

Optimally Locating Biorefineries: A GIS-Based Mixed Integer Linear Programming Approach

Yuanchang Xie*, Ph.D.

Assistant Professor

James E. Clyburn University Transportation Center
South Carolina State University
Orangeburg, SC 29117
Tel: (803) 536-8321
Fax: (803) 516-4693
E-mail: yxie@scsu.edu

Kaiguang Zhao, Ph.D.

Post-Doc Research Fellow

Spatial Science Lab

Texas A&M University
College Station, TX 77843
E-mail: zhaokg@tamu.edu

Shaun Hemingway

Graduate Research Assistant

South Carolina State University

Orangeburg, SC 29117

Email: shaunhemingway@hotmail.com

* Corresponding author

Paper submitted for consideration of presentation at the 51st Transportation Research Forum

Word count: 5,035 + 2 Tables + 5 Figures = 6,785 Words

Submission Date

October 1, 2009

1 1. INTRODUCTION

2 Diverse and affordable energy is critical for America's future. To reduce the dependence on
3 foreign oil and also mitigate the environmental impacts (e.g., climate change, pollution) of using
4 fossil fuel, a significant amount of research has recently been devoted to methods on producing
5 biofuels. Less attention has been given to the cost associated with transporting bulky biomass
6 feedstock to biorefinery plants. The biomass transportation cost however is very significant
7 compared to the biofuel production cost. For this reason, a majority of existing biorefinery plants
8 in the United States are located in the Midwest where corn and soybean supplies are abundant.

9 With the soaring and unstable gasoline price and the increasing environmental concern,
10 many other states in the U.S. are now seeking the opportunity to use biomass feedstocks such as
11 switchgrass for producing biofuels. Also, under the Energy Independence and Security Act of
12 2007, the U.S. Environmental Protection Agency (EPA) has developed a Renewable Fuel
13 Standard program (RFS) to ensure that gasoline in the U.S. contains a minimum percentage of
14 renewable fuel. The latest RFS (1) "*will increase the volume of renewable fuel required to be*
15 *blended into gasoline from 9 billion gallons in 2008 to 36 billion gallons by 2022.*" Therefore,
16 there is an immediately demand for biomass transportation cost analysis model to help optimally
17 locate the new biorefineries. In this study, a GIS-based decision support system is developed for
18 this purpose. A mixed integer linear programming (MILP) model is integrated into the system.
19 Given the distribution of biomass feedstocks, transportation network, and candidate plant
20 locations, the GIS tool and the MILP model will be able to help identify the best locations for
21 biorefineries.

22

23 2. LITERATURE REVIEW

24 Only few pioneer studies have attempted to develop models for optimally locating biorefineries.
25 Graham et al. (2) developed a GIS tool that takes raster data as input. The authors used a brutal
26 search to find the best biorefinery locations. In their study, a 1km x 1km resolution cropland map
27 was used as one of the input data. The entire study area was also separated into 1km x 1km
28 pixels, and each pixel was a candidate location for biorefineries. Each of the candidate locations
29 and the cropland pixels was connected to the road network by a shortest straight line between the
30 pixel and the road network. The unit transportation costs from candidate biorefineries to any
31 pixels in the cropland map were calculated and sorted in ascending order. Based on the yield of
32 each pixel and the demand of each candidate biorefinery, the first n pixels (in the cropland map
33 satisfying the demand of the candidate biorefinery) were selected and the total transportation cost
34 for the candidate biorefinery was calculated. A similar strategy was used by Ravula (3). The
35 difference was that Ravula only considered a single-biorefinery scenario and straight line
36 distances was used in lieu of actual distances along road network. For multi-biorefinery scenarios,
37 Graham et al. (2) assumed that energy crop resources in one pixel can only be assigned to one
38 biorefinery. Based on this assumption, a sequential method was used. Starting from the first
39 biorefinery, once its best location has been identified and cropland pixels have been assigned to
40 it, those cropland pixels were removed from further consideration. The next best biorefinery
41 location was then determined following the same procedure using the remaining cropland pixels.
42 This process was repeated until the last biorefinery location was determined. Two key issues
43 were not appropriately addressed by Graham et al. (2). First, for a large study area, the entire
44 area can be separated into hundreds of thousands of 1km x 1km pixels. For instance, the state of
45 Texas can be divided into 696,241 1km x 1km pixels. Such a huge number of pixels require

1 tremendous computation time for finding the shortest distances and the lowest transportation cost
2 for each candidate biorefinery. Second, the procedure used in Graham et al. (2) cannot guarantee
3 to find the optimal locations in multi-biorefinery scenarios.

4 Instead of finding the best locations for biorefineries, Dubuc (4) developed a GIS tool to
5 identify the best locations for feedstock farms used for supplying bioenergy crops to power
6 plants. In his study, Dubuc considered a scenario with only one power plant. Given the location
7 of the power plant and the locations of candidate farm locations, the GIS tool was used to
8 identify the best locations for feedstock farms. However, it is unclear if the developed system
9 can be used for multiple power plants. In (4), the locations for the power plant and the candidate
10 feedstock farms were given, and the author did not treat each pixel as a candidate farm and use
11 brutal search to evaluate its transportation cost. Compared to the brutal search in (2), this method
12 makes more sense in many cases because finding the best locations is not a pure mathematical
13 problem. A lot of other issues that cannot be readily quantified also need to be considered in this
14 process.

