in F}S_f\, y_{fi}^2\left(1-p_i^b\right)\right]\gamma_f. \quad (4)$$

Let the nonnegative deviation variables, $\delta_{TLC}^+, \delta_{MDWCD}^+, \delta_{EABF}^-, and\ \delta_{EPAB}^-$, denote the amounts by which each value of *TLC*, *MDWCD*, *EABF*, and *EPAB* deviates from the minimum values of *TLC* and *MDWCD*, TLC_{min} and $MDWCD_{min}$ and maximum values of *EABF* and *EPAB*, $EABF_{max}$ and $EPAB_{max}$, respectively. Then, the deviation variables are given by

$$\delta_{TLC}^+ = TLC\ in\ (1) - TLC_{min}, \quad (5)$$

$$\delta_{MDWCD}^+ = MDWCD\ in\ (2) - MDWCD_{min}, \quad (6)$$

$$\delta_{EABF}^- = EABF_{max} - EABF\ in\ (3), \quad (7)$$

and

$$\delta_{EPAB}^- = EPAB_{max} - EABP\ in\ (4). \quad (8)$$

Now, the weighted sum of the percentage deviations is defined as

$$G(\alpha) = \alpha_1\frac{\delta_{TLC}^+}{TLC_{min}}+\alpha_2\frac{\delta_{MDWCD}^+}{MDWCD_{min}}+\alpha_3\frac{\delta_{EABF}^-}{EABF_{max}}+\alpha_4\frac{\delta_{EPAB}^-}{EPAB_{max}}, \quad (9)$$

where α_g, g=1, 2, 3, and 4, is a weight factor ranging from 0 and 1 and $\sum_g \alpha_g = 1$. All constraints come from Hong et al. (2014).

CASE STUDY USING EPA TRACKED SITES IN SOUTH CAROLINA

We use and follow the scenario illustrated in Figure 1 (EPA Tracked sites in South Carolina (SC) with biorefinery facility siting potential, 2013). Sixteen (16) counties, whose biomass resources are classified 'good' or better as shown in Figure 1, are selected as the harvesting sites (HSs). Then, one city from each county using a centroid approach was chosen as the candidate location for collection facility (CF) location. Five (5) locations {Branchville, Cayce, Lake City, Prosperity, Ridgeland} and ten locations (10) throughout SC are selected as candidate sites for BRFs and blending stations (BSs), respectively. The potential locations for BRFs are selected based upon low population density, easy access to interstate highways, etc. We use the actual distances between cities representing HSs, CFs, BRFs, and BFs. Note that we assume that the conversion rates, β_f and γ_f, are 70% and 30%, respectively. As shown in Figure 1, the minimum, maximum, and average amounts of biomass yield at each HS are summarized in Table 1. Based on these input data, an Excel spreadsheet model for MILP problem is developed and solved using Analytic Solve Platform with VBA (Visual Basic for Applications).

Table 1: Biomass Yield

No	Harvest Site	Minimum Yield (Thousand Metric Tons)	Average (Thousand Metric Tons)	Maximum Yield (Thousand Metric Tons)	Risk Probability
1	Allendale	100	150	200	0.36
2	Berkeley	150	200	250	0.44
3	Chester	150	225	300	0.28
4	Colleton	100	200	300	0.36
5	Darlington	150	225	300	0.40
6	Dorchester	150	225	300	0.36
7	Florence	150	225	300	0.48
8	Georgetown	250	400	550	0.48
9	Greenwood	150	225	300	0.24
10	Hampton	150	225	300	0.28
11	Horry	100	175	250	0.64
12	Lexington	100	175	250	0.44
13	Newberry	250	400	550	0.36
14	Orangeburg	150	225	300	0.44
15	Richland	250	400	550	0.44
16	York	150	225	300	0.32

Before solving the GP model, it is necessary to determine the target values for each goal parameter. These values can be found by solving the GP model with the objective function of the corresponding equations. In fact, each of these target values could be obtained by setting the corresponding weight equal to 1 and solving the GP problem. For example, setting $\boldsymbol{\alpha}$ = (1, 0, 0, 0) and solving the model yields the target value of *TLC*, TLC_{min}. Similarly, setting $\boldsymbol{\alpha}$ = (0, 0,

1. 0) and solving the model yields the target value of *EABF*, $EABF_{max}$. Now, we solve and summarize the target values for the performance metrics in Table 2.

