A Two-Stage GIS-Based Suitability Model for Siting Biomass-to-Biofuel Plants and its Application in West Virginia, USA

Jinzhuo Wu
Jingxin Wang*
Michael P. Strager

Abstract

Woody biomass has been considered of low value because the cost of removal generally exceeded market price. New, valued-added markets to offset removal costs are necessary for utilization to be effective. In recent years the use of biomass as feedstock for biofuel production in the United States has been on the rise. A variety of liquid fuels can be produced from woody biomass; ethanol is one of the most promising. This study presents a two-stage approach to selecting woody biomass-based biofuel plants using Geographical Information System (GIS) spatial analysis and the multi-criteria analysis ranking algorithm of compromise programming. Site suitability was evaluated to minimize direct cost for investors and potential negative environmental impacts. The first step was to create a site suitability index using a linear fuzzy logic prediction model. The model involved 15 variables in three factor groups: (1) general physical conditions, (2) costs, and (3) environmental factors. The weights of the cost factors were determined using pairwise comparisons in the Analytical Hierarchy Process (AHP). The value of site suitability was reclassified into three categories (non-suitable, low-suitable, and high-suitable) using different classification methods. With a feasible plant location defined as an industrial site within the most suitable area, the second stage of the analysis used compromise programming to compare the potential sites. The criteria used to rank the potential sites included fuzzy distance to woody biomass, highways, railways, commercial airports, communities, and available parcel size. The AHP was used to compute the relative importance of each criterion. The top ten suitable sites were determined, and sensitivity analyses were conducted to derive the most preferred sites. The approach was successful in taking a large amount of non-commensurate spatial data and integrating a site-based ranking algorithm to find the top locations for biomass plants. It also has great potential and applicability to other suitability and site selection studies.

Keywords: multi-criteria analysis (MCA), Geographical Information System (GIS), compromise programming, biofuel, site selection.

Introduction

Interest in the use of cellulosic biomass as feedstock for bioenergy or biofuels in the United States in order to reduce dependence on fossil fuels has been increasing. As one of the larger unexploited sources of cellulosic biomass, woody biomass is identified as a potentially important feedstock for biofuels (Perlack et al. 2005). Of the variety of liquid fuels that can be produced from woody biomass (Zerbe 1991), ethanol is one of the most promising products (Badger 2002, Sun and Cheng 2002, Hamelinck et al. 2003). Site selection for a woody biomass-based ethanol plant is one of the major steps in the plant design process. The process itself can have significant impacts on the viability and profitability of the facility. The goal of this process is to find the optimum location, one that can minimize direct cost for investors and social cost for a local community. In the short run, the direct cost of developing an ethanol facility includes excavation cost of the site, capital costs, and expenses associated with purchasing land resources. In the long run, direct costs include feedstock cost, energy cost, and transportation cost. Woody biomass-based ethanol plants could benefit from being near raw materials and highways/railways to keep transportation costs low (Hamelinck et al. 2003). Locating the ethanol plant near existing industrial or power-generating facilities is an efficient method of lowering energy expenses. Favorable sites will also have sufficient water supply, sewer

The authors' affiliations are: Northeast Forestry University, College of Engineering and Technology, Harbin, China, jinzhuo2009@gmail.com; West Virginia University, Division of Forestry and Natural Resources, Division of Resource Management, Morgantown, WV, jxwang@wvu.edu; mstrager@wvu.edu.

* Corresponding author: Jingxin Wang. Tel: (304) 293-7601. Fax: (304) 293-2441. This manuscript is published with approval of the Direction of West Virginia Agric. Forestry Experimental Sta. as Scientific Paper No. 3127.

International Journal of Forest Engineering
Volume 22, No. 2

treatment, and highly qualified and suitable workforce. The site size should be large enough for truck traffic when moving feedstock and performing product storage. Finally, a properly located plant should also minimize environmental impacts (i.e., outside of wetland and endangered species areas) and reduce potential problems (dust, noise, light pollution, etc.) for the community and residents who live in proximity to the plant (USDA 2006).

Several conflicting and contradicting interests exist among economic, environmental and social criteria that are part of the decision-making process of site selection. Multi-criteria decision analysis (MCDA) can be used to evaluate these interests among different stakeholders. The most common procedure for multi-criteria evaluation is Weighted Linear Combination (WLC). The determination of criterion weights is a very important part of this method. MCDA provides four methods for assessing criterion weights: ranking, rating, pairwise comparison, and trade-off analysis. According to Malczewski (1999), "Which method to use depends on the trade-offs one is willing to make between ease of use, accuracy, the degree of understanding on the part of the decision maker, and the theoretical foundation underlying a given method, …". The pairwise comparison technique, developed by Saaty in the 1970s and 1980s in the context of Analytical Hierarchy Process (AHP), is a preferred method of calculating criteria weights because of its stronger theoretical foundations as compared to point allocation and rank ordering methods (Malczewski 1999).

With advancements in technology, spatial models have become tools to aid in land suitability assessment (Campbell et al. 1992, Eastman et al. 1995, Vatalis and Manoliadis 2002, Strager and Rosenberger 2006, 2007, Vahidnia et al. 2009). Geographic Information Systems (GIS) are very efficient and effective tools for handling a large amount of spatial data. They provide capabilities for modeling, optimization, and simulation. The combination of GIS and MCDA has been proved to be an efficient way of assessing land suitability (Pereira and Duckstein 1993, Joerin et al. 2001, Gomes and Lins 2002). A suitability index can be calculated using GIS map algebra techniques throughout an entire study area, and a value that represents the relative usefulness for a particular land use can be generated for each cell or pixel. The weighted linear combination (WLC) is usually applied in the spatial modeling processes. The advantage of the classical WLC is that it is very straightforward. However, it is usually quantitative, and crisp input values are usually used in the modeling process, which do not reflect exactly real-life problems. To eliminate the explicit shortcoming of this method, Zadeh (1965) introduced the concept of fuzzy sets to model vagueness or uncertainty in the real world. Fuzzy logic and membership functions can be created to provide a way of obtaining conclusions from vague, ambiguous or imprecise information (Clementini et al. 1997). Fuzzilized values, instead of crisp inputs, can then be used in the spatial modeling process to derive more practical results.