15 Celli et al. (5) applied a genetic algorithm to find the best locations for biomass power
16 plants. The authors again discretized the entire study area into small cells. In the genetic
17 algorithm, each chromosome (individual solution) is of length equal to the total number of cells.
18 For a specific chromosome, if one element in it is equal to 1 (or 0), that means the corresponding
19 cell is selected (or not selected) for power plants. For large study areas such as one state, the
20 chromosomes will become extremely long, which will significantly degrade the computation
21 efficiency.

22 In a recent study conducted by Panichelli and Gnansounou (6), a BIOAL algorithm was
23 developed to find the optimal locations for gasification units (GUs). The authors initially
24 identified (6) nine candidate locations. The BIOAL algorithm was then used to choose two out of
25 them. Therefore, there are a total of 36 combinations. The BIOAL algorithm is basically an
26 enumeration method. The authors evaluated each of the 36 combinations and selected the one
27 that is of the lowest cost. Compared to optimization methods, the BIOAL algorithm is
28 computationally inefficient and cannot guarantee the optimal solution. For instance, the best
29 solution might be choosing three candidate plants instead of two.

30 The literature review shows that most previous studies did not use an explicit optimization
31 model to jointly determine the best quantities and locations of biorefineries. Almost all existing
32 methods used brutal search or enumeration methods to solve the multi-biorefinery problem. In
33 this study, the optimal biorefinery location problem is formulated as a MILP model. This MILP
34 model assumes that the candidate biorefinery locations, road network, and biomass distribution
35 data are given. As discussed previously, the optimal locations for biorefineries also depend on a
36 number of other issues that are difficult to quantify and model mathematically. Therefore,
37 compared to discretizing the study area into small cells and treating every cell as a candidate
38 biorefinery location, allowing users to specify the biorefinery locations not only makes sense but
39 also save considerable computation time especially when the study area is large. The proposed
40 MILP model is integrated into a GIS framework and a GIS-based tool is developed. The GIS tool
41 is used to manage the road network and biomass resources distribution data. Another important
42 usage of the GIS tool is to generate input data to the MILP model. Given the input data, the
43 MILP model is formulated in a special way such that the optimal quantities and locations of
44 biorefineries can be determined simultaneously.

1 3. OPTIMIZATION MODEL DEVELOPMENT

2 3.1. Modeling Single-Biorefinery Problem

3 If the problem is to find the optimal location of a single biorefinery, then it becomes
 4 straightforward to solve and no complicated mathematical models are required. For each
 5 pre-selected candidate biorefinery, one just needs to find the shortest paths from it to all biomass
 6 feedstock farms. After sorting the farms based on the shortest path values in ascending order,
 7 data from the first n farms that satisfy the subject candidate biorefinery's demand will be used to
 8 calculate its lowest possible transportation cost using Eq. (1). When looking for the best location
 9 for a single biorefinery, there is no need to consider the construction and annual operational cost
 10 of the biorefinery, as such cost does not vary significantly by location for a given study area.

$$11 \left[\sum_{i=1}^{n-1} P_i d_{ij} + \left(T_C - \sum_{i=1}^{n-1} P_i \right) d_{nj} \right] u \quad (1)$$

13 where

- 14 P_i = yield of farm i (tons);
 15 d_{ij} = weighted distance between farm i and candidate biorefinery j (miles);
 16 T_C = total amount of feedstock required by biorefinery j ; and
 17 u = unit transportation cost (\$/ton-mileage).

18
 19 Eq. (1) has two major parts. The first part calculates the transportation cost for the first $n-1$
 20 farms, while the second part calculates the transportation cost for the n th farm separately. The
 21 reason for this separation is that in many cases only part of the yield of the n th farm is supplied
 22 to the biorefinery. Also notice that weighted distances between farms and the candidate refinery
 23 are used in Eq. (1). This is because on different types of roadways the vehicular operational
 24 characteristics are quite different, which results in different costs for the same distance traveled.
 25 The use of the weighted distances in Eq. (1) enables the developed tool to model the switchgrass
 26 transportation cost more accurately. u in Eq. (1) can be ignored, as it does not affect the final
 27 results.

28 29 3.2. Modeling Multi-Biorefinery Problem

30 For multi-biorefinery scenario, the problem is formulated as a mixed-integer linear programming
 31 (MILP) model shown in Eq. (2). The solution to this MILP model will answer how many
 32 candidate biorefineries should be selected and how the produced biomass feedstock should be
 33 allocated among the selected biorefineries to minimize the total transportation, construction, and
 34 operations cost. The multi-biorefinery scenario is essentially a location-allocation problem (7).

$$35 \quad \text{Min} \quad \sum_j C C_j x_j + u \sum_i \sum_j y_{ij} P_i d_{ij} \quad (2)$$

36

37 **s.t.**

38 $y_{ij} - x_j \leq 0, \quad \forall j$

39 $\sum_i y_{ij} P_i - C_j x_j = 0, \quad \forall j$

$$\begin{aligned}
1 & \quad \sum_j y_{ij} \leq 1, \quad \forall i \\
2 & \quad \sum_j C_j x_j \geq T_C \\
3 & \quad x_j = \{0,1\} \\
4 & \quad 0 \leq y_{ij} \leq 1, \quad \forall i, j \\
5 & \quad i = 1, \dots, M \quad j = 1, \dots, N
\end{aligned}$$

