To the Graduate Council:

I am submitting herewith a thesis written by Bradly Scott Wilson entitled "Modeling Cellulosic Ethanol Plant Location Using GIS." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.

_____

Shih-Lung Shaw, Major Professor

We have read this thesis

and recommend its acceptance:

_____

Bruce Ralston

_____

Burton English

Accepted for the Council:

_____

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

MODELING CELLULOSIC ETHANOL PLANT LOCATION USING GIS

A Thesis
Presented for the
Master of Science
Degree
University of Tennessee, Knoxville

Bradly S. Wilson
August 2009

ABSTRACT

Domestically produced renewable transportation fuels have become a priority for researchers, policy makers, and investors in recent times. Ethanol, particularly cellulosic ethanol made from woody or fibrous plant material, has emerged as one such fuel that could help to ease the current dependence on foreign oil. The questions of where to build the facilities that convert feedstocks to cellulosic ethanol as well where the feedstock will come from to supply these biorefineries are important ones. This paper examines methods for using GIS to model feedstock availability and ideal biorefinery location in an economically feasible manner.

The software developed as part of this study is referred to as BIOFLAME (Biofuels Facility Location Analysis Modeling Endeavor). Expanding on similar efforts that came before it, The model allows the user to perform an analysis on any combination of counties within a 16 state region in the southeast U.S. given parameters such as biorefinery capacity, crop prices, transport cost rate, feedstock yield adjustments, hay land availability, driving distance limit, required profit, and more. A suitability analysis can be performed using an array geographic features that a biorefinery might be situated near or away. A feedstock supply analysis then evaluates the costs involved in siting a facility in all candidate sites within the suitable areas. An ideal site is identified that minimizes transportation and farmgate costs. A report is generated that details the annual costs involved as well as how much and what kind of traditional cropland would be converted to switchgrass production under the scenario. The siting algorithm supports single or multiple facilities.

TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION ---- **1**

CHAPTER 2 - LITERATURE REVIEW ---- **5**

CHAPTER 3 – METHODOLOGY ---- **15**

3.1 – Design Decisions ---- *15*

3.2 – Broad Methodology Explanation ---- *17*

3.3 – Data Acquisition ---- *19*

3.4 – Data Processing ---- *22*

CHAPTER 4 – SOFTWARE DESIGN ---- **28**

4.1 – User Interface ---- *28*

4.2 – Performance ---- *33*

CHAPTER 5 – ANALYSIS AND RESULTS ---- **35**

5.1 – Scenario ---- *35*

5.1 – Analysis ---- *36*

5.3 – Sensitivity ---- *41*

CHAPTER 6 – CONCLUSIONS ---- **48**

6.1 – Summary ---- *48*

6.2 – Limitations ---- *49*

6.3 – Future Development ---- *51*

LITERATURE CITED ---- **53**

VITA ---- **56**

LIST OF FIGURES

Figure — Page

Figure 3.1 – States available for use by the model 20
Figure 3.2 – Crop zones derived from the union of soil map units and county boundaries 23
Figure 3.3 – The model selects lowest cost feedstock supply from crop zones until facility capacity is met 26
Figure 4.1 – Screenshot of BIOFLAME user interface with tasks organized within horizontal tabs labeled Study Area, Suitability Analysis, Feedstock Analysis, and Facility Siting 29
Figure 4.2 – Screenshot of BIOFLAME with expandable suitability factors within the Suitability Analysis tab 30
Figure 4.3 – An example of Feedstock Analysis output showing lowcost feedstock supplying a biorefinery 31
Figure 4.4 – A biorefinery siting report showing annual feedstock costs and crop area converted to switchgrass production 31
Figure 4.5 – The Model Parameters dialog box for Facility Siting 32
Figure 5.1 – Study area selected for the scenario 37
Figure 5.2 – Output map resulting from the first suitability analysis 39
Figure 5.3 – The model's ideal biorefinery location for the first scenario along with its associated feedstock supply 39
Figure 5.4 – Biorefinery siting report generated for the first scenario 40
Figure 5.5 – Suitability map show a shift to rural areas 42
Figure 5.6 – Biorefinery location and feedstock supply map resulting from the facility siting analysis in which the modified suitability map was used 44
Figure 5.7 – Biorefinery report showing increased transportation costs 44
Figure 5.8 – Biorefinery location and feedstock supply map showing a shift away from cotton acreage and toward soybean acreage for switchgrass 45
Figure 5.9 – Biorefinery siting report showing a shift to soybean acreage for switchgrass production 45
Figure 5.10 – Biorefinery location and feedstock supply map resulting from the change to allow for more hay acreage to be considered for switchgrass production 47
Figure 5.11 – Biorefinery siting report showing a shift to hay acreage for switchgrass production and a decrease in feedstock costs 47

CHAPTER 1 - INTRODUCTION

Domestically produced, renewable transportation fuels have become a priority for policy makers and researchers in the United States in recent years. Investing in such fuels would help to diminish the country's vulnerability to high oil prices and oil supply reductions while reducing greenhouse gas emissions from the transportation sector. Ethanol, particularly cellulosic ethanol that is derived from fibrous or woody biomass, has emerged as a viable option that could help to achieve these goals (Kaylen et al 1999; Tiller 2007; NCEP 2004). While much research has gone into evaluating the various feedstocks that can be converted to this type of fuel, one of the major challenges facing the emerging ethanol industry is how to supply the conversion facilities with feedstock in an economically feasible manner. Facility location is a critical factor in balancing the input supply, product demand and infrastructure needs of the industry and GIS can potentially provide valuable insight as we address these issues.

Biofuels have a number of benefits over traditional fossil fuels. The 25x25' Action Plan: *Charting America's Energy Future* estimates that 25 percent of our gasoline and diesel use could be displaced by biofuels by the year 2025, reducing our dependence on imported oil while not constraining domestic food production. This would result in a smaller trade deficit and new markets for agricultural products; increasing farm income and creating new jobs. In addition to domestic economic growth, cleaner burning biofuels could mean fewer smog causing emissions and a healthier atmosphere. In the case of cellulosic ethanol, CO^2 emissions could be reduced to nearly zero, helping to reduce the effects of global warming (English et al 2006).

A variety of abundant feedstocks can be converted to cellulosic ethanol. These range from agricultural waste materials such as corn stover and wheat straw to forest residues like brush and wood chips (Kaylen et al 1999; Tiller 2007). Waste materials from the pulp and paper industry, particularly in the northeast, are being considered for conversion (Menard 2007; NRBP 1994). Dedicated energy crops such as switchgrass are expected to be a major feedstock contributor in the emerging ethanol industry in the Southeastern U.S. DuPont Danisco Cellulosic Ethanol LLC (DDCE), partnering with the University of Tennessee, is currently in the process of building the country's first cellulosic ethanol plant that utilizes switchgrass as a feedstock (DDCE 2009). Plans are underway to construct similar facilities around the U.S. in the coming years that take advantage of cellulosic feedstocks that particular regions have to offer (Renewable Energy Access 2007; Menard 2007).

During the past decade, over 100 ethanol plants have been constructed, primarily in the Midwest, and almost 100 more have been announced (Lambert et al 2007). Up to this point almost all of these have been grain-based, utilizing corn as a feedstock. The National Commission on Energy Policy found that even if 100 percent of the current U.S. corn crop were used to make ethanol, only 25 percent of the nation's gasoline use would be displaced (NCEP 2004). There is a limit to how much corn ethanol we can sustainably produce without disrupting the agricultural sector (Tiller 2007). The Renewable Fuel Standard requires that 7.5 billion gallons of ethanol be blended with gasoline by 2012 (EPA 2007). While that amount can be met with the existing corn ethanol infrastructure, the consensus is that higher amounts of ethanol from corn will not be feasible. By the year 2025, it is

expected that ethanol from biomass will make up the bulk of renewable fuels, bringing the total to as high as 86 billion gallons (English et al 2006). The life cycle green house gas emissions are also lower with cellulosic ethanol primarily due to the fact that perennial feedstocks require considerably less fertilizer and energy inputs to grow than corn (Tiller 2007; NCEP 2004). Biomass conversion facilities can also co-produce electricity from byproducts, reducing the amount of power from fossil fuels needed in the conversion process (NCEP 2004).