Land-suitability assessment could provide a technical basis for the planning of facility locations at the regional level; however, it may not identify the best alternative from a set of potential alternatives. Regional planning agencies may divide the land into different parcels of various sizes on the basis of a regional development plan, but potential site alternatives for ethanol plants need to provide cost-effective infrastructure (i.e., road access, water, gas, power, and sewer), adequate parcel size, and other conveniences (USDA 2006). Compromise programming (CP), a ranking algorithm for multi-criteria evaluation, can be used to identify the best compromise solution from a set of potential alternatives (Zeleny 1973, 1974, Nirupama and Simonovic 2002, Manoliadis et al. 2007). Some examples of CP applications include preference ranking of irrigation technologies, natural resource management, site selection, portfolio selection, and others (Duckstein and Opricovic 1980, Goicoechea et al. 1982, Gershon and Duckstein 1983, Tecle and Yitayew 1990, Baliestero and Romero 1996, Manoliadis et al. 2007, Pantouvakis and Manoliadis 2008). For example, Goicoechea et al. (1982) used compromise programming to evaluate a set of water-quality management alternatives subject to multiple criteria. Manoliadis et al. (2007) presented a framework for site selection of construction temporary facilities. The CP method was used as a comprehensive tool to compare different facility alternatives. Pantouvakis and Manoliadis (2008) developed a model for borrow pit (BP) selection using compromise programming. They concluded that the CP approach is appropriate and valid for BP selection and may also be used for other multiple-objective construction-related site-selection problems. As with the GIS-based MCDA method, the relative importance of evaluation criteria in the CP model can be determined using AHP.

The goal of this paper was to integrate components of a two-stage spatial multi-criteria analysis using GIS/spatial analysis, AHP, and a CP model to perform site selections for woody biomass-based ethanol plants. The approach was applied as a case study in the central Appalachia hardwood region of West Virginia, USA. The results will be beneficial to further analysis of the economic feasibility of woody biomass-based ethanol plants in the region.

Materials and Methods

Description of Study Area

The study was carried out in the central Appalachian hardwood region found throughout the entire state of West Virginia (Figure 1), which has the third highest percentage of forested land in the United States. The state is geographically located between the 37° 10' to 40° 40' of north latitude and 77° 40' to 82° 40' of west longitude. The harvesting process in West Virginia yields approximately 2.19 million tonnes (dry weight) of wood residues annually, including logging residue, mill residue, urban trees, and pallet residue (Wang et al. 2006). A significant portion (68%) of mill residue was utilized because mill residues are clean, concentrated at specific locations and relatively homogeneous; however, most of the logging residue, the largest proportion of wood residues (56%) in the state, was still underutilized (Wang et al. 2006). The abundant woody biomass in the state could support several medium-sized (100-200 million litres/year) ethanol fuel plants, given the conversion rate of 290 litres per dry tonne of woody biomass (Wooley et al. 1999). These plants could provide a source of tax revenue to meet the ever-increasing demands for energy and help to create more job opportunities for the local community in West Virginia.

Figure 1. Study area-West Virginia.

First Stage Modeling – Site Suitability Analysis

In the first stage, a site suitability index was computed for the entire study area using a linear fuzzy-logic prediction model (Equation 1). To account for criteria measured at different scales, variables in the prediction model are standardized and transformed using fuzzy-logic membership functions so that a positive change in the value of a criterion is always associated with a positive change in the suitability or desirability of an outcome.

$$SSI = \sum \left(f_m w_m \right) * \prod b_n \qquad [1]$$

Where SSI is the site suitability index, f_m is the fuzzy value of criteria m, w_m is the weight of criteria m, b_n is the criteria score of constraint n (Boolean value), and Π is the product.

The variables that are inputs to assessing site suitability for woody biomass-based ethanol facilities may include general physical conditions (topography, elevation, land cover/land use), proximity to local infrastructure (highways, railways), utilities (electric power, water/sewer, natural gas), raw materials, and environmental factors. Delivery of the output of the plants to consumers is not considered in siting the ethanol plants because the products are assumed to be consumed locally and in neighboring states. In this study, 15 variables were grouped into three factors on the basis of their specific relationships with the assessment of land suitability, namely (1) general physical conditions, (2) costs and (3) environmental factors (Table 1). Site suitability was evaluated to minimize the direct cost for investors and potential negative environmental impacts. The variables in Group 2 were considered as evaluation criteria. Fuzzy-logic membership functions were constructed for these variables, which were expressed as distance metrics and then normalized on the basis of site preference (Table 1). Variables in Groups 1 and 3 were considered constraints. Boolean values (0 and 1) were assigned to the variables in Groups 1 and 3 based on the site preference or acceptable range, where 1 means suitable and 0 means non-suitable.

The AHP approach was used to determine the weights of the evaluation criteria, depended on the importance of each variable in comparison with others. In AHP, Saaty (1980)

Table 1. Site suitability criteria for woody biomass-based ethanol plants.

Factors	Attributes	Acceptable range	References
General physical conditions	Topography	Flat to slightly rolling topography. Slope gradient: 0-10 percent	Stans et al. (1969)
	Elevation	100-700 m above sea level	Jensen and Christensen (1986), Hendrix and Duckley (1992)
	Aspect (orientation)	Southern, eastern, or western aspects	
	Land cover/land use	Shrub land, pasture, grassland, row crops	Ready and Guignet (2010)
Costs	Distance from woody biomass sources	0-80,000 m	Bain et al. (2003)
	Distance from highways	10-3,200 m	Apawootichai (2001), Koikai (2008)
	Distance from railways	10-5,000 m	Koikai (2008)
	Distance from power lines	10-1,600 m	Koikai (2008), FPL (2009)
	Distance from water bodies	50-5,000 m	Apawootichai (2001), Koikai (2008)
	Distance from sewer treatment plants	0-1,600 m	Pueblo County Board (2010)
	Distance from communities	800-5,000 m	Apawootichai (2001)
Environmental impacts	Wildlife habitat/endangered species areas	Avoid land with more wildlife biodiversity or presence of endangered or threatened species habitat	Deloitte and Touche (2001)
	Flood plain	Avoid flood-prone area	Lee and Pitchford (1999)
	Wetland area	Outside wetland area	Deloitte and Touche (2001)
	Public land	Outside public land	Deloitte and Touche (2001)

suggested a scale of one to nine for making subjective pairwise comparisons, with 1 indicating equally important and 9, extremely important. Strager and Rosenberger (2006) suggested the use of a three-point weighting scheme based on the 9-point scale by Saaty (1980) to lessen the difficulty regarding separation of differences between preferences. Therefore, in order to score comparisons, we used 1 for equal, 3 for prefer, and 6 for greatly prefer. A consistency test was performed to assure that choices were not randomly entered. The judgment rule is that pairwise comparisons are consistent if the consistency ratio (CR) is less than 0.1; otherwise, the pairwise comparisons must be redone until the consistency condition is accomplished (Saaty 1980).