6 where

7 CC_j = annualized construction and operational cost of candidate biorefinery j (\$);

8 x_j = 1 if candidate biorefinery j is selected (0 if not selected);

9 y_{ij} = percentage of feedstock produced by farm i that is transported to candidate
10 biorefinery j ;

11 T_C = total amount of feedstock required;

12 C_j = demand for feedstock at candidate biorefinery j (tons);

13 M = the total number of farms; and

14 N = the total number of candidate biorefineries.

15
16 The objective of the model in Eq. (2) is to minimize the biomass feedstock transportation
17 cost as well as the biorefinery construction and operations cost. The first set of constraints in Eq.
18 (2) is to ensure that if biorefinery j is not chosen, then y_{ij} must be zero. The second set of
19 constraint is to guarantee that the demand of each selected biorefinery must be satisfied. The
20 third set of constraints is to make sure that the total amount of feedstock shipped out from a farm
21 is less than what it produces. When planning for biorefineries, the total biofuel production
22 capacity is usually pre-specified and needs to be satisfied. Given this total production capacity
23 value, the minimum total amount of biomass feedstock (T_C) can easily be determined. To
24 satisfy the pre-specified biofuel production capacity is equivalent to ensuring that the summation
25 of biomass feedstocks required by each selected biorefinery is greater than T_C , and this is why
26 the last constraint is included.

27 Several unique features of the model in Eq. (2) distinguish it from some of the existing
28 biorefinery location models reviewed in this paper:

- 29
30 1. The developed model use road network distances between farms and candidate
31 biorefineries for calculating the transportation cost. Compared to using the distances of
32 straight lines linking farms and biorefineries, the road network distances can give
33 much better and more accurate transportation cost estimates.
- 34 2. Instead of dividing the entire study area into small pixels and treating each pixel as a
35 candidate location for biorefineries, users are given the flexibility to specify candidate
36 biorefinery locations. This can save considerable computation time by getting rid of
37 some obviously infeasible locations such as lakes, residential areas, and wildlife
38 refugees. Users can also specify as many candidate biorefineries as they want to cover
39 all pixels in the study area.
- 40 3. None of the existing studies has developed an explicit optimization model to find the
41 best locations for multi-biorefinery scenarios. Only empirical and heuristic methods

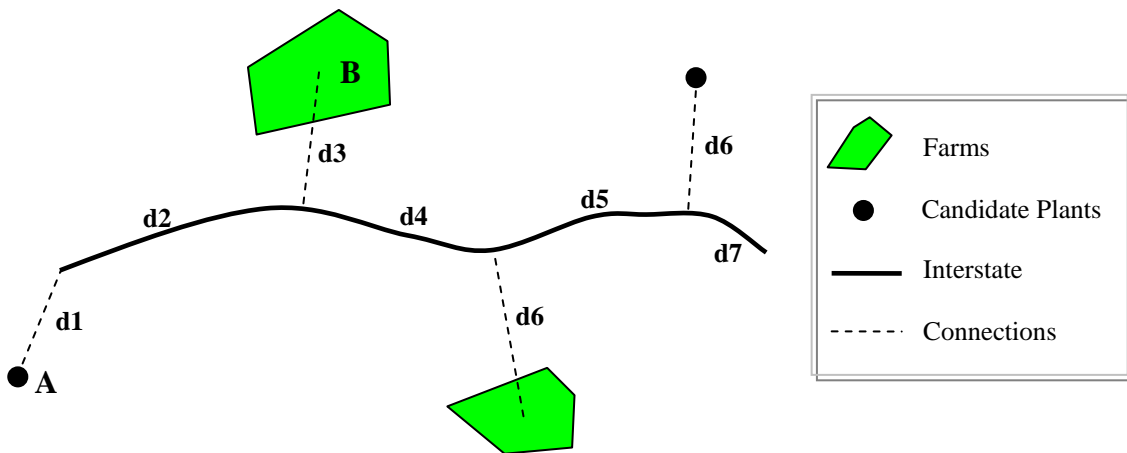
1 were used, and these methods cannot guarantee to find the global optimal solution. For
 2 the MILP model introduced in this study, a global optimal solution can always be
 3 reached given the necessary input data.
 4

5 **4. GIS TOOL DEVELOPMENT**

6 A GIS tool in the form of an ArcMap® toolbar is developed to facilitate the implementation of
 7 the MILP model. The GIS tool is developed using ArcObjects and the Visual Basic for
 8 Applications (VBA) environment provided by ArcMap. This tool is not an independent program
 9 and can only be used along with ArcMap. It has two major functions: (1) input data editing and
 10 (2) MILP input data preparation. Since the toolbar is built based on the ArcMap environment, all
 11 tools provided in ArcMap can still be used for managing the GIS input data, and this saves a lot
 12 of time and efforts to write codes for data management. Some tasks pertaining to this study
 13 such as the MILP input data preparation cannot be done easily using the tools provided in
 14 ArcMap. For this reason, several customized toolbar menus are created to
 15