Switchgrass has been found to be a viable dedicated energy crop for cellulosic ethanol in the Southeastern U.S [Tiller 2007]. It's a perennial native plant to the region and has the ability to thrive in a variety of marginal soil conditions with very little fertilizer or irrigation. It is believed that switchgrass could simultaneously benefit the regional farm economy while also meeting the feedstock needs of the growing ethanol industry. Unlike corn, it is a bulky material that is not easily packed and shipped via rail or barge. It is typically cut, rolled into square or round bales and shipped by truck. Since the total cost structure of these supply-oriented bio-refinery firms is dominated by feedstock procurement, the location of facilities is crucial to minimizing transportation costs (Lambert et al 2007). Switchgrass yields have been estimated for much of the country, and it is generally well known where the crop can be grown. We can also reasonably predict at what price point farmers will switch from growing traditional crops to growing switchgrass (English et al 2006). Therefore it should be possible to model ethanol facility location as a function of feedstock supply.

[Lambert et al 2007] found several significant determinates for ethanol facility location in addition to feedstock supply and the transportation issues associated with it. Product markets and by-product markets also drive siting decisions. Facilities must be able to effectively deliver the final products to fuel blending and retail gasoline businesses (Reynolds 2002). Not unlike other manufacturing operations such as food processing, these facilities must also be situated near electrical, water and natural gas distribution centers in order to minimize variable cash operating expenses. Labor availability, proximity to metropolitan areas, regional fiscal policies and zoning are other factors that must be carefully evaluated when siting ethanol plants. There are unique considerations involved in cellulosic ethanol plant location, but there are enough similarities with corn ethanol to warrant an investigation of siting determinates that were used in the establishment of the existing corn ethanol industry (Lambert et al 2007). To be useful and relevant today, a GIS based cellulosic ethanol facility location model would need to at least be comprehensive enough to incorporate most of these known siting determinates. The main goal of this study is to evaluate existing methods for solving such facility location problems and apply the knowledge gained in such a way as to provide a comprehensive GIS modeling tool for siting cellulosic ethanol plants in the Southeast U.S. with switchgrass as a primary feedstock.

CHAPTER 2 - LITERATURE REVIEW

When thinking about the problem of siting biorefineries, it is useful to examine some basic principals of industrial location theory. Weber's Least Cost Theory suggests that an industry will locate where it can minimize its costs and therefore maximize its profits (Weber 1929). These costs can be categorized as transportation, labor and agglomeration with transportation being the most important. Using this theory, a logical model can be constructed that illustrates how industrial location is influenced by spatial variations in cost. In Weber's model, raw materials, labor concentration and product market are represented as a triangle of three points on a map. Given transport costs of both moving raw materials to the plant and moving finished products to the market, and given costs involved with moving farther away from the concentrated labor, a set of concentric circles can be drawn around the points. From this, the total cost per unit of production at any point can be calculated and a series of isopleths joining points of equal costs can be derived. Each corner of the triangle exerts a force pulling the least cost location toward it (Smith 1966). Visualizing cost variation within this "locational triangle" is a useful way to frame the problem of facility location as available methods are evaluated.

The facility location problem is common in the realm of GIS, and scientists have been developing and improving location science methods for decades. McHarg, in the late 1960s, proposed the approach of overlaying a series of colored thematic maps, each representing land-based geographic features. These clear acetate sheets were colored from light to dark representing less suitable to more suitable areas in regards to a particular theme. When the sheets were superimposed on top of each other, the composite map

revealed darker areas that were more suitable for a particular route or location for a specific function. This basic mapping idea is utilized in GIS today and at the very least it serves as a way to screen out infeasible or undesirable sites. Modern suitability analysis, using a raster data model, allows for the reclassification of the features on each layer to a common suitability scale, and weights can be applied to give preference to particular layers. The composite output map can be adjusted to show, for example, only the top 15 percent or top 1 percent of the areas that are found to be most suitable (Church 2002).

One of the commonly applied facility location problems is the so-called p-median problem. It requires the siting of *p* facilities among a set of *m* candidate sites to serve *n* demand sites, where *p* is a known number. When this problem is solved on a network, both the candidate and demand sites are restricted to network nodes (Miller and Shaw 2001). When the number of candidate and demand sites is large, this becomes a very complex problem to solve and requires heuristic techniques to compute a near optimal solution in a reasonable amount of time. There are several existing algorithms that do this well, including functions within ArcINFO/Workstation and TRANSCAD software packages (Church 2002). In the case of siting biorefineries, we would allow feedstock supply locations to act as the demand sites. One concern with this approach is how to represent feedstock supply in discrete space since the p-median problem requires demand sites to be situated on the network as points. Feedstock supply might be better represented in continuous space, therefore some sort of spatial aggregation would need to be performed to create the condensed demand points for feedstock. Specifically if feedstock supply is more naturally represented as a raster dataset or a set of polygons, one challenge would be

how to process this data into a set of points instead so that it could be used with one of the network-based location problems.

Another frequent facility location problem is known as the set covering location problem (SCLP). This differs from the p-median problem in that it finds both the number and location of facilities such that demand points are within a specified distance or travel cost of a facility (Miller and Shaw 2001). In the context of siting biorefineries, it would find the amount of facilities needed to exhaust the available feedstock supply in an area. SCLP assumes an unlimited budget for opening new facilities, which might not be realistic for this application. A variation of the set covering location problem, known as maximal covering location problem (MCLP), addresses this concern by maximizing the number of demand points covered by a pre-specified number of facilities (Miller and Shaw 2001). So if it were known, for example, that 5 facilities are needed, MCLP would site those facilities in proximity to the highest supplies of feedstock. Although the objective functions of SCLP, MCLP and the p-median problem aim to solve the core biorefinery siting problem, some hybrid method might be required to address the case of continuous feedstock supply coverage, as well as the case of balancing feedstock availability with ethanol demand constraints.

Most location modeling development has concentrated on discrete network-based models as opposed to continuous planar models. The primary reason for this bias is that a network can capture a number of fine distinctions of spatial variability that are often assumed away in an unbounded, continuous, planar representation. For example a network might accurately represent travel distances in an urban area whereas a planar

model may not when numerous barriers exist. There is continuing and growing interest in the planar approach though, because representing demands as continuous areas instead of points makes sense in many applications. A few hybrid network and planar paradigms do exist in literature, but significant spatial detail is required to describe a real world problem (Church 2002).

This literature review revealed three previous real world implementations of GIS-based models for siting biorefineries. The three models approached the problem in a similar manner that generally conforms to Weber's Least Cost Theory. However, there were significant differences in the implementations in terms of the data preparation, siting algorithms and results generated. Careful analysis of the three approaches should provide useful insight prior to the development of a solution that achieves the goals outlined for this particular study.

The first paper discusses methodologies used in the Regional Integrated Biomass Assessment (RIBA) project, which analyzed the economic feasibility of locating switchgrass-to-ethanol conversion plants within an 11 state region in the United States (Noon et al, 2002). The system was designed to predict both where the energy crop could be grown and the marginal cost of supplying the feedstock to specific locations. It is a loosely coupled GIS modeling system rooted in ArcINFO/Workstation with custom software modules developed in C that perform four basic functions; mapping cropland availability, calculating expected yields and farmgate price, mapping the marginal cost of delivered energy crop feedstocks, and mapping sites where biorefineries might be co-

located. It was designed to evaluate individual states but could be modified to evaluate larger or smaller geographic regions.