After running the suitability model in the context of GIS using map algebra techniques, the value of the site suitability index (ranging from 0 to 1) for each cell or pixel can be derived. Classification can be performed using various methods, depending upon the nature of the criteria to be evaluated, or according to the utility to be classified as a suitability surface. In this study, the site suitability index was reclassified into three categories (non-suitable, low-suitable, and high-suitable) using different methods of classification (Equal Interval, Quantile, and Natural Breaks (Jenks 1967)) in order to get more robust results. With the Equal Interval classification method, each class has the same range. With the Quantile classification method, each class has the same number of observations. The Natural Breaks classification method reduces the variance within classes and maximizes the variance between classes (Jenks 1967). Since the availability of utilities (electricity, gas, water, and sewer) and access to pre-existing infrastructure are the conditions necessary for startup and operation of plants, the alternative feasible woody biomass-based ethanol plants were selected from industrial sites within the highest suitable area.

Second-Stage Modeling - Comparing the Alternative Sites Using CP

At the second stage, potential suitable sites were ranked using the compromise programming method. A new set of evaluation criteria that incorporates both spatial and non-spatial data is extended from the first-stage modeling, including fuzzy distance to highways, railways, commercial airports, and communities; woody biomass; and available parcel size. As in the first stage, the weights of the criteria were determined using pairwise comparisons in AHP and a consistency test was performed.

An ideal solution for the compromise programming algorithm, as defined by Tecle and Yitayew (1990), is the vector of objective functions' values, $f^* = (f_1^*, f_2^*, \ldots, f_I^*,)$ where the individual maximum values for criterion i, f_i^*, and minimum or worst value for criterion i, f_i^{**}, are defined using Equations 2 and 3:

$$f_i^* = \text{Max}\,(f_{ij}),\ i=1,2,\ldots,I \text{ and } j=1,2,\ldots,J, \qquad [2]$$

$$f_i^{**} = \text{Min}\,(f_{ij}),\ i=1,2,\ldots,I \text{ and } j=1,2,\ldots,J, \qquad [3]$$

The L_p metric as a compromise solution with respect to P can be expressed as Equation 4:

$$\text{Min} \quad \left\{ L_p\left(A_j\right) - \left[\sum_{i=1}^{N} \left(W_i\right) \left[\frac{\left(f_i^* - f_{ij}\right)}{\left(f_i^* - f_i^{**}\right)} \right]^p \right]^{1/p} \right\} \qquad [4]$$

Where $L_p(A_j)$ is the distance metric, a function of the decision alternative A_j and the parameter p (Tecle and Yitayew 1990). W_i is the standardized form of the criterion weight, which represents the decision maker's relative preference among the criteria, while f_i^* and f_i^{**} represent the best and worst value for criterion i. The parameter p reflects the importance of the maximal deviation from the ideal point (Tecle and Yitayew 1990, Duckstein and Opricovic 1980), and it can be assigned a value of from zero to infinity. For $p = 1$ all deviations are weighted equally and $L_p(A_j)$ is called Manhattan metric. In the case of $p = 2$ each deviation is weighted in proportion to its magnitude and $L_p(A_j)$ is called the Euclidean metric. The greater the deviation, the greater the weight will be. In case of $p = \infty$ the result is a min-max problem in which the compromise solution minimizes the maximum difference between the ideal point and the solution with respect to all indicators. $L_p(A_j)$ now is the Tchebycheff metric. In this case, Equation [4] is transformed to form Equation 5:

$$\text{Min} \quad \left\{ L_p\left(A_j\right) - \left[Max\left(W_i\right) \left[\frac{\left(f_i^* - f_{ij}\right)}{\left(f_i^* - f_i^{**}\right)} \right] \right] \right\} \qquad [5]$$

In this study, the compromise programming Equation 4 will be run for parameter values of $p = 1$ and 2. In order to analyze the robustness of the results, we allowed the variable weights in the CP model to vary within an interval (0-0.3) and carried out 10 simulations to compute the probability of each alternative being the preferred one. The simulated weights ($W_i^{'}$) was changed on the basis of predetermined weights (W_i) as: $W_i^{'} = \dfrac{W_i + \delta_i}{1 + \delta}$, with $\delta = \sum_i \delta_i$, so that the sum of the weights continues to equal 1.0 (Escobar and Moreno-Jiménez 1997). The alternative with the lowest value for the L_p metric was the best compromise solution because it was the nearest solution with respect to the ideal point.

Data Source and Manipulation

Data Source and Analysis

The spatial and categorical data used in the study were collected from the West Virginia Division of Forestry (WVDOF 2006), Appalachian Hardwood Center (Bragonje et al. 2006), West Virginia Development Office (WVDO 2008), and West Virginia GIS Technical Center (WVGISTC 2010). The counties with woody biomass inventory greater than