- 16 1. **Specify locations of candidate biorefineries:** A simple tool is developed to allow
 17 users to specify the locations of candidate biorefineries on a point feature layer. With
 18 this tool activated, each click on this point feature layer will create a biorefinery at that
 19 location.
- 20 2. **Match plants and farms to road network:** Plants and farms may not be located
 21 exactly on the road network. To calculate the shortest distances between farms and
 22 candidate biorefineries, they must be connected to the road network. In this study, a
 23 shortest link between the farm/biorefinery and the road network is used to connect
 24 them. For each farm, a shortest link from its centroid is connected to the road network.
 25 Since candidate biorefineries are represented by point features, shortest links directly
 26 from those points to the road network are used. Several examples of the shortest links
 27 are shown in Figure 1. Such links are created automatically by a VBA code. For the
 28 example in Figure 1, the weighted shortest distance from biorefinery A to farm B is
 29 $d1*coe_connection+d2*coe_interstate+d3*coe_connection$.

30



31

32

Figure 1 Match farms and biorefineries to road network.

- 1
- 2 3. **Prepare input data for the MILP model:** Dijkstra's shortest path algorithm is used in
- 3 this study for finding the shortest distances between farms and biorefineries. To use the
- 4 dijkstra's method, matrices that store the intersection, road links, and link distances
- 5 information must be provided. Such information usually cannot be directly derived
- 6 from publicly available GIS data. A special ArcObjects and VBA program is
- 7 developed in this study to generate these data required by the dijkstra's algorithm.
- 8 4. **Run the single-biorefinery and multi-biorefinery optimization procedures:** The
- 9 optimization algorithms in Eqs. (1) and (2) are implemented using the ArcObjects and
- 10 VBA programming language. These two procedures take the attribute information
- 11 from each GIS input layer and also write the modeling results into the attribute table of
- 12 the candidate biorefinery layer.

14 5. CASE STUDY

15 The developed GIS tool and the single-biorefinery and multi-biorefinery algorithms are applied

16 to a four-county area in South Carolina, which includes Orangeburg, Calhoun, Dorchester, and

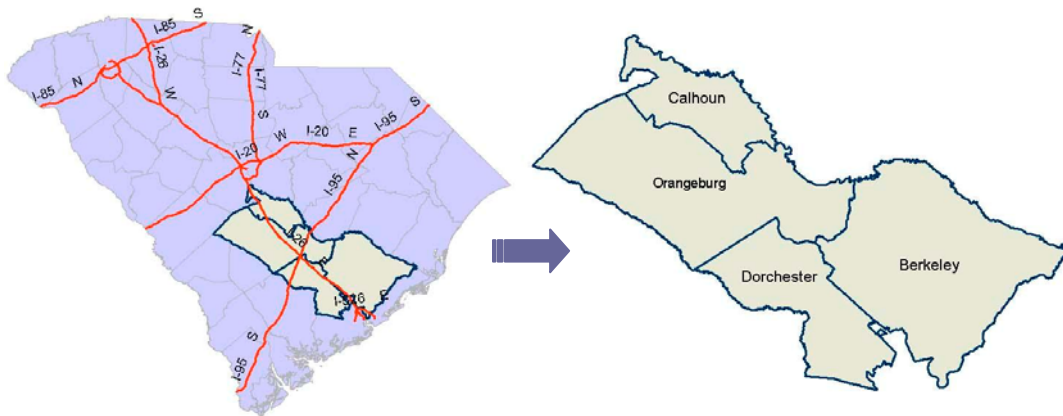
17 Berkeley. These counties are close to or part of the Interstate 95 (I-95) corridor in South Carolina.

18 The economy along this I-95 corridor is heavily dependent on traditional crops such as cotton

19 and tobacco. With the decline in market for cotton and tobacco, the corridor has become

20 economically depressed. The four-county area and the I-95 corridor are shown in Figure 2

21 below.



22

23 Figure 2 The four-county study area and the I-95 corridor.

24

25 Switchgrass can produce 540% more renewable energy than the energy required for the

26 production process (8). Unlike corn, using switchgrass for ethanol does not compete with food

27 supply. Moreover, switchgrass can grow almost everywhere. Growing switchgrass requires less

28 fertilizer and pesticide compared to growing corn, thus causing less impact on the environment.

29 Switchgrass is also very useful for soil conservation and amendment due to its deep and

30 extensive root system. The merits of switchgrass are not limited to producing ethanol.

31 Switchgrass has an excellent capability of storing carbon in its root system. It removes more

1 carbon from the atmosphere than is released by producing and using the switchgrass ethanol.
2 Ethanol production from switchgrass is thus called “carbon negative” (9).