The spatial resolution of the cropland map used by the system is 1 km^2, and it was created by overlaying county boundary, soil-group, and land-use maps. The first system component calculates variables that define what proportion of each pixel is cropland suitable for growing switchgrass. This is done by linking county level data on the relative dominance of conventional crops in each county to the map. A farmgate price, which represents an estimate of the price per unit required by a farmer to sell a crop, is calculated in component two along with expected crop yields. These values are determined using switchgrass field trial results, observed conventional crop production data, and soil yield indexes from the crop simulation model EPIC. With the availability of energy crop, potential yield, and farmgate price known, the model is able to map the marginal cost of delivered feedstock in component three. It does this by converting the cropland map pixels to points, calculating the shortest path distance between each point and every other candidate node on the map, and determining a cost per ton mile of feedstock. The marginal cost algorithm ranks the delivered cost (farmgate price + transport price) for all supply points in a state to the target destination and calculates the cumulative potential supply of feedstock. The marginal cost is the cost of the most expensive ton of delivered feedstock to the destination. When this algorithm is run for all destinations in the state, a marginal cost surface map is created. Component four is able to site potential biorefineries using this cost surface. When a plant is sited at the least cost location, the feedstock supply points used by that plant are no longer available to other plants. As the process continues, the

next best least cost location is identified and plants are sited until all available feedstock is utilized in the region (Noon et al 2002).

Central to the RIBA model is the ability to generate a cost surface for siting facilities. The approach taken utilizes a large set of supply nodes throughout the state and a subset of candidate nodes that reside on the transportation network to essentially solve a set covering location problem. It fits the SCLP paradigm because the optimal number of facilities and their locations are determined based on available supply nodes. In an attempt to minimize processing time, the label-correcting Pallottino shortest path algorithm was implemented with two queues to find the shortest travel distance between candidate nodes and supply nodes in the region. The issue of feedstock supply being represented as a continuous surface is dealt with by generating points from raster pixels. When larger geographic areas are used, the number of supply points would increase dramatically. This appears to be the primary factor in the computational processing time of the system. The results of the model illustrate graphically the number and location of the sited facilities based on minimized marginal costs of delivered feedstock as well as the feedstock supply allocated to each facility (Noon et al 2002).

The second real world implementation of a GIS-based siting model found in literature was developed to identify ideal locations for biofuel-based energy production using logging residues in Finland. With a focus on increasing forest fuel usage, the project analyzed the availability of residues by means of both resource supply and demand-side approaches in order to identify the most suitable regions. The analysis included availability comparisons between power plant sites and resource allocation in a least-cost manner according to

feedstock supply cost, and between existing plant structure under given demand and supply constraints. By comparing the feedstock supply potential and the power plant demand potential, the model identified deficit and surplus areas around the country (Ranta 2004).

Forest stand data used in the project consisted of sparse samples of forest inventory plots. Natural neighborhoods were built around the plots using Delanay triangulation resulting in a network of Thiesen polygon sub-regions. The center of each polygon represented the aggregated forest stand supply point, and it was assumed for estimation and presentation of logging residue potential that residue was uniformly spread throughout the sub-region. To incorporate the effect of regional geography into these supply potentials, the transport distance along the road network within a 100 km radius of each point was calculated. Prior to the distance calculations, the potentials were aggregated again to a 10 km grid of points. This was done to reduce the intensity of the distance calculations. The points were visualized on a map using contour regions classified according to supply potential. A linear programming model was developed to allocate facilities based on the generalized supply regions. The model was capable of being run both with and without demand constraints expressed in terms of giga-watt hours. When usage targets were set, the results would reveal surplus and deficit areas in regards to new biorefineries meeting energy demand (Ranta 2004).

This GIS modeling system was loosely coupled and based on the MapInfo software package and extended with the MapBasic programming language. Data aggregations were in part calculated by the Vertical Mapper MapInfo extension, and the transport LP-model

was developed using the Premium Solver Platform with plug-in Large-scale LP Solver engine for Excel. RoutView Pro, another MapInfo extension, was used for transport distances (Ranta 2004). Methods used in this model differ from those used in RIBA primarily in the way data is aggregated for the purposes of faster execution. The Finland model does allow for the incorporation of both supply and demand constraints, but this comes with a cost of much coarser resolution due to aggregation. While both approaches afford worthwhile insight into the process of siting these types of facilities, the techniques used in the RIBA project are probably better suited to meet the needs of this particular study.

The third biorefinery siting model reviewed was developed by The University of California, Davis as part of the Western Governors' Association's Strategic Assessment of Bioenergy Development in the West project. Of the three models reviewed, this system is the most comprehensive in terms of number of feedstocks utilized, modes of transportation, and area covered. It includes 17 states in the western U.S. from North Dakota south to Texas and westward. Twenty-three feedstocks are utilized including several types of crop residues and forest residues as well as soybean and canola oils, and municipal solid waste. Modes of transportation considered by the model are road, rail, and barge. The system is loosely coupled and incorporates a modern GIS to supply inputs to a GAMS optimization model. ArcGIS is used to determine available feedstock, potential biorefinery locations, and transportation costs while GAMS determines the optimal spatial distributions of biomass supply and fuel demands, competitions between technologies for

resources, and economies of scale of conversion technologies in determining ideal biofuel supply chains (Western Governers' Association 2008).

Due to the complexity of the multiple feedstocks, transportation modes, and size of the analysis area being considered within the U.C. Davis model, a number of trade-offs were rationalized to ensure the system performs the siting analysis in a reasonable amount of time. Since feedstock data existed at the county level, it was decided that transportation costs would be calculated at the county level as well. This greatly simplifies the complexity of the network analysis since only travel distances and times between county centroids are processed resulting in an origin-destination cost matrix that is orders of magnitude smaller than a system using a sub-county points. Another trade-off that was made was to choose representative locations for potential biorefineries based on existing cities and towns. City data was analyzed to determine which areas could support a biorefinery. Assumptions about infrastructure were made based on population data and places not meeting the assumed criteria were culled. Also, if cities or towns were within 50 km of one another, they were merged into a single representative point. Setting up the GIS portion of the model in this fashion allows for a more manageable amount of data to be passed to the GAMS model, which is a mixed integer linear program that determines the optimal biorefinery location and identifies feedstock that could supply it. (Western Governors' Association 2008)

Compared to the RIBA and Finland models, the U.C. Davis model does a much more comprehensive job of looking at the problem of identifying available feedstock to supply a biorefinery on a small-scale basis. If the goal is to do a large-scale analysis at the state level,

the RIBA method might be more desirable since it looks at a much finer grained array of supply points and more detailed road-based transportation network. An ideal GIS solution for solving the biorefinery location problem would incorporate the feedstocks, transportation modes, and geographic coverage of the U.C. Davis model and perform the analysis at the relatively large scale of the RIBA model. The trade-off is functionality versus run time and finding this balance in a new GIS model is part of goal of this study.

CHAPTER 3 – METHODOLOGY

3.1 – Design Decisions

After looking at the facility location problem and specifically the biorefinery location and feedstock supply problem in-depth as part of the literature review, a number of key design questions were identified as being important to answer prior to model construction. Since the focus of this study is on switchgrass as a feedstock and switchgrass is typically transported by truck, the road network was chosen for transportation cost calculation. The question of how to represent switchgrass supply was a bit more difficult. The RIBA method of using discrete points 1km apart would be desirable but scaling them beyond the state level to the south eastern U.S. would likely be difficult to process in a reasonable amount of time. Using county level supply on the other hand, such as in the U.C. Davis model, would not capture enough spatial variability for the purposes of this study. Since the smallest geographic unit of data available for estimating crop yields is soil boundary data, it was determined that swithgrass supply would be represented using areas defined by the intersection of soil boundaries and county boundaries. This would allow for a larger scale analysis of cropland than county level but still involve orders of magnitude less records of data to manage than with discrete points 1 km apart.