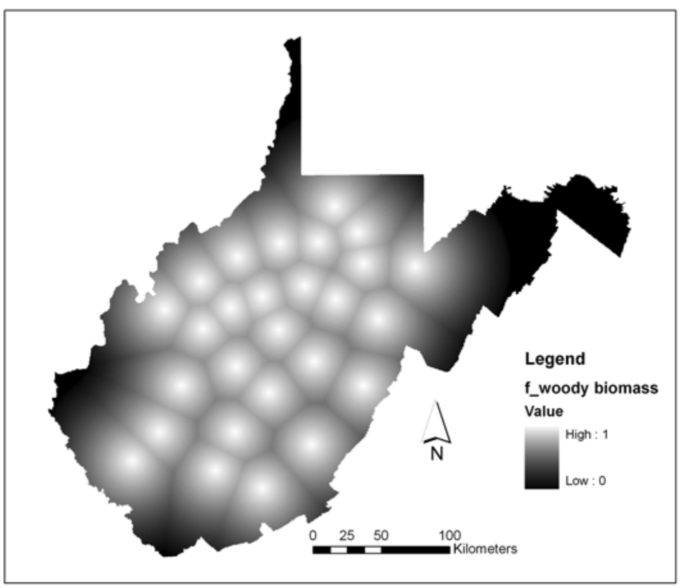
30,000 t.yr^{-1} were selected as raw material supply sources, which can support several medium-sized woody biomass-based ethanol facilities. As indicated in the Methodology section, three factor groups involving 15 variables were considered in this study, namely (1) general physical conditions, (2) costs, and (3) environmental factors.

To create distances associated with each of the variables in the cost group, fuzzy-logic membership functions were constructed based on the site preference (Table 1). GIS spatial analysis was used to calculate distance 'away from' each of the spatial features in Group 2 (woody biomass locations, highways, railways, power lines, water bodies, sewer plants, and communities) in meters. Then, the *con* function in ArcMap, which allows for 'if then' scenarios, was used to create the fuzzy membership functions. For example, the function of fuzzy distance to woody biomass sources would be:

Con([d_woodybiomass] > 80000, 0,
(80000 - [d_woodybiomass]) / 80000),

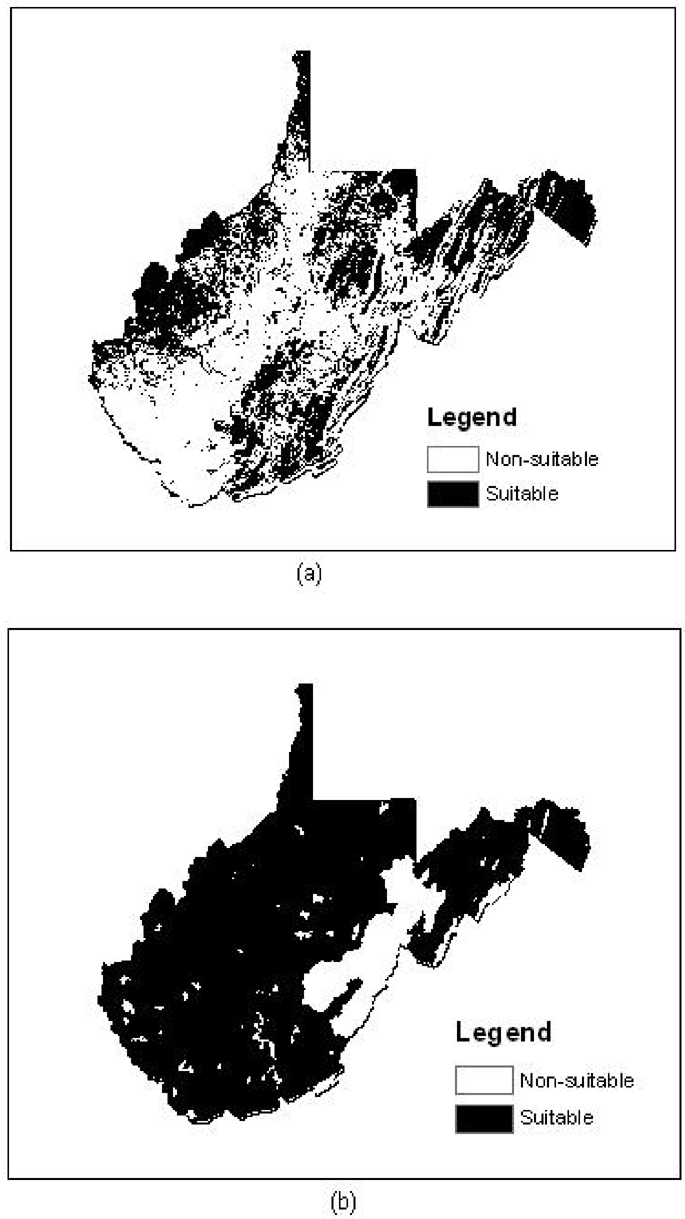
where *d_woodybiomass* is the distance away from woody biomass sources, and 80,000 is the maximum distance in meters that would be economically feasible for woody biomass delivery (Bain et al. 2003). The resulting fuzzy distance to woody biomass is shown in Figure 2.

Figure 2. Fuzzilized distance to woody biomass with 1 highly suitable and 0 non-suitable.



Boolean values (0 and 1) were assigned to the general physical conditions (topography, elevation, aspect, and land cover) and environmental factors (wildlife area, public land, flood plain, and wetland area). Topography (slope) and aspect were derived from the elevation dataset using GIS surface analysis and reclassified into two categories (suitable and non-suitable) on the basis of site preferences in Table 1. The resultant slope data were mapped in Figure 3a. Landcover categories that are suitable for plant construction include grassland/shrub and row crop agriculture. Reclassification of the landcover data is needed to create two classes: suitable and non-suitable. The spatial features (wildlife area, public land, flood plain, and wetland area) are in the shapefile format, which were converted to raster grid, with 1 indicating suitable area and 0, non-suitable area. Because wildlife area, public land, flood plain, and wetland area are not suitable for biomass plant siting, they will be defined as non-suitable area. Public land is mostly concentrated in the national forests and recreation areas located in eastern West Virginia. The raster data of public land is shown in Figure 3b.

Figure 3. Example of constraints (a) slope and (b) public land.



Weight Preference

A pairwise comparison matrix was created to reflect the preference or importance of the evaluation criteria in the suitability model (Table 2). The weights and tests for inconsistency were derived using the eigenvector method (Saaty 1994). The CR value (0.036) is lower than 0.1, and therefore the judgments are consistent. As expected, the criteria distance from woody biomass (0.360) was the most important factor, because shortening the hauling distance of woody biomass from the forest to the biomass-based ethanol plants would greatly reduce the transportation cost. A guaranteed supply of woody biomass at a competitive price within a reasonable radius of the plant is critical for the profitability and viability of the plant. The other criteria, such as distance from highways (0.205), distance from power lines (0.146), and distance from water bodies (0.130), also received higher attention during the process of site suitability evaluation. The criteria distance from railways (0.041) appears to be the factor of least importance. Compared to truck transportation, rail is more suitable for long hauling of bulk goods (Mahmudi and Flynn 2006). However, the dispersed nature of biomass requires that it start its transportation to a processing plant on a truck. Rail transshipment may be preferred in cases in which road congestion precludes truck delivery (Mahmudi and Flynn 2006).