3 South Carolina has abundant switchgrass resources along the I-95 corridor. However,
4 these resources have not been fully utilized. Areas along the I-95 corridor are still among the
5 poorest parts of the State. In the mean time, South Carolina consumes about 2.5 billion gallons
6 of gasoline every year, all of which is imported from other states and countries. Replacing
7 gasoline by ethanol produced from switchgrass thus appears to be a very promising strategy for
8 boosting the local economy and making South Carolina energy more independent. Assessing
9 the feasibility of building biorefinery plants in these counties is very important. The result of this
10 research could be used to make informed decisions that help people in these counties both
11 economically and environmentally.
12

13 5.1. GIS Data

14 Table 1 shows the GIS data that need to be collected or provided for applying the developed
15 model. The road network data (dataset 1) is used to calculate the shortest distances between
16 farms and candidate biorefineries. Data for all highways maintained by the South Carolina
17 Department of Transportation (SCDOT) are collected. Only highways within the four-county
18 area are used for this study. Furthermore, some roadways such as those unpaved segments are
19 not used for modeling. The resultant roadway segments are categorized into four groups:
20 interstate, primary road, secondary road, and ramp.

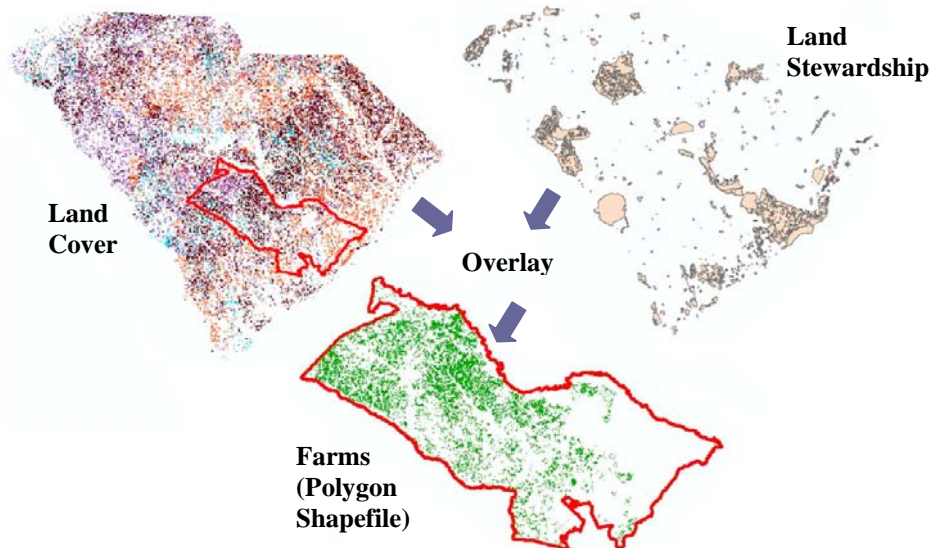
21 The collected land cover data are in raster format. For the original raster data, the entire
22 South Carolina is divided into approximately 173 million cells and each cell is 900 square meters.
23 Each pixel has a value indicating its land cover type. There are totally 28 different types of land
24 covers, including urban development, urban residential, grassland/pasture, swamp, etc. Pixels
25 with certain types of land covers, such as aquatic vegetation, are obviously not suitable for
26 growing switchgrass and are thus excluded from further consideration. Finally pixels with the
27 following five types of land covers are considered as potential sites for growing switchgrass,
28 which are dry scrub/shrub thicket, sandy bare soil, open canopy/recently cleared forest,
29 grassland/pasture, and cultivated land. The “clip” tool provided in ArcToolbox is used to remove
30 those pixels with other types of land covers from the original raster data.
31

32 Table 1 Collected data and sources.

Dataset	Type of Data	Data Source
1	Road Network	South Carolina Geographic Information Systems http://gis.sc.gov/data.html#DOT
2	Land Cover	South Carolina Department of Natural Resources http://www.dnr.sc.gov/GIS/gap/mapping.html
3	Land Stewardship	GIS Data Server at the University of South Carolina http://www.cas.sc.edu/gis/dataindex.html
4	County Boundary	GIS Data Server at the University of South Carolina http://www.cas.sc.edu/gis/dataindex.html

33 The land stewardship data are collected to identify lands that are reserved for various
34 purposes and cannot be used for growing switchgrass. For example, some lands are reserved by
35 the Department of Defense or Public Service Authority. The resulting land cover data from the
36 previous paragraph are overlaid with the land stewardship and the four-county boundary to
37

1 obtain a new land cover layer. Pixels in this layer are all suitable for growing switchgrass, not
 2 reserved by federal or local governments, and are within the four-county study area. By
 3 combining pixels with the same type of land cover, this layer is further converted into a polygon
 4 shapefile. The entire process of preparing the farm input data is illustrated in Figure 3.
 5