Another key question to consider in model design was how to represent potential biorefinery locations or candidate nodes. Rather than constrict these locations to existing

cities and towns or even to locations on the transportation network, it was decided to distribute an even grid of points across the southeastern U.S. 5 miles apart. This would allow for a more fair analysis of potential sites that isn't influenced by population. Because these nodes do not necessarily fall on or close to the transportation network, a network based transportation analysis such as what was used in the RIBA and U.C. Davis models would not be ideal for this study. To fit the network-based paradigm, off road travel distances would need to be calculated for each node to the nearest road and then a very large origin-destination cost matrix constructed to maintain travel costs between candidate nodes and supply regions. To overcome these complications, it was decided a raster based shortest path method would be used to determine the travel cost of transporting feedstock from the supply areas to the potential facility sites. Each candidate node would have an associated table that contains information about the feedstock supply surrounding it out to a limited distance. These individual tables would contain the travel cost data as opposed to one large O-D cost matrix.

Rather than build a loosely coupled model like the models described in the literature review, it was thought that building the entire system inside one executable program could reduce complexity in the long run. The advantage of this approach is that there would be only one code-base to maintain; the disadvantage could be the lack of highly optimized solvers that for example GAMS could provide. Generally, the RIBA methods of generating supply curves and siting facilities in a linear fashion were found to be better suited to the goals of this study. It was thought that an algorithm could be created using those basic ideas and be enhanced to evaluate more detailed scenarios.

3.2 – Broad Methodology Explanation

At the core of the model developed for this study is the evaluation of feedstock availability across a region. By examining where traditional crops are currently growing and what yields the associated soils are capable of, it is possible to estimate how much switchgrass could be grown in these areas. Given the market prices of these traditional crops, along with their respective costs of production, it is also possible to estimate at what price point a farmer might start producing switchgrass as an alternative. The model calculates this price point using a break-even formula and the result is referred to as the farm gate value or the net value of the feedstock when it leaves the farm. This break-even formula derives this price estimation by comparing current crop production to potential switchgrass production for an area. Starting with the assumption that a producer must earn at least as much as he or she would with a traditional crop then:

*Crop price * Crop Yield – Crop Cost = Switchgrass Price * Switchgrass Yield – Switchgrass Cost – Required Profit*

Where the required profit is the additional amount required before a producer will convert from an annual traditional crop to a perennial crop like switchgrass. Then setting the breakeven price equal to the switchgrass price and rearranging, the following is found:

*Breakeven Price = (Crop Price * Crop Yield – Crop Cost + Switchgrass Cost + Required Profit) / Switchgrass Yield*

Current crop price and yield can dramatically affect the results of the formula. If crop price and yield are relatively low, the area becomes much more attractive for converting to switchgrass production.

In order to determine the total cost of feedstock that will supply a biorefinery, the cost of transporting the switchgrass bales from the farm to the facility must be determined. The model does this by multiplying a user-defined cost per ton-mile by the shortest linear distance along the road network from the farm to a potential biorefinery. The sum of farm gate cost and transportation cost is the total feedstock cost.

Conceptually, this analysis is performed for every potential biorefinery location and surrounding cropland. The assumption is that a potential investor will build a facility near the most abundant source of cheap feedstock. The literature revealed that although supply and transportation are the two dominate factors in biorefinery location, other important geographic features should be taken into consideration as well [Lambert et al 2007]. The model incorporates a suitability analysis to identify a set of candidate facility locations that match criteria that is important to a user's particular scenario. It is this set of candidate facilities that are used in the feedstock supply and transportation part of the analysis.

Given a particular biorefinery capacity goal, the model iterates through candidate facilities evaluating feedstock costs. Depending on the objective function specified by the user, the model identifies the facility with the lowest total feedstock cost, lowest average cost per ton, or the lowest marginal cost. In addition to providing an annual feedstock cost summary, the model also calculates the number of acres of each traditional crop that are

converted to switchgrass production under the conditions of the scenario, allowing for the impact on regional agriculture to be evaluated.

3.3 – Data Acquisition

The model was developed to encompass a 16 state area of the southeastern U.S that includes Missouri, Illinois, Indiana, Ohio, West Virginia, Virginia, North Carolina, South Carolina, Tennessee, Kentucky, Arkansas, Mississippi, Alabama, Georgia, Florida, and Louisiana (Figure 3.1). This decision was made primarily for two reasons. First, the literature review indicated that most research up to this point on switchgrass as a viable bioenergy feedstock has generally been confined to the southeastern states. Secondly, a considerable amount of computational processing time is required to pre-process the data for use with the model and this increases linearly with geographic area. The only real limitations to expanding this model from the southeast to the entire country are pre-processing time and disk space. For the purposes of this study, the 16 state area was considered a reasonable trade-off given the resources available. The spatial and tabular data used by the model are summarized in Tables 3.1 and 3.2 respectively. Each dataset covers the entire U.S. but the initial version of the model only incorporates the states mentioned above.

On the spatial side, most of the basic infrastructure features such as roads, railroads, waterways, cities, and political boundaries were acquired from the ESRI Maps



Figure 3.1 – States available for use by the model

Table 3.1 – Spatial data used in model development

Table 3.2 – Agricultural data used in model development

and Data 2007 DVD. Generalized soil map unit boundaries came from the SSURGO dataset and when coupled with the associated tabular data provided the model with its smallest geographic unit of analysis for crop yields. FEMA hazard maps were also utilized, as was the USGS National Land Cover Dataset.

Most of the tabular agricultural data exists at the county level. The exceptions were crop prices, which were found at the national level although hay prices were found at the state level. Traditional crop acreages and yields were extracted from the National Agricultural Statistics Service database. Crop budgets including costs for switchgrass came from the Agricultural Policy Analysis Center's POLYSYS databases. Current crop prices were found at the USDA's ESMIS website. SSURGO tabular data provided potential crop yield data at the sub-county level. The Model contains crop data for 9 traditional crops and switchgrass (Table 3.3).

A vector dataset containing candidate facility nodes was created to cover the 16 state area. The points were spaced equally apart every 5 kilometers and total to 77,255. The decision to space the nodes in this fashion was made after some initial performance testing showed that real time model performance scaled linearly with number of candidate nodes. For the purposes of this study, keeping the total number of nodes under 100,000 was a suitable tradeoff to evaluate the modeling framework. Theoretically the size of the candidate node dataset would only be limited by processing time and disk space. A feedstock node vector dataset was also generated containing points spaced 0.5 kilometers apart. Using the NLCD land cover raster as a filter, all nodes that fell outside of agricultural lands (urban, water,

Table 3.3 – Crops used by the model

Crop Name
Barley
Corn
Cotton
Grain Sorghum
Hay
Oats
Rice
Soybeans
Wheat
Switchgrass

forest, etc.) were removed. The resulting dataset is used by the model when displaying low-cost feedstock that supplies a biorefinery.

The acquired spatial data was projected to Albers Equal Area Conic, clipped to the 16 state area and cataloged on an ArcSDE server. Shapefiles were extracted from the vector datasets while rasters were extracted and saved locally in GRID format. This was done to allow the model to work in both networked and offline environments. Tabular data was stored in tables both remotely on a SQL Server as well as locally in an Access database for the same reason.

3.4 – Data Processing

In the interest of allowing the model to perform its analysis at the smallest geographic unit possible given the data available, a new dataset of boundaries was generated and will be referred to in this study at *crop zones*. This dataset was derived by performing a union of county boundaries and soil map units (figure 3.2). The number of resulting zones per county varies depending on the underlying soil unit layout. Because the area of each crop zone is known relative to the area of each associated county that contains it, area weighting can be used to distribute county level crop data to the crop zone level. Soil level data is also adjusted to these derived zones using a similar method.