Table 2. Pairwise comparison matrix and relative weights of the variables in the fuzzy logic prediction model.

Criteria	Woody biomass	Highways	Railways	Power	Water	Sewer	Community	Weight
Woody biomass	1.000 (0.393)	3.000 (0.495)	7.000 (0.304)	3.000 (0.429)	3.000 (0.333)	5.000 (0.263)	5.000 (0.306)	0.360
Highways	0.333 (0.131)	1.000 (0.165)	5.000 (0.217)	1.000 (0.143)	3.000 (0.333)	5.000 (0.263)	3.000 (0.184)	0.205
Railways	0.143 (0.056)	0.200 (0.033)	1.000 (0.043)	0.333 (0.048)	0.333 (0.037)	1.000 (0.053)	0.333 (0.020)	0.041
Power	0.333 (0.131)	1.000 (0.165)	3.000 (0.130)	1.000 (0.143)	1.000 (0.111)	3.000 (0.158)	3.000 (0.184)	0.146
Water	0.333 (0.131)	0.333 (0.055)	3.000 (0.130)	1.000 (0.143)	1.000 (0.111)	3.000 (0.158)	3.000 (0.184)	0.130
Sewer	0.200 (0.079)	0.200 (0.033)	1.000 (0.043)	0.333 (0.048)	0.333 (0.037)	1.000 (0.053)	1.000 (0.061)	0.051
Community	0.200 (0.079)	0.333 (0.055)	3.000 (0.130)	0.333 (0.048)	0.333 (0.037)	1.000 (0.053)	1.000 (0.061)	0.066

Note: ( ) is normalized value.

Results

Site Suitability

Using GIS spatial analysis, the site suitability index (*SSI*) in the study area was computed based on Equation 1, which ranged from 0 to 0.911703. The site suitability index was reclassified into three categories using three classification methods. If the Natural Breaks (Jenks 1967) classification method was used, the three categories would be: non-suitable (0-0.135331), low-suitable (0.135331-0.391747), and high-suitable (0.391747-0.911703). If the Equal Interval classification method was used, the groups would be: non-suitable (0-0.303901), low-suitable (0.303901-0.607802), and high-suitable (0.607802-0.911703). If the Quantile classification method was used, the three categories would be: non-suitable (0-0), low-suitable (0-0.331205), and high-suitable (0.331205-0.911703). The most suitable areas were found to be 0.19-1.68% of the total land area in all the scenarios. It was noticed that the quantile classification method provided more high-suitable area compared to the other two methods. As stated earlier, wood residue-based ethanol plants should be located in the most suitable area and have access to utilities (electricity, gas, water, and sewer) and pre-existing infrastructure. Currently, a total of 183 industrial sites are in West Virginia (WVGISTC 2010), of which 69 can provide all the required utilities. The potential industrial sites that fell within the most suitable areas identified by the quantile method are summarized in Table 3. Altogether 20 industrial sites fell within the highest suitable areas, which were selected as potential sites of woody biomass-based ethanol facilities.

Top-Ranked Sites

Compromise programming was applied to compare the suitability of these alternative sites derived from the first stage. The priorities of the variables in the CP model obtained from the pairwise comparison matrix were as follows: distance from highways-0.301, distance from railways-0.061, available parcel size-0.127, distance from commercial airports -0.059, distance from communities-0.083, and distance from woody biomass-0.370. The consistency ratio of the comparison was 0.054, less than 0.10; therefore the judgments were consistent and the weights can be used in the CP model.

A spatial analysis tool called zonal statistics was used to compute the average distances of each alternative site to features such as highways, railways, commercial airports, communities, and woody biomass. Zonal statistics calculates sta-

Table 3. Industrial sites that fell within the highest suitable area.

No.	Name	County	Electric	Gas provider	Water supplier	Sewer
1	Porter Farm Site	Harrison	AP	Dominion	Sun Valley PSD	Sun Valley PSD
2	Flatwoods-John Skidmore Development Site	Braxton	AP	Dominion	Flatwoods Canoe Run PSD	Flatwoods Canoe Run PSD
3	Suarez Site	Harrison	AP	Dominion	Anmoore city	Anmoore city
4	Mink Shoals Site	Kana-wha	APC	MGC	West Virginia American Water	Charleston city
5	Saltwell Road Site	Harrison	AP	Dominion	Bridgeport city	Bridgeport city
6	Ross Site No.1	Upshur	AP	MGC	Buckhannon city	Buckhannon city
7	Morris Farm Site	Braxton	AP	Dominion	Flatwoods Canoe-Run PSD	Flatwoods Canoe-Run PSD
8	Mark Carroll Site	Upshur	AP	MGC	Buckhannon city	Buckhannon city
9	Charton Management Site	Jackson	APC	CGUC	Jackson County PSD	Jackson County PSD
10	Ross Site No.2	Upshur	AP	MGC	Buckhannon city	Buckhannon city
11	Lee's Hill Site	Wood	AP	Dominion	Claywood Park PSD	Claywood Park PSD
12	Grafton County Club Road Site #1	Taylor	AP	EGC	Grafton city	Grafton city
13	Parsons Site	Jackson	APC	On site	Jackson County PSD	Jackson County PSD
14	Bosley Site	Wood	AP	Dominion	Parkersburg city	Parkersburg city
15	Pettyville Site	Wood	AP	Dominion	Mineral Wells PSD	Mineral Wells PSD
16	Henderson Site	Putnam	APC	MGC	Hurricane city	South Putnam PSD
17	Mid-Ohio Valley Regional Airport South Ramp	Wood	AP	Dominion	Union Williams PSD	Union Williams PSD
18	Deerfield Site 1	Mason	APC	MGC	Point Pleasant city	Mason County PSD
19	Thompson Site	Mason	APC	MGC	Point Pleasant city	Mason County PSD
20	Borgman Site	Preston	AP	MGC	Kingwood city	Kingwood city