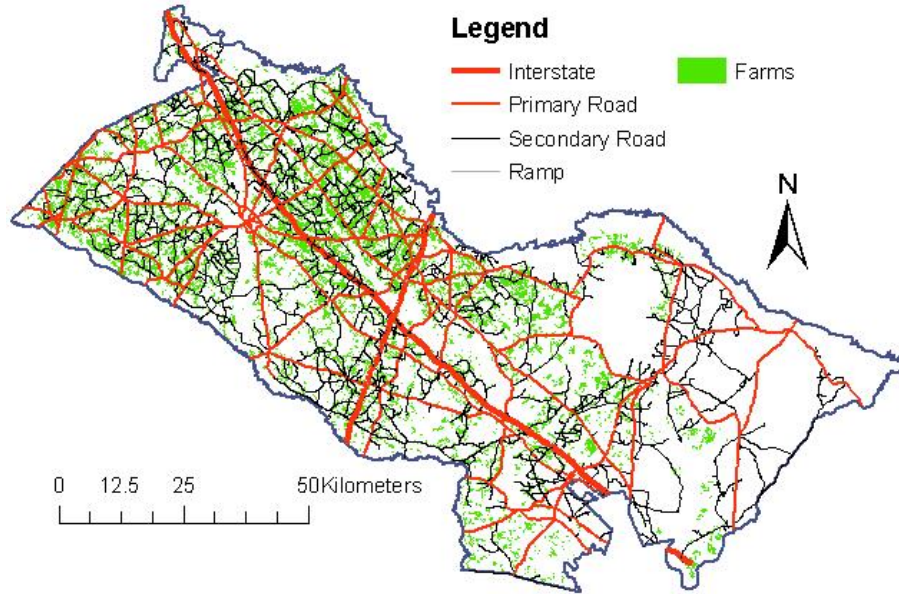


6
 7 **Figure 3 Preparation of farm input data.**
 8

9 **5.2. Other Input Data for MILP Model**

10 Other than the GIS data, the MILP model require several other inputs, including travel speeds on
 11 each type of roadways, switchgrass yields, construction and operations cost of each biorefinery,
 12 and capacity of each biorefinery. A recent research report released by the U.S. Department of
 13 Agriculture Iowa State office (10) shows that the average yield of switchgrass is between 4 to 7
 14 tons per acre. This result is based on extensive field tests in three states: Iowa, Illinois, and
 15 Missouri. Based on this result, the switchgrass yield rate is assumed to be 5.5 tons per acre (1
 16 acre is equal to 4,047 square meter) in this study.

17 As shown in Figure 4, there are four types of roadways considered in this study. For the
 18 collected road network data, there is no speed limits information available, and such information
 19 is not available in other publicly accessible road network GIS data either. In this study, the
 20 speeds on interstate, primary roads, secondary roads, and ramps are assumed to be 70 mph
 21 (112.7 km/h), 60 mph (96.6 km/h), 50 mph (80.5 km/h), and 30 mph (48.3 km/h), respectively.
 22 For the connection roads linking farms and candidate biorefineries to the road network, their
 23 speed limits are also assumed to be 30 mph (48.3 km/h). Based on the assumed speed limit data,
 24 the original length of each road segment is converted to an equivalent interstate length. The
 25 converted lengths are then used in the subsequent calculation of shortest distances. For example,
 26 a secondary road of length 10 miles (16.1 km) is converted to $70 \cdot (10/50) = 14$ miles (22.5 km).
 27



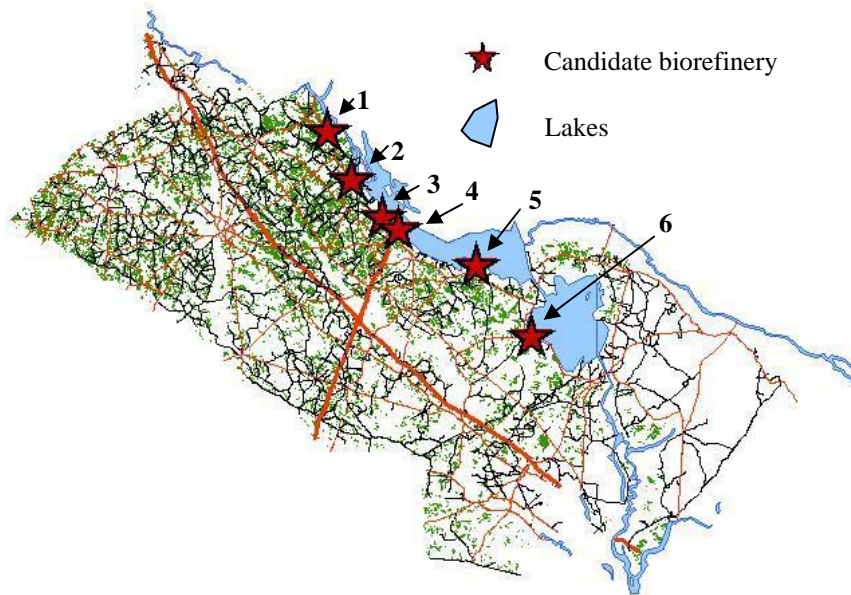
1
2
3
4
5
6
7
8
9
10

Figure 4 Input data for the four-county study area.

5.3. Model Results

Single-Biorefinery Scenario

In April 2008, the U.S. Department of Energy announced the selection of 3 small-scale biorefineries for federal funding (11). One of the three biorefineries is in Vonore, Tennessee. The total investment for this biorefinery is \$136 million. It requires 85 tons of switchgrass per day (equivalent to 31,025 tons annually) and can produce 2 million gallons of ethanol per year.



11
12

Figure 5 Prespecified candidate locations for biorefineries.

1 For this case study, it is assumed the same biorefinery plant is to be built in the four-county area.
 2 The developed GIS-based MILP model is used to find the best location for it. Since biorefineries
 3 require a substantial amount of water for producing ethanol, it is desirable to keep the
 4 biorefineries close to major surface water sources. In this case study, the six candidate
 5 locations in Figure 5 are prespecified.