The SSURGO soil database contains numerous properties of the soils contained in each soil map unit, one of which is the potential yield of various crops. While this is a useful way to view cropland at the sub-county level, the estimated yield values in this



Figure 3.2 – Crop zones derived from the union of soil map units and county boundaries

dataset occasionally stray well out of the range of expected values for a particular location. To take advantage of the spatial detail of SSURGO and attempt to adjust its yield outliers back into a more reasonable range, an adjustment factor was calculated using a three-year average of county level NASS crop yields. This adjustment factor brought the outliers back in line while keeping the spatial variability of the SSURGO dataset.

The suitability analysis portion of the model allows for the selection and weighting of one or more geographic features that are relative to a particular scenario. The model combines these features and generates a suitability map that indicates areas that are the most suitable. Each suitability layer actually consists of two raster layers that are the results of a straight-line distance function and reclassification to a common suitability scale. One raster is classified to show areas near the geographic feature as favorable and the other is classified the opposite way, showing areas away from the feature as favorable. This allows the user to choose to situate a biorefinery near or away from a number of geographic phenomena that may be important in the siting decision. A weight between 1 and 100 can be assigned to each layer to signify its relative importance. The output suitability map is created by multiplying the pixel values of each suitability layer by the user specified weight, then summing the resulting products and dividing by the total number of suitability layers. The resulting suitability map can be adjusted to show more or less suitable areas and can be used as a filter during the facility location process to only consider potential facility locations that fall within the suitable areas.

A feedstock analysis is performed at the crop zone level for all zones surrounding a potential biorefinery site within a 50 mile straight-line buffer. The break-even formula

incorporates the crop zone level agricultural data as well as user-defined crop prices and optional switchgrass yield adjustments to calculate a farm gate value. The model uses a base set of county level switchgrass yields, but these can be adjusted by the user to simulate, for example, improvements in plant genetics or drought conditions. Another optional parameter used in the break-even formula is required profit, which allows for the simulation of additional profit the farmer may demand. It should be noted that the spatial reference is consistent throughout the spatial datasets used with a linear unit set to meters with the exception of the 50 mile buffer layer. It is assumed that a biorefinery will not procure switchgrass that is more than 50 miles straight-line distance from the facility.

Transportation costs by crop zone are calculated using a raster based cost weighted distance function. The road network is assigned weights based on road class with off road areas assigned a relatively large value. These weights are used by the cost weighted distance function along with a potential biorefinery location to create an output raster containing pixel values that are shortest path distance along the road network to the site. The average linear distance from each crop zone to the potential site is found by summarizing the pixels within each zone. The cost of transporting feedstock is found by multiplying this distance by a user-defined cost per ton-mile. Total feedstock cost is the sum of the farm gate value and the transportation cost.

With total feedstock cost known by crop zone, the model sorts by increasing cost and selects feedstock from crop zones until a given biorefinery capacity is met (Figure 3.3). Using the selected feedstock attributes, total annual cost, average cost per ton, and marginal cost can be calculated along with the amount of traditional crop acreage that is

being used to grow switchgrass under the scenario. The crop zones containing the low-cost feedstock can be used to select points from the feedstock node layer to display on an output map. It should be noted that an optional parameter can be used in this selection process to constrain trucks to a limited driving distance. Even though the maximum area around the plant for the consideration of feedstock is 50 miles, the actual driving distance along the road network can greatly exceed this distance.

The facility siting part of the model performs a feedstock analysis for every candidate facility node. The set of candidate nodes used here are confined to the selected study area and optionally the suitable areas defined in the suitability analysis. The model calculates the total annual cost, average cost per ton, and marginal cost of the least expensive feedstock surrounding each candidate node. The ideal facility location is



Figure 3.3 – The model selects lowest cost feedstock supply from crop zones until facility capacity is met.

determined by minimizing one of the above values, depending on the objective function selected by the user. The low-cost feedstock map is displayed along with a report showing a breakdown of annual feedstock costs as well as crop acreage converted to switchgrass production.

In the case of siting multiple biorefineries, the model allows the user to queue up a number of facilities each with its on capacity and optional driving distance restrictions. The model sites each facility in order, finding the best location followed by the next best and so on. When a biorefinery is sited, its associated low-cost feedstock is flagged as unavailable for the next facility in the queue. By doing this, the model is simulating a contract between the biorefinery and the farmers for the switchgrass for a period of time, forcing the next biorefinery to look elsewhere for low-cost feedstock. The locations of the facilities are shown on a map and a report is generated listing the annual costs and converted crop acreage associated with each.

CHAPTER 4 – SOFTWARE DESIGN

4.1 – User Interface

The software developed in this study is named BIOFLAME (Biofuels Facility Location Analysis Modeling Endeavor). Containing over 11,000 lines of code, it was written in Visual Basic.NET using Visual Studio.NET 2005 and ArcGIS Engine 9.3. Tabular data is stored locally in an Access database for offline use and remotely on a SQL Server for use in a networked environment. Spatial data is stored locally in Shapefile and GRID format and remotely on an ArcSDE server. Local XML files, one per candidate facility node, contain cached data that was pre-processed to improve run-time performance. BIOFLAME was designed to be as database driven and modular as possible allowing agricultural data, spatial data, and suitability layers to be updated by the user without the need to recompile the code.

The user interface is divided into four main sections or tabs, each representing a task; study area selection, suitability analysis, feedstock analysis, and facility siting (Figure 4.1). Each tab contains a map window on the right, toolbar above the map, and legend to the left. The toolbar buttons change depending on which tab is currently active. All toolbars contain buttons for basic map operations such as zooming, panning, full extent, and identify. A menu bar at the top of the screen contains menu items for each tab's toolbar commands. The file menu also contains a menu item that opens a "Data Preparation" tab for pre-processing data.

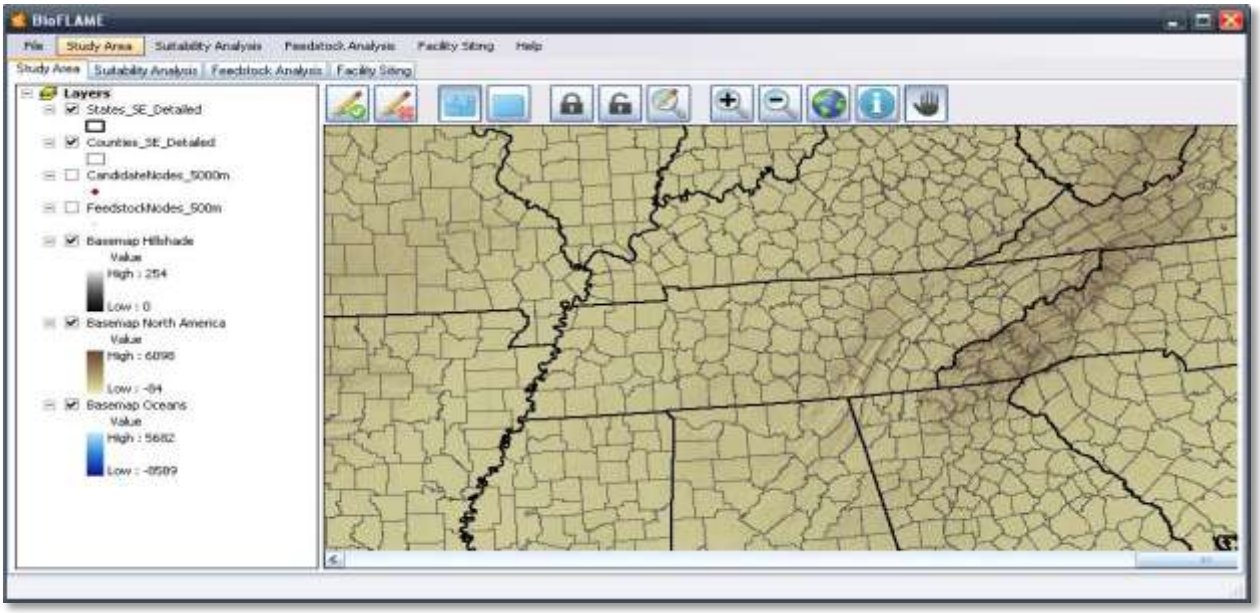


Figure 4.1 – Screenshot of BIOFLAME user interface with tasks organized within horizontal tabs labeled Study Area, Suitability Analysis, Feedstock Analysis, and Facility Siting.