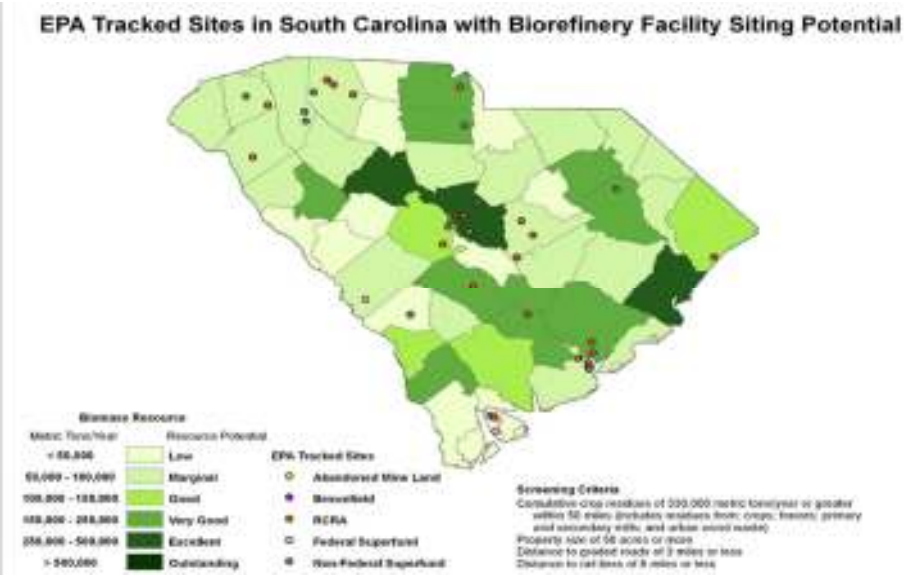
Table 2: The Target Values of Four Performance Measures

	α=(α_1, α_2, α_3, α_4 **)**	Target Value
TLC_{min}	(1, 0, 0, 0)	$4,021.97
$MDWCD_{min}$	(0, 1, 0, 0)	89,596
$EABF_{max}$	(0, 0, 1, 0)	2,701 MT
$EAPB_{max}$	(0, 0, 0, 1)	1,491 TG

Using the values in Table 2, the GP model is solved for various values of the weight, α_g, where each weight changes between 0 and 1 with an increment of 0.1. There are 286 configurations arising out of the combinations of the setting of ***α*** under the condition, $\sum_{g=1}^{4} \alpha_g = 1$. After we solve the model, we reduce 286 configurations into 99 consolidated configurations, based upon the values of the four performance measures. Considering each configuration as DMU, which represents the optimal locations of BRFs and CFs and assignments of BRFs to BSs for each given set of weights, we apply data envelopment analysis (DEA) method to find the efficiency score (ES) for each decision making unit (DMU) (see Zhu, 2014). Then, using context-dependent DEA (see Seiford & Zhu, 2003), we stratify 99 DMUs into 14 levels, identify DMUs in each level, and compute attractiveness scores (ASs) for the DMUs in Level 1 against the DMUs in all other 13 levels. In Table 3, we present the DMUs belonging to Level 1 with efficiency score (ES), the set of weights, the values of four performance measure, and the average attractiveness scores (AASs) for the DMUs in Level 1. We rank them based upon AAS and report the rankings too.