6 As discussed in Section 3.1, for single-biorefinery location modeling, minimizing the total
 7 cost is equivalent to minimizing the total transportation cost. Furthermore, the transportation cost
 8 can be measured by ton-mileage instead of dollars. The flexibility provided here is very useful,
 9 especially when there is no reliable unit transportation cost estimate (\$/ton-mileage). In this
 10 study, the best biorefinery location is selected based on the ton-mileages of each candidate
 11 biorefineries. Based on the final output from the MILP model shown in Table 2, the optimal
 12 location for this case study is biorefinery 2 (see Figure 5). Data in Table 2 also suggest that if
 13 biorefinery 6 is selected instead of biorefinery 2, this will increase the total travel by 117,362
 14 ton-mileages.
 15

16 Table 2 MILP output for single-biorefinery scenario (One-way trip).

Rank	Candidate Biorefinery	Ton-Mileage	Average Miles/Ton	Ton-Kilometer
1	2	228,123	7.4	367,128
2	1	248,481	8.0	399,892
3	5	294,986	9.5	474,733
4	4	297,350	9.6	478,538
5	3	307,095	9.9	494,221
6	6	345,485	11.1	556,004

17

18 *Multi-Biorefinery Scenario*

19 The same six candidate biorefineries are considered for the multi-biorefinery scenario. Truck
 20 transportation cost was estimated by Levinson et al. (12) to be \$1.11/mile (\$0.69/km). Assume
 21 each truck can carry 20 tons of switchgrass and also take into account the return trip, the cost of
 22 truck transportation is assumed to be \$0.111 per ton per mile. The annualized construction and
 23 operations cost of each biorefinery is assumed to be \$10 million. Assume the goal is to produce 4
 24 million gallons of switchgrass; the total demand for switchgrass would be 62050 tons. Given all
 25 these inputs, the candidate biorefineries 1 and 2 (see Figure 5) constitute the optimal solution.
 26 The optimal objective function (Eq. 2) value is \$20,057,218. It can be seen that for this particular
 27 example, the transportation cost is relatively insignificant compared to the construction and
 28 operations cost. One reason is that this study considers all suitable cropland for growing
 29 switchgrass, and this can considerably reduce the transportation distances. Also, as the size of the
 30 biorefinery grows, the transportation distances typically increase at a much faster rate than the
 31 size of the biorefinery. By assuming different biorefinery sizes, the impact of the transportation
 32 cost will become more obvious.
 33

1 6. SUMMARY AND CONCLUSIONS

2 Many studies have attempted to investigate the optimal biorefinery location problem. These
 3 studies used either brutal search or heuristic algorithms to solve the single-biorefinery problem.
 4 For the multi-biorefinery problem, empirical methods, enumeration, and heuristic optimizations
 5 were used to find the best locations. These methods are easy to understand and implement.
 6 However, they are usually computationally intensive. More important, these empirical and
 7 heuristic methods cannot guarantee to find the global optimal solutions.

8 In this study, a GIS-based biorefinery location tool is developed. This GIS tool provides
 9 users more flexibility by allowing them to prespecify candidate biorefinery locations. In this way,
 10 considerable computation time can be saved. Another major advantage of doing this is that it
 11 allows users to better incorporate some important factors that are difficult to quantify but have
 12 significant impact on the biorefinery location decision making. A Mixed Integer Linear
 13 Programming (MILP) model is integrated into the developed GIS tool for solving the
 14 multi-biorefinery location optimization problem. The MILP model can jointly optimize the
 15 quantities and locations of biorefineries. Given biomass distribution, road network, and candidate
 16 biorefineries data, the MILP model can guarantee the globally optimal solution with the lowest
 17 total cost.

18 The developed GIS tool is applied to a four-county area in South Carolina to find the best
 19 locations for switchgrass biorefineries, and the results for single-biorefinery and
 20 multi-biorefinery scenarios are presented. For this case study, land cover and road network data
 21 are obtained from South Carolina Department of Natural Resources and South Carolina
 22 Department of Transportation. Other data, such as speed limits for each road segments, unit
 23 transportation cost (\$/ton-mileage), construction and operations cost of biorefineries, and
 24 switchgrass yield are assumed based on data in relevant documents. Some of the assumed
 25 numbers may not be very accurate. Nevertheless, the case study is used mainly for demonstrating
 26 the applicability of the GIS tool and the MILP model for biorefinery location optimizations. By
 27 supplying accurate and real (instead of the hypothetical data used in the case study) input data,
 28 this GIS-based MILP model should be able to provide more informative and useful results.

29

30 7. FUTURE WORK

31 Similar to other biorefinery location models reviewed in this study, the model developed in
 32 Section 3 uses deterministic input data and also generates a single output solution. However, for
 33 practical applications there are a lot of uncertainties involved in the biorefinery location
 34 modeling. For example, the yield of bioenergy crops such as switchgrass depends on many
 35 factors, including varieties, soil, and rainfall. All these factors are difficult to control and vary
 36 year by year and also from locations to locations. It is desirable to incorporate such uncertainties
 37 into the biorefinery location modeling and to see how the variations in the input data can affect
 38 the model output. In this case, the optimization objective is to minimize the expected total
 39 transportation and construction costs as shown in Eq. (3).