The user begins by selecting a study area that consists of any combination of states or counties within the 16 state area. Once the selection is made, clicking the "Lock Study Area" button saves it and the area becomes shaded in blue on the map. The next tab, "Suitability Analysis" displays a set of expandable suitability factors in the legend area (Figure 4.2). The user can check any factors that are relevant to the scenario and expand them to view additional options. Each suitability factor can be set as "near" or "away from" depending on whether the particular geographic feature is desirable or undesirable. A slider bar can also be used to adjust a relative importance weight from 1 to 100. When the "Run Suitability Analysis" button is clicked, a portion of the map is shaded red to indicate suitable areas based on the factors and weights chosen. A slider bar can be adjusted to show more or less suitable area, with a default value of 1%. The third and forth tabs have similar layouts but allow for manual and automatic facility siting respectively.

On the “Feedstock Analysis” tab, a facility location can be set manually by clicking within the study area. A “Model Parameters” button opens a dialog box with scenario specific parameters the user can adjust such as facility capacity, crop prices, and driving distance limit. After the “Run Feedstock Analysis” button is clicked, low-cost feedstock for the biorefinery is displayed on the map (Figure 4.3). The “View Report” button displays a report showing annual feedstock costs and converted crop acreages (Figure 4.4). This tab and the tools contained within it are optional and are provided as a way for the user to fine tune their analysis. It is expected that the user will typically go from the Suitability Analysis to the Facility Siting tab, only using the manual Feedstock Analysis to quickly evaluate the consequences of moving a biorefinery to different locations around the area, independent of the model’s ideal location.

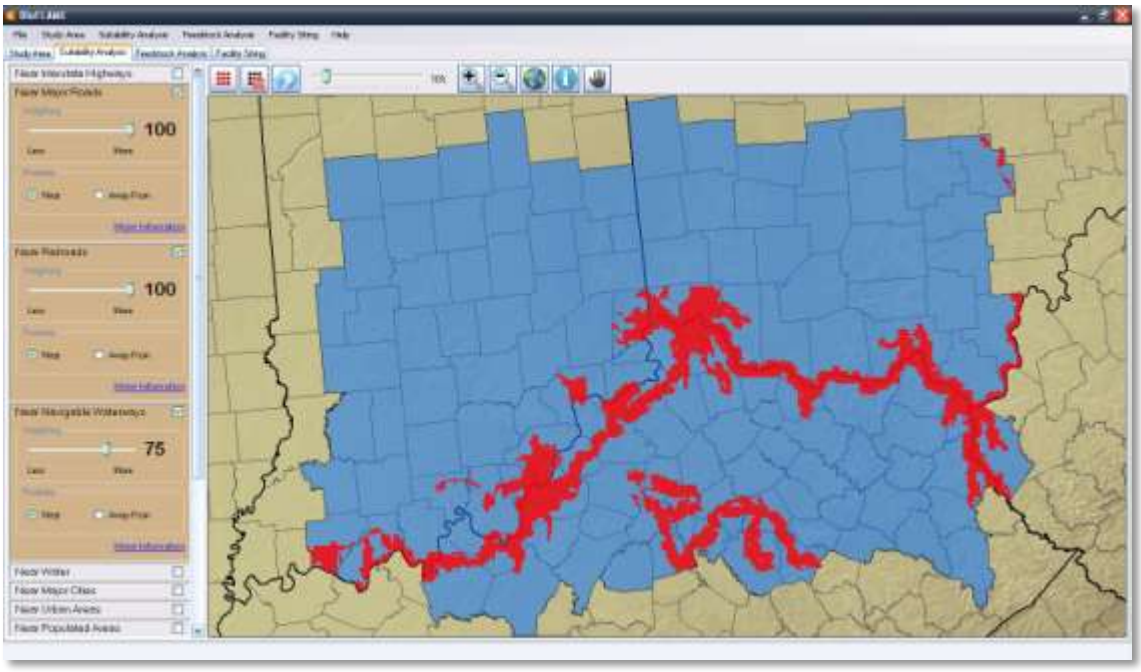


Figure 4.2 – Screenshot of BIOFLAME with expandable suitability factors within the Suitability Analysis tab.

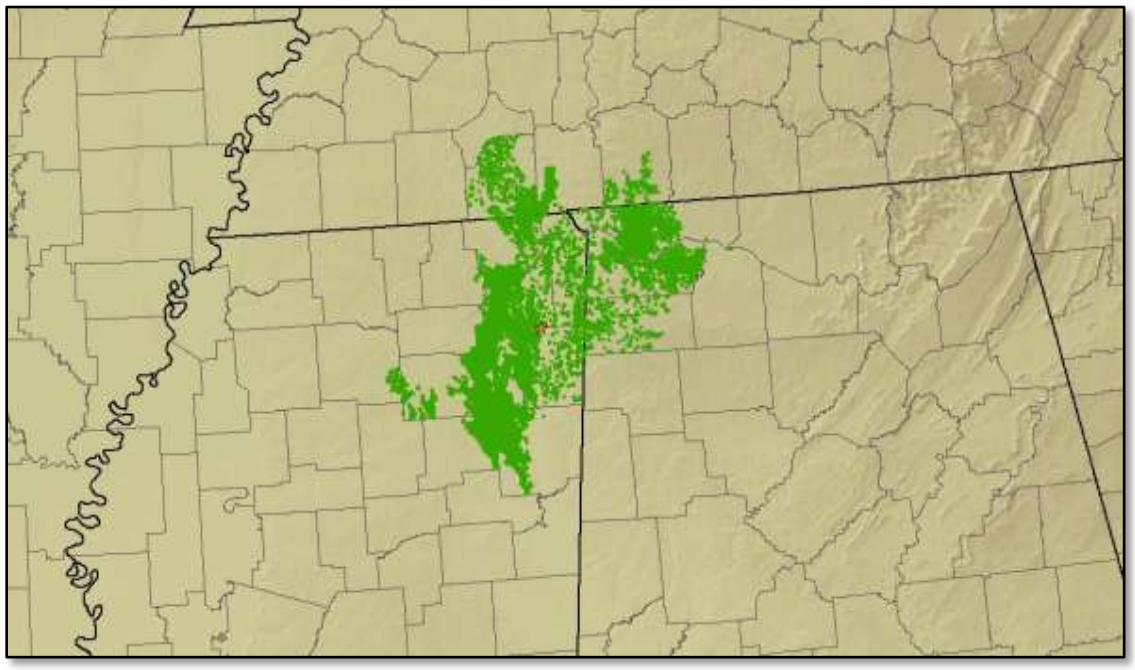


Figure 4.3 – An example of Feedstock Analysis output showing lowcost feedstock supplying a biorefinery.

BioRefinery Siting Report

Feedstock Summary

Total Feedstock Cost	$31,438,850
Total Farmgate Cost	$21,260,589
Total Transportation Cost	$10,178,272
Total Switchgrass Supply	701,001
Marginal Cost	$39.06
Average Feedstock Cost Per Ton	$44.85
Average Farmgate Cost Per Ton	$30.33
Average Transportation Cost Per Ton	$14.52

Cropland Converted To Switchgrass

Barley Acres	
Corn Acres	6,000
Cotton Acres	206
Hay Acres	6,265
Oats Acres	
Rice Acres	
Sorghum Acres	338
Soybeans Acres	67,310
Wheat Acres	1,213

Figure 4.4 – A biorefinery siting report showing annual feedstock costs and crop area converted to switchgrass production.