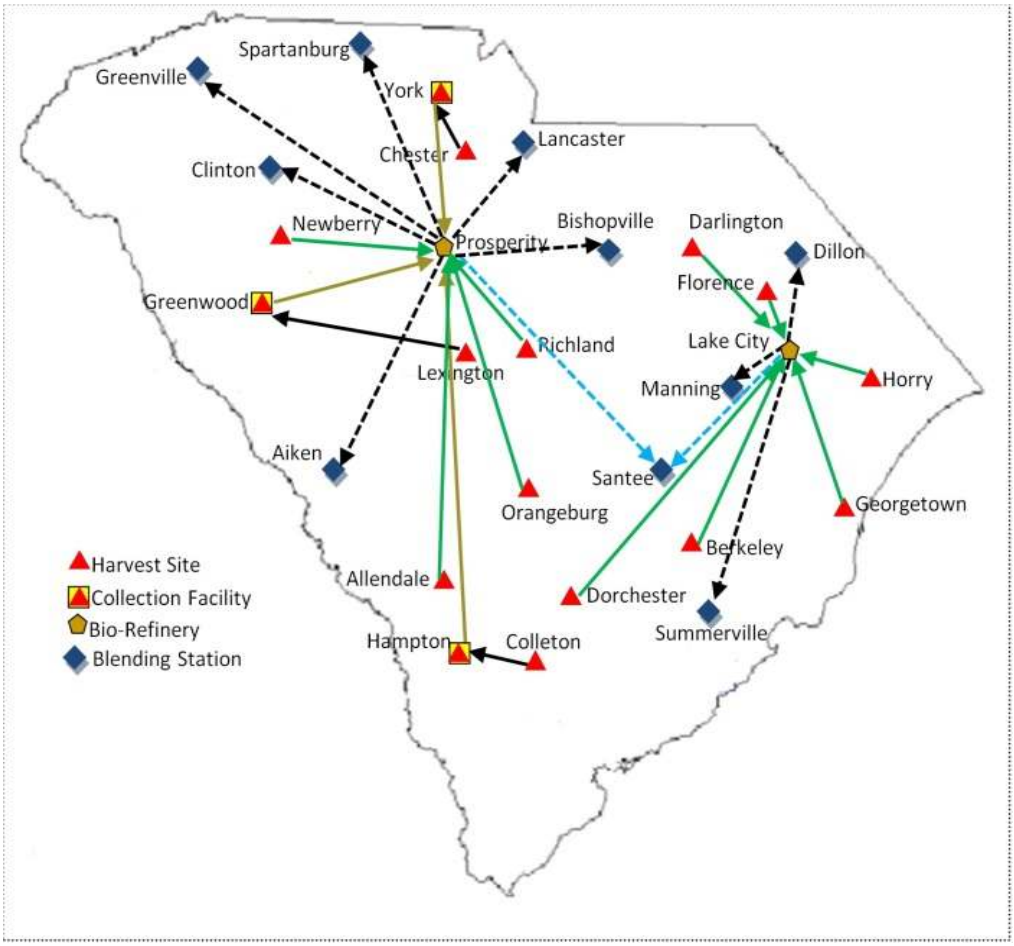
In Table 3, we see that DMU #79, generated from solving GP model with sets of weights presented in Table 3.A, obtains the highest AAS of 1.9297. DMU #89, generated with ***α*** = (0.3, 0.5, 0.0, 0.2) or ***α*** = (0.5, 0.2, 0.0, 0.3), produces 1.9296 as AAS, which is slightly less than the top ranked DMU #79. The other two efficient DMUs in Level 1, DMUs #94 and #98, yield lower TLCs and higher PPBs, but a much higher MDWCD than DMUs #79 and #89.

Figure 1: EPA Traced sites in South Carolina with biorefinery facility siting potential



The most efficient supply chain scheme in Level 1, DMU #79, is depicted in Figure 2. We find that all schemes select {Prosperity, Lake City} as BRFs and allocation from BRFs to BSs are same for all schemes. As we put more weights on *TLC*, the direct shipments from HSs to BRFs increase to reduce *TLC*. Note that, in the efficient schemes to generate more *BPO*, the transportation route does not look as efficient.

Figure 2: Efficient biomass-biofuel supply chain networks (DMU #79)



From the above observation of the results generated by GP and DEA/SDEA, we can claim that the traditional single-objective models of minimizing the total cost or maximizing the biofuel production may not generate the efficient supply chain schemes. Thus, our innovative approach proposed in this paper would help decision-makers design the efficient biomass to biofuel supply chain schemes and find the robust biofuel facility location-allocation decisions.

SUMMARY AND CONCLUSIONS

To design more balanced biomass-to-biofuel supply chain (*BBSC*) design scheme under the risk of disruption, four different performance metrics are considered. To deal with such multiple objectives, we use Goal-Programming (GP) approach which would be able to provide various *BBSC* scheme alternatives for decision makers. To evaluate these supply chain alternatives, we

propose data envelopment analysis (DEA) method which enables us to find efficient supply chain schemes from the various alternatives generated by GP. In addition, we introduce stratification/context-dependent DEA (SDEA) methodology to stratify the alternatives generated by GP into different efficiency levels. Using these levels, we compute the average attractiveness scores (AASs) for the efficient schemes found by DEA and rank them based upon AASs to identify the most efficient supply chain scheme. Through a case study, we demonstrate the applicability of our innovative approach of combining GP into DEA method for BBSC design problem. We observe that our approach can help decision makers evaluate various BBSC options and develop efficient and robust design schemes.