[a]. The orders of the sites do not reflect the site preference.
[b]. All the sites have 0% flood plain; AP-Allegheny Power; APC-Appalachian Power Company; MGC - Mountaineer Gas Company; CGUC - Consumer Gas Utility Company; EGC - Equitable Gas Company.
[c]. Source: WVGISTC (2010).

tistics on values of a raster grid within another zone dataset (ESRI 2010). Here, the zone dataset will be the potential sites, the zone field is the unique site name, and the value raster is the straight distance to highways, railways, commercial airports, communities or woody biomass supply sources. The available parcel size of the sites was obtained from West Virginia GIS Technical Center. The values for each alternative site were normalized, and the best and worst values for each evaluation criteria were determined. Next, the CP Equation 4 was run for parameter values of $p = 1$ and 2 to offer a level of sensitivity analysis as suggested by Tecle and Yitayew (1990). The alternative sites were ordered based on the L_p metric (from low to high) for each run of $p = 1, 2$ (Table 4). It was noted that the ranking of the sites changed with respect to different p values. The final ranks of these sites were determined by summing all the ranks together, and the top five ranked potential locations in West Virginia were: Porter Farm Site (Site 1) (Harrison County), Saltwell Road Site (Site 5) (Harrison County), Flatwoods - John Skidmore Development Site (Site 2) (Braxton County), Suarez Site (Site 3) (Harrison County), and Ross Site No. 1 (Site 6) (Upshur County). The locations of the top 10 ranked sites are shown in Figure 4.

In order to analyze the robustness of the results, we allowed the variable weights in the CP model to vary within an interval (0-0.3) and carried out 10 simulations to compute the probability of preference for each alternative. The simulation results with respect to different p values were summarized in Table 5. Site 1 (Porter Farm Site) and Site 5 (Saltwell Road Site) were the most preferred when the value of p equaled 1 and 2, respectively.

Table 4. Industrial sites ranking.

Site No.	Rank when p=1	Rank when p=2	Sum rank	Final rank
1	1	2	3	1
2	2	5	7	3
3	3	4	7	3
4	4	6	10	6
5	5	1	6	2
6	6	3	9	5
7	7	10	17	8
8	8	7	15	7
9	9	9	18	9
10	10	8	18	9
11	11	13	24	11
12	12	15	27	13
13	13	14	27	13
14	14	11	25	12
15	15	12	27	13
16	16	19	35	17
17	17	16	33	16
18	18	17	35	17
19	19	18	37	19
20	20	20	40	20

Figure 4. Top industrial site locations.

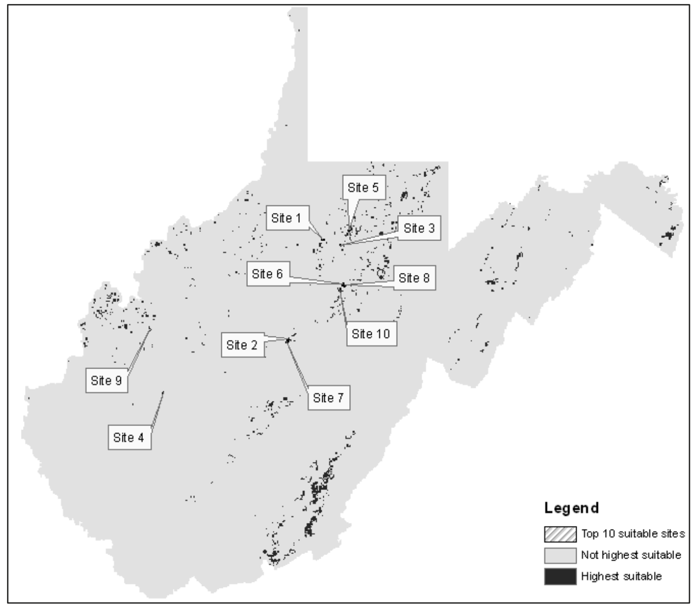


Table 5. Statistics of L_p metric with respect to different p values.

Site No.	p=1 Mean	p=1 Variance	p=1 Prob. Most pref.	p=2 Mean	p=2 Variance	p=2 Prob. Most pref.
1	0.1491	0.001301	1	0.3219	0.002287	0
2	0.1703	0.003214	0	0.3613	0.004427	0
3	0.1852	0.001635	0	0.3368	0.002490	0
4	0.2145	0.001697	0	0.3717	0.002813	0
5	0.2166	0.000682	0	0.2884	0.000711	1
6	0.2380	0.000943	0	0.3245	0.001140	0
7	0.2796	0.003325	0	0.4659	0.003931	0
8	0.2588	0.001569	0	0.3786	0.002025	0
9	0.2768	0.001893	0	0.4262	0.002658	0
10	0.2844	0.001324	0	0.3878	0.002047	0
11	0.3692	0.002464	0	0.4935	0.001849	0
12	0.3672	0.001511	0	0.5394	0.001254	0
13	0.4152	0.001218	0	0.5335	0.001063	0
14	0.4002	0.001603	0	0.4760	0.001377	0
15	0.4309	0.001469	0	0.4923	0.001355	0
16	0.5055	0.004031	0	0.6883	0.002107	0
17	0.5109	0.002282	0	0.6042	0.001697	0
18	0.5974	0.000567	0	0.6457	0.000334	0
19	0.6271	0.000378	0	0.6680	0.000217	0
20	0.6563	0.001804	0	0.7350	0.000793	0

Summary and Discussion

This study has presented a two-stage approach to selecting a woody biomass-based ethanol plant location using GIS spatial analysis and compromise programming. The land use suitability was evaluated to minimize the direct cost for investors and the potential negative environmental impacts. A fuzzy-logic prediction model incorporating AHP was used to predict the site suitability index of every single cell. Industrial sites that could provide utilities (electricity, gas, water, and sewer) and easy access to pre-existing infrastructure, and which were within the most suitable area, were selected as potential sites for 100-200 million $l.yr^{-1}$ woody biomass-based ethanol facilities. The compromise programming algorithm was then applied to rank the potential suitable sites identified from the first stage. This methodology incorporated a large number of economic and environment related factors that are essential to identifying suitable sites. Decision rules for locating suitable sites for woody biomass-based ethanol plants were key to the success of the application. In fact, many other factors could be involved in the site selection process to enhance the robustness of the results, but the most important factors were taken into consideration in the study.