The fourth tab "Facility Siting" allows the user to set similar parameters as with the Feedstock Analysis but also has a Facility Selection option for automatically determining the ideal location of one or more facilities. The Model Parameters dialog box (Figure 4.5) provides a way to site a single biorefinery or add multiple facilities to a queue, each with its own capacity. The "Run Facility Siting" button starts the siting process after which the bio-refinery location and associated low-cost feedstock is shown on the map. The analysis is performed for all areas within the study area shaded as suitable from the Suitability Analysis. The View Report button again opens a dialog box showing an annual cost summary by facility as well as a break down of the traditional crop acres converted to switchgrass under the scenario.

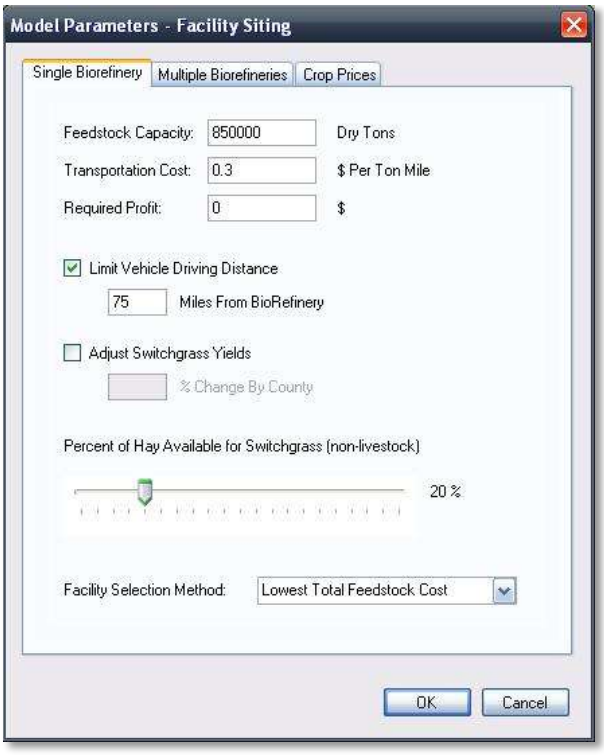


Figure 4.5 – The Model Parameters dialog box for Facility Siting.

4.2 – Performance

During the software development process, much attention was paid to runtime performance. When examining similar existing models it was found that slow performance hindered the user's ability to evaluate a variety of scenarios in any sort of reasonable time. One goal with BIOFLAME all along was to allow the user to put together a scenario, instantly see results, adjust some parameters and quickly see how those adjustments affect the output. Due to the scale of the spatial data and the time consuming nature of the geoprocessing techniques used by the model, providing a way for the software to site facilities in near real time was one of the central challenges of developing this software.

The approach taken to minimize runtime involves pre-processing as much "static" data as possible. Examples of static data include linear distances along the road network, crop planted area, and average crop yields. The idea was to store these values for all crop zones that fall within a 50-mile buffer of each candidate node. At runtime the model reads this data from disk for the subset of candidate nodes it needs and avoids spending processing time performing lengthy GIS tasks. This speed gain comes at the cost of disk space used for the cached XML files and the time spent pre-processing the data.

The sub routine that builds the cached data files contains a large loop that handles the bulk of the GIS tasks needed for each candidate node. Some tasks performed inside this loop include extracting crop zones and roads within a 50 mile buffer of the candidate node, running a cost-weighted distance function, calculating the average linear distance from

each crop zone to the node, area weighting the agricultural data, and saving to an XML file. In the case of the initial version of the software this loop executed those tasks 77,225 times, once for each candidate node. The Data Preparation dialog box in BIOFLAME allows for this process to be distributed across multiple computers with each computer responsible for a batch of nodes. Total processing time across four computers was close to 2 weeks and the resulting cached files occupy over 100 gigabytes of hard disk space. This routine would only have to be run again when changes are made to the static data used by the model.

With the cached data in place, BIOFLAME simply reads an XML file from disk for a requested candidate node and calculates output values using cached values, user parameters, and some other "dynamic" values stored in Access tables such as state level hay prices. Runtime performance obviously varies based on the size of the study area the user is evaluating, but this caching technique brings a process that would otherwise take hours or days down to seconds or minutes. The size of the displayed suitable area in the Suitability Analysis affects runtime performance as well because this essentially filters out a number of candidate facilities. In one test scenario where the entire state of Tennessee was selected as a study area with 1% suitable area displayed, the model sited a biorefinery in 1 minute 52 seconds. In the same scenario the suitable area was then set to 20% and the model took 6 minutes and 15 seconds to solve. Based on this performance testing of a relatively large area, it was considered that the goal of using pre-processing to achieve near real-time model results was achieved.

CHAPTER 5 – ANALYSIS AND RESULTS

5.1 – Scenario

In order to gauge the effectiveness of the model at evaluating feedstock availability and siting biorefineries a scenario was constructed that resembles what an investor might face in the real world. By analyzing these results, strengths and weaknesses can be indentified which can help to highlight parts that might need further revision as well as the directions future development might take. By observing the consequences of adjusting various user parameters for the scenario the sensitivity of the model can also be measured.

In this fictional scenario, an investor is interested in building a cellulosic ethanol plant that will process approximately 700,000 dry tons of switchgrass annually. The plant will be constructed somewhere within a rectangular selection of counties ranging from central Arkansas eastward to middle Tennessee extending southward to cover the northern portions of Mississippi and Alabama. This study area encompasses the metropolitan areas of Little Rock, Memphis, Birmingham, and Huntsville as well as cropland where cotton, soybeans, corn and hay are dominant. Three interstate highways, three major rivers, and a sprawling railroad network extend through this landscape.

In addition to choosing a site that is near an abundant supply of cheap switchgrass the investor also would like to situate the plant near infrastructure that is vital to its operation. In this case existing roads and power lines are top priority as well as bodies of water such as lakes and streams. In an attempt to minimize the transport costs of moving

the ethanol product from the plant to urban demand centers, the site would ideally not be far from major cities. To a lesser extent proximity to railroads is desired for potential distribution of the product to areas farther away.

For the purposes of this analysis, current national crop prices are assumed for all crops except for hay for which current state level prices are to be used. The investor assumes that the majority of the hay grown in this area will be used for livestock consumption, only 20% of hay acreage will be considered for possible conversion to switchgrass. Arrangements with trucking companies restrict trucks from traveling more than 70 miles from the biorefinery to obtain bales of switchgrass. The assumed cost of transporting the feedstock is 50 cents per ton-mile.

5.1 – Analysis

Using the BIOFLAME software, the counties within the study area in question were selected (Figure 5.1). Major roads, water, and powerlines were toggled on for the suitability analysis and their importance weights were set to the maximum value of 100. Urban areas and railroads, being slightly less important, were selected and their weights set to 75 and 50 respectively. All of the chosen suitability factors had the "Near" option set to instruct the model to show suitable areas near these geographic features. Within the siting parameters dialog box, the plant capacity was set to 700,000 dry tons and default crop prices were accepted since they were current. Driving distance limit was checked, a value of 70 miles was entered, and the cost per ton-mile was set to $0.50. The hay acreage



Figure 5.1 – Study area selected for the scenario.

slider bar was also adjusted to 20%. The facility selection method used for this scenario was Lowest Total Feedstock Cost.

After running the suitability analysis and setting the output slider bar to show 5% suitable area, the map in Figure 5.2 was generated. Areas shaded as suitable that immediately stand out include the areas around Little Rock Arkansas where interstate 40 dips southward, a nearly vertical line stretching from just north to just south of Memphis Tennessee, and the area connecting Decatur and Huntsville Alabama. Overlaying roads, railroads, and water features on this map, it becomes obvious why the model chose these areas. Since urban areas were weighted 75, it found a combination of these features as close to major cities as possible. Powerlines generally follow roads so their influence

generally overlaps. Other shaded areas speckle the landscape showing pockets of suitable land that meets the criteria of the analysis.