40

$$41 \quad \mathbf{Min} \quad E \left[\sum_j CC_j x_j + u \sum_i \sum_j y_{ij} P_i d_{ij} \right] \quad (3)$$

42

1 The yield of biomass under different conditions can be considered as a random variable.
 2 By collecting enough biomass yield data under different scenarios, an empirical probability
 3 density function can be estimated for this random variable. Given this probability density
 4 function, the optimization problem in Eq. (3) can be solved empirically by using a simulation
 5 method. Assuming that N candidate biorefinery locations are specified by users, a single run of
 6 the optimization model in Eq. (2) will generate a vector S as in Eq. (4). The length of this vector
 7 is N . Each element in S represents whether the corresponding candidate biorefinery is selected
 8 (=1) or not (=0). The basic idea of the simulation method is to run the optimization model in Eq.
 9 (2) K times, each time with different biomass yield data generated by the estimated probability
 10 function. In this way, K solutions can be generated. By averaging the K solutions, a solution
 11 vector \bar{S} can be obtained and Eq. (5) shows a simple example of \bar{S} . Intuitively, values in Eq.
 12 (5) represent the preferences for each candidate biorefinery to be selected. The final model
 13 output can be obtained by sorting the candidate biorefinery locations in terms of the preference
 14 values in descending order and select the first n candidate biorefineries that meet the minimum
 15 feedstock demand T_C in Eq. (2).

$$S = [0 \ 1 \ \dots \ 1 \ 0 \ 0 \ \dots \ 0]_{1 \times N} \quad (4)$$

$$\bar{S} = \frac{1}{K} \sum_{i=1}^K S_i = [0.11 \ 0.97 \ \dots \ 0.87 \ 0.13 \ 0.05 \ \dots \ 0.01]_{1 \times N} \quad (5)$$

20
 21 The empirical method just introduced is straightforward and easy to implement. However,
 22 it cannot guarantee the global optimal solutions for the biorefinery model that take uncertainties
 23 into consideration. Additional research is needed to develop new methods to solve Eq. (3). A
 24 good starting point to solve this problem is (13).
 25

26 8. ACKNOWLEDGMENTS

27 The authors would like to thank the James E. Clyburn University Transportation Center at South
 28 Carolina State University for supporting this research.
 29

30 REFERENCES

- 1 United States Environmental Protection Agency. Renewable Fuel Standard Program. Available at <http://www.epa.gov/OTAQ/renewablefuels/index.htm>. Accessed on July 26, 2009.
- 2 Graham, R.L., B.C. English, C.E. Noon. A Geographic Information System-based modeling system for evaluating the cost of delivered energy crop feedstock. *Biomass and Bioenergy*, Vol. 18, No. 5, 2000, pp. 309-329.
- 3 Ravula, P.P. Design, Simulation, Analysis and Optimization of Transportation System for a Biomass to Ethanol Conversion Plant. Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2007..

- 4 Dubuc, S. GIS-based accessibility analysis for network optimal location model: An application for bio-energy projects of the mineral industry in the UK. 2007, Online at <http://www.cybergeo.eu/pdf/12653>. Accessed on July 26, 2009.
- 5 Celli, G., E. Ghiani, M. Loddo, F. Pilo, S. Pani. Optimal location of biogas and biomass generation plants. In *Proceedings of the 43rd International Universities Power Engineering Conference*, Padova, Italy, 2008, pp. 1-6.
- 6 Panichelli, L., and E. Gnansounou. GIS-based approach for defining bioenergy facilities location: A case study in Northern Spain based on marginal delivery costs and resources competition between facilities. *Biomass and Bioenergy*, Vol. 32, No. 4, 2008, pp. 289-300.
- 7 Francis, F.L., L.F. McGinnis, and J.A. White. *Facility layout and location: an analytical approach*. second edition, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1992.
- 8 Ritter, S. K. Ethanol from Switchgrass Deemed Feasible. <http://pubs.acs.org/cen/news/86/i02/8602news3.html>. Accessed on July 18, 2009.
- 9 Fuels Made From Prairie Biomass Reduce Atmospheric Carbon Dioxide. <http://www.sciencedaily.com/releases/2006/12/061207161136.htm>. Accessed on July 18, 2009.
- 10 Kaiser, J, S. Bruckerhoff. Switchgrass for biomass productivity by variety selection and establishment methods for Missouri, Illinois, and Iowa. 2009. Available online at <http://efotg.nrcs.usda.gov/references/public/IA/AgronomyTechNote35.pdf>. Accessed on July 30, 2009.
- 11 U.S. Department of Energy (USDOE). DOE Selects 3 Small-Scale Biorefinery Projects for up to \$86 Million of Federal Funding in Maine, Tennessee and Kentucky. 2008. Available online at <http://www.energy.gov/news/6164.htm>. Accessed on July 28, 2009.
- 12 Levinson, D., M. Corbett, M. Hashami. Operating costs for trucks. Available online at <http://nexus.umn.edu/Papers/TruckOperatingCosts.pdf>. Accessed on July 28, 2009.
- 13 Balachandran, V, and S. Jain. Optimal Facility Location under Random Demand with General Cost Structure. *Naval Research Logistics Quarterly*, Vol. 23, No. 3, 1976, pp. 421-436.