In this paper, we consider the simple average conversion rates for different types of biomass, For future research, it would be interesting to consider identifying blends of different types of biomass whose conversion rates will be significantly different. It would also be interesting to consider other goals, such as CO_2 emission due to biorefinery operations and social impacts by estimating the number of jobs created due to various *BBSC* activities.

ACKNOWLEDGEMENTS

This material is based upon work that is supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, Evans-Allen project number SCX-313-03-15.

REFERENCES

References available upon request.

Table 3: Results for DMUs in Level 1 and Level 2										
Level	DMU	**α =** (α_1, α_2, α_3, α_4)	TLC	MDWCD	EABF	EPAB	PBB	ES	AAS	Rank
1	79	See Table 3.A	$4,348.87	89,596	2,103.76	1,050.72	0.4994	1.0000	1.9297	1
	89	(0.3, 0.5, 0.0, 0.2), (0.5, 0.2, 0.0, 0.3)	$4,304.43	89,596	2,061.84	1,027.68	0.4984	1.0000	1.9296	2
	94	(0.5, 0.0, 0.0, 0.5), (0.5, 0.0, 0.1, 0.4), (0.6, 0.0, 0.0, 0.4), (0.6, 0.0, 0.1, 0.3), (0.7, 0.0, 0.0, 0.3), (0.7, 0.0, 0.1, 0.2)	$4,077.86	206,309	1,782.52	1,060.74	0.5950	1.0000	1.5387	3
	98	(0.8, 0.0, 0.0, 0.2)	$4,038.59	206,309	1,731.32	1,022.34	0.5904	1.0000	1.5384	4

Table 3.a: Set of Weights for DMU #79	
DMU	**α =** (α_1, α_2, α_3, α_4)
79	(0.1, 0.7, 0.0, 0.2), (0.1, 0.8, 0.0, 0.1), (0.1, 0.8, 0.1, 0.0), (0.2, 0.4, 0.0, 0.4), (0.2, 0.4, 0.1, 0.3), (0.2, 0.5, 0.0, 0.3), (0.2, 0.5, 0.1, 0.2), (0.2, 0.5, 0.2, 0.1), (0.2, 0.5, 0.3, 0.0), (0.2, 0.6, 0.0, 0.2), (0.2, 0.6, 0.1, 0.1), (0.2, 0.6, 0.2, 0.0), (0.3, 0.1, 0.0, 0.6), (0.3, 0.1, 0.1, 0.5), (0.3, 0.2, 0.0, 0.5), (0.3, 0.2, 0.1, 0.4), (0.3, 0.2, 0.2, 0.3), (0.3, 0.2, 0.3, 0.2), (0.3, 0.2, 0.4, 0.1), (0.3, 0.3, 0.0, 0.4), (0.3, 0.3, 0.1, 0.3), (0.3, 0.3, 0.2, 0.2), (0.3, 0.3, 0.3, 0.1), (0.3, 0.3, 0.4, 0.0), (0.3, 0.4, 0.0, 0.3), (0.3, 0.4, 0.1, 0.2), (0.3, 0.4, 0.2, 0.1), (0.3, 0.4, 0.3, 0.0), (0.4, 0.1, 0.0, 0.5), (0.4, 0.1, 0.1, 0.4), (0.4, 0.1, 0.2, 0.3), (0.4, 0.1, 0.3, 0.2), (0.4, 0.1, 0.4, 0.1), (0.4, 0.1, 0.5, 0.0), (0.4, 0.2, 0.0, 0.4), (0.4, 0.2, 0.1, 0.3), (0.4, 0.2, 0.2, 0.2), (0.4, 0.2, 0.3, 0.1), (0.4, 0.2, 0.4, 0.0), (0.4, 0.3, 0.0, 0.3), (0.4, 0.3, 0.1, 0.2), (0.4, 0.3, 0.2, 0.1), (0.5, 0.1, 0.0, 0.4), (0.5, 0.1, 0.1, 0.3), (0.5, 0.1, 0.2, 0.2)