With the suitability map in place for use as a filter of potential candidate facility sites, the siting analysis was run. Figure 5.3 shows the model's ideal biorefinery location for this scenario along with the estimated location of the cropland that would convert to switchgrass production and supply the plant with the cheapest feedstock available. The site is just outside of Huntsville Alabama, 8 miles north of Guntersville Lake, situated near the intersection of several railroad lines. The majority of the available switchgrass is shown to be in the agricultural areas of Madison County with Limestone, Marshall, and Morgan counties supplementing. A relatively small amount of production is shown to be in the higher elevations of Jackson County to the east.

A biorefinery siting report was also generated (Figure 5.4) that shows estimated costs of the feedstock. The model calculated a total annual feedstock cost for this biorefinery of nearly $21.6 million, $14.4 million of which is estimated to be the cost of purchasing the feedstock from the farmer. That leaves $7.1 million, or roughly one third,

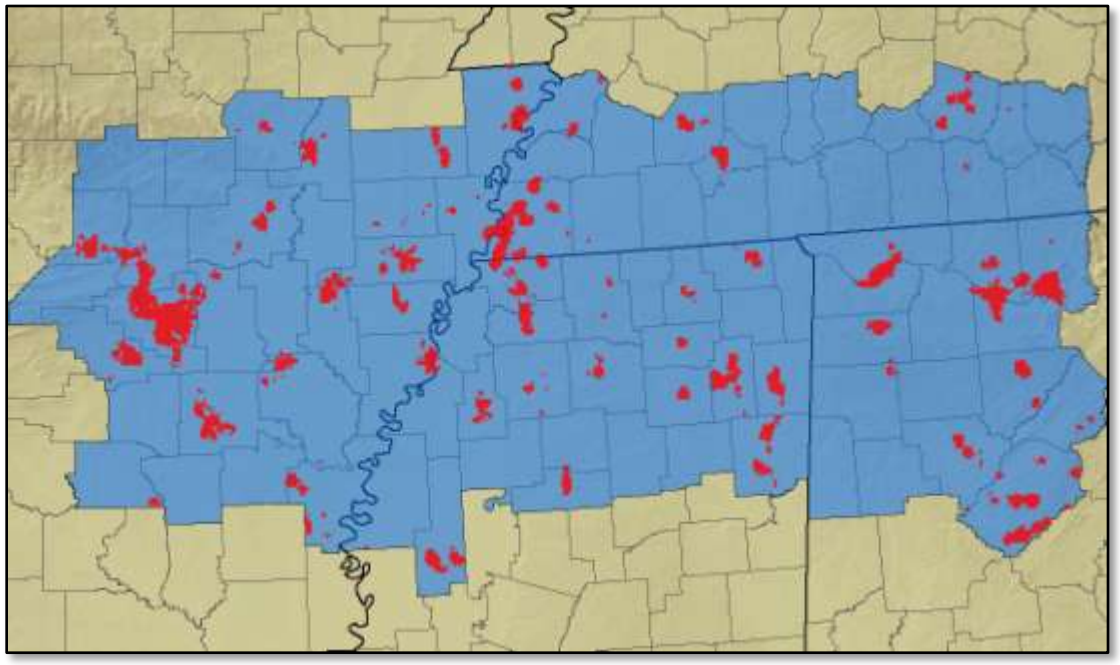


Figure 5.2 – Output map resulting from the first suitability analysis.

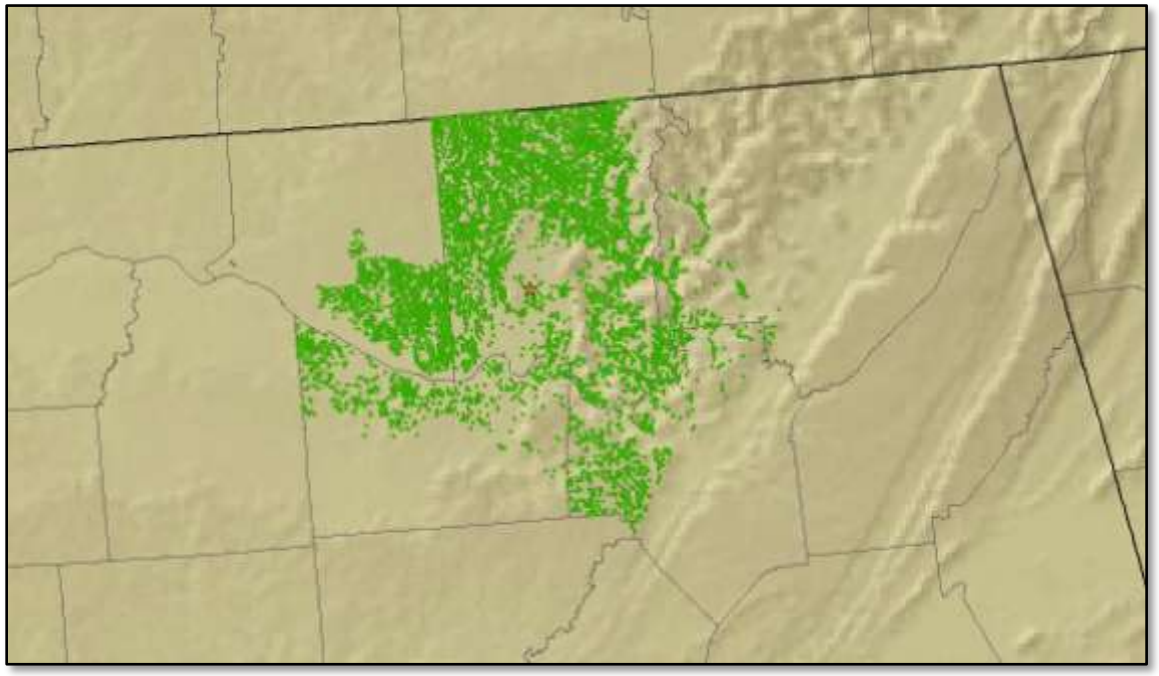


Figure 5.3 – The model's ideal biorefinery location for the first scenario along with its associated feedstock supply



BioRefinery Siting Report

Feedstock Summary

Total Feedstock Cost	$21,599,036
Total Farmgate Cost	$14,420,546
Total Transportation Cost	$7,178,498
Total Switchgrass Supply	711,837
Marginal Cost	$20.38
Average Feedstock Cost Per Ton	$30.34
Average Farmgate Cost Per Ton	$20.26
Average Transportation Cost Per Ton	$10.08

Cropland Converted To Switchgrass

Barley Acres	
Corn Acres	
Cotton Acres	70,656
Hay Acres	8,403
Oats Acres	
Rice Acres	
Sorghum Acres	
Soybeans Acres	
Wheat Acres	506

Figure 5.4 – Biorefinery siting report generated for the first scenario.

for transportation costs of hauling the switchgrass bales from the fields to the site. The average cost per ton is shown to be around $30 per ton. Also included in the report is a summary of the acres of traditional crops that were converted to switchgrass production. In this case the majority were cotton, over 70,000 acres, with hay land totaling about 8,400 acres.

The model's results for this scenario generally appear to be reasonable. The suitability analysis filtered out areas that did not meet the user's criteria for required infrastructure. Based on current crop prices, cotton acreage became the most attractive for conversion to switchgrass. The model kept transportation costs as low as possible by selecting crop zones within the county where the biorefinery was sited and pulled only enough feedstock from neighboring counties to meet the specified plant capacity, not exceeding the 70 mile driving distance limit.

5.3 – Sensitivity

For the purposes of testing the model's sensitivity, it was useful to change certain user parameters and evaluate the consequences of those changes. The first change to the initial scenario that was made was setting the urban suitability factor from "near" to "away from". In this case, the investor might be weary of situating an ethanol plant near densely populated areas due to the potentially offending odor. With all other parameters the same, the suitability analysis ran and a new output map was generated (Figure 5.5). As expected,