The approach was applied in West Virginia, United States, and was successful in taking a large amount of non-commensurate spatial data and integrating a regional and a local scale site-based ranking algorithm to find the top locations for biomass plants. It also has great applicability to other suitability and site selection studies. In order to test the robustness of the results, sensitivity analyses were conducted, and the results showed that Site 1 (Porter Farm Site) and Site 5 (Saltwell Road Site) were the most preferred when decision makers have different concerns about the maximum deviation from ideal points (p = 1, 2). Because site selection is a complex decision-making process, it is very possible that the preferences of the evaluation criteria in the suitability model or CP model may change significantly, different stakeholders (investors, local residents, and environmentalists) may be involved in this decision process and most likely will have different interests. Strager and Rosenberger (2006) recommended accommodating different and potentially conflicting preferences among groups by statistically validating the within, and across-group preferences.

A biofuel plant would have social, economic, and environmental impacts on the sites in the study region. Future research will be needed to evaluate the economic viability and sustainability of the plants in these top-ranked sites. For example, it is critical to stabilize the feedstock (wood residues) supply for a woody biomass-based biofuel plant. Forest growth in the state of West Virginia continues to outpace removals at nearly a 1.77 to 1 ratio, which indicates that the forestlands are currently managed sustainably (Oswalt and Turner 2009). The study region can produce high quantities of woody biomass and residues on an annual basis, which can assure the long-term raw material supply for the biofuel plant. The local-level economic impacts (jobs, wages, income) and environmental impacts (land use change, water quality) of the biofuel facility could be analyzed using life cycle analysis (LCA). The LCA will also help us understand and further explore the opportunities for reducing carbon emissions and evaluate whether the biofuel plant will result in net carbon storage or carbon generation.

The framework provided in this study provides a viable screening and ranking approach to finding a woody biomass-based ethanol plant location. While there are limitations with the criteria and potential subjectivity with the preference weights, by varying both and mapping the change, we can indicate which areas are least spatially sensitive.

Acknowledgements

This research was funded by a grant from the West Virginia EPSCOR Agreement No. EPS2006-35. The authors would like to thank the editors and the anonymous reviewers for their insightful comments and suggestions.

Literature Cited

Apawootichai, S. 2001. Inclusion of environmental criteria for light industrial estate site selection in Supanburi Province. Master Degree of Philosophy in Environment. The Joint Graduate School of Energy and Environment, King Mongkut's University of Technology Thonburi, Bangkok.

Badger, P.C. 2002. Ethanol from cellulose: a general review. In: Janick J. and Whipkey A., (eds.), Trends in New Crops and New Uses; ASHS Press, Alexandria, VA. 17-21 pp.

Bain, R., W. Amos, M. Downing, and R. Perlack. 2003. Biopower technical assessment: state of the industry and the technology. NREL Rep. No. TP-510-33123. Golden, CO: National Renewable Energy Laboratory. 277 p.

Baliestero, E. and C. Romero. 1996. Portfolio selection: A compromise programming solution. The Journal of the Operational Research Society. 47(11): 1377-1386.

Bragonje, R., S. Grushecky, B. Spong, B., J. Slahor and L. Osborn. 2006. West Virginia wood byproducts available and needed. West Virginia University, Morgantown, WV. P. 10-33.

Clementini, E., P.D. Felice and D. Hernández. 1997. Qualitative representation of positional information. Artificial Intelligence. 95: 317-356.

Campbell, J.C., J. Radke, J.T. Gless and R.M. Wirtshafter. 1992. An application of linear programming and geographic information systems: cropland allocation in Antigua. Environment and Planning A. 24: 535-549.

Deloitte and Touche. 2001. Industrial/Business park standards: I-80 corridor region. Accessed Sept. 22, 2010. http://www.neded.org/files/businessdevelopment/library/i80region.pdf.

Duckstein, L. and S. Opricovic. 1980. Multiobjective optimization in river basin development. Water Resources Research. 16 (1): 14-20.

Eastman, J.R., W. Jin, P.A.K. Kyem and J. Toledano. 1995. Raster procedures for multi-criteria/multi-objective decisions. Photogrammetric and Remote Sensing. 61: 539-547.

ESRI. 2010. ArcGIS 9.2 Desktop Help. Accessed Oct. 8, 2010. http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=welcome.

Escobar, M. and J. Moreno-Jiménez. 1997. The Hierarchical compromise programming. Top. 5: 253-281.

FPL. 2009. Electric service standards. Accessed Sept. 18, 2010. http://nexteraenergyresource.com/doingbusiness/builder/pdf/Ess13Safety.pdf.

Gershon, M. and L. Duckstein. 1983. Multiobjective approaches to river basin planning. *Journal of Water Resources Planning and Management*. 109: 13-28.

Gomes, E.G. and M.P.E. Lins. 2002. Integrating Geographical Information Systems and multi-criteria methods: A Case Study. Annals of Operations Research, Amsterdam. 116 (1): 243–269.

Goicoechea, A., D.R. Hansen and L. Duckstein. 1982. Multi-objective Decision Analysis with Engineering and Business Applications. John Wiley and Sons, New York, NY. pp. 326-360.

Hamelinck, C.N., G.V. Hooijdonk, and A.P.C. Faaij. 2003. Prospects for ethanol from lignocellulosic biomass: techno-economic performance as development progresses. Accessed Dec. 20, 2009. http://www.senternovem.nl/mmfiles/149043_tcm24-124362.pdf.

Hendrix, W.C. and D.J.A. Duckley. 1992. Use of a geographic information system for selection of sites for land application of sewage waste. *Journal of Soil and Water Conservation*. 47(3):271-275.

Jenks, G.F. 1967. The data model concept in statistical mapping. In: International Cartographic Association, ed. International Yearbook of Cartography. University of Ulm, Ulm, Germany. 7: 186-190.

Jensen, J.R. and E.J. Christensen. 1986. Solid and hazardous waste disposal site selection using digital geographic information system techniques. The Science of the Total Environment. Elsevier Publishing Company, Amsterdam. 56: 265-276.

Joerin, F., M. Thériault and A. Musy. 2001. Using GIS and outranking multicriteria analysis for land-use suitability assessment. *International Journal of Geographical Information Science*. 15: 153-174.

Koikai, J.S. 2008. Utilizing GIS-based suitability modeling to assess the physical potential of bioethanol processing plants in Western Kenya. Accessed June 25, 2010. http://www.gis.smumn.edu/GradProjects/KoikaiJ.pdf.

Lee, V.W. and G.E. Pitchford. 1999. Creating industrial sites. Accessed Sept. 20, 2010. http://www.aces.edu/crd/publications/CRD-14.html.

Malczewski, J. 1999. GIS and Multicriteria Decision Analysis. John Wiley and Sons, New York, NY, pp. 156-182.

Mahmudi, H. and P.C. Flynn. 2006. Rail vs. truck transport of biomass. Applied Biochemistry and Biotechnology. 129: 88-103.

Manoliadis, O., J. Pantouvakis and I. Tsolas. 2007. Compromise programming model in site selection for construction temporary facilities. Operational Research. 7: 381-400.

Nirupama, N. and S. Simonovic. 2002. A spatial fuzzy compromise approach for flood disaster management. The Institute of Catastrophic Loss Reduction, London, Ontario, Canada.

Oswalt, C.M. and J.A. Turner. 2009. Status of hardwood forest resources in the Appalachian region including estimates of growth and removals. USDA Forest Service, Southern Research Station, Resource Bulletin SRS -142.

Pantouvakis, J.P. and O.G. Manoliadis. 2008. A compromise programming model for site selection of borrow pits. Construction Management and Economics. 26: 433–446.

Perlack, R.D., L.L. Wright, A.F. Turhollow, R.L. Graham, B.J. Stokes and D.C. Erbach. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. U.S. Department of Energy and U.S. Department of Agriculture, Washington, DC. 60 p. http://feedstockreview.ornl.gov/pdf/billion_ton_vision.pdf .

Pereira, J.M.C. and L. Duckstein. 1993. A multiple criteria decision-making approach to GIS-based land suitability evaluation. *International Journal of Geographical Information Systems*. 7: 407-424.

Pueblo County Board. 2010. Local regulations of site selection and construction of major new domestic water and sewage treatment systems and major extensions of existing domestic water and sewage treatment systems. Accessed Sept. 23, 2010. http://www.codes.co.pueblo.co.us/data/title17/Chapter_17_164_SITE_SELECTION_AND_/17_164_010_Designation_.html.

Ready, R. and D. Guignet. 2010. Methodology for projecting land cover change in the CARA Region. Accessed Sep. 20, 2010. http://www.cara.psu.edu/about/LndUse_Methodology.pdf.

Saaty, T.L. 1980. The Analytic Hierarchy Process: Planning, Priority, Setting, Resource Allocation. McGraw-Hill, New York, NY. 437 p.

Saaty, T.L. 1994. Fundamentals of Decision Making and Priority Theory with the Analytical Hierarchy Process.RWS publications, Pittsburgh, PA. pp. 204-220.

Stans, M.H., R.C. Siciliano, and R.A. Podesta. 1969. How to make an industrial site survey, U.S. Government Printing Office, Washington, DC.

Strager, M.P. and R.S. Rosenberger. 2006. Incorporating stakeholder preferences for land conservation: weights and measures in spatial MCA. Ecological Economics. 57: 627-639.

Strager, M.P. and R.S. Rosenberger. 2007. Aggregating high-priority landscape areas to the parcel level: an easement implementation tool. *Journal of Environmental Management*. 82: 290-298.

Sun, Y. and J. Cheng. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresource Technology. 83(1): 1-11.

Tecle, A. and M. Yitayew. 1990. Preference ranking of alternative irrigation technologies via a multicriteria decision making procedure. American Society of Agriculture Engineers. 33(5):1509-1517.

U.S. Department of Agriculture (USDA). 2006. A guide for evaluating the requirements of ethanol plants. Accessed Aug. 18, 2010. http://www.ne-ethanol.org/industry/evalreq.pdf.

Vatalis, K. and O. Manoliadis. 2002. A two-level multicriteria DSS for landfill site selection using GIS: case study in western Macedonia Greece. *Journal of Geographic Information and Decision Analysis*. 6: 49-56.

Vahidnia, M.H., A.A. Alesheikh, and A. Alimohammadi. 2009. Hospital site selection using fuzzy AHP and its derivatives. *Journal of Environmental Management*. 90: 3048-3056.

Wang, J., S. Grushecky and J. McNeel. 2006. Biomass Resources, Uses, and Opportunities in West Virginia. Working Report, Biomaterials Center, Division of Forestry and Natural Resources, West Virginia University, Morgantown, WV.

West Virginia Division of Forestry (WVDOF). 2006. West Virginia Logging Sediment Control Act – 2006 Statistics. West Virginia Division of Forestry, Charleston, WV.

West Virginia Development Office (WVDO). 2008. Real estate. Accessed Jan. 20, 2010. http://www.wvdo.org/realestate/index.html.

West Virginia GIS Technical Center (WVGISTC). 2010. Digital elevation dataset (30-meter NED). Accessed July 21, 2010. http://wvgis.wvu.edu/index.php.

Wooley, R., M. Ruth, J. Sheehan, K. Ibsen, H. Majdeski and A. Galvez. 1999. Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis Current and Futuristic Scenarios. Research Report NREL/TP -580-26157, National Renewable Energy Laboratory, Golden, CO.

Zadeh, L.A. 1965. Fuzzy sets. Information and Control. 8: 338-353.

Zeleny, M. 1973. Multiple Criteria Decision Making. University of South Carolina Press, Columbia. 262-301 pp.

Zeleny, M. 1974. Linear Multi-objective Programming. Springer-Verlag, New York, NY. pp. 197-220

Zerbe, J.I. 1991. Liquid fuels from wood-ethanol, methanol, diesel. World Resource Review. 3: 406-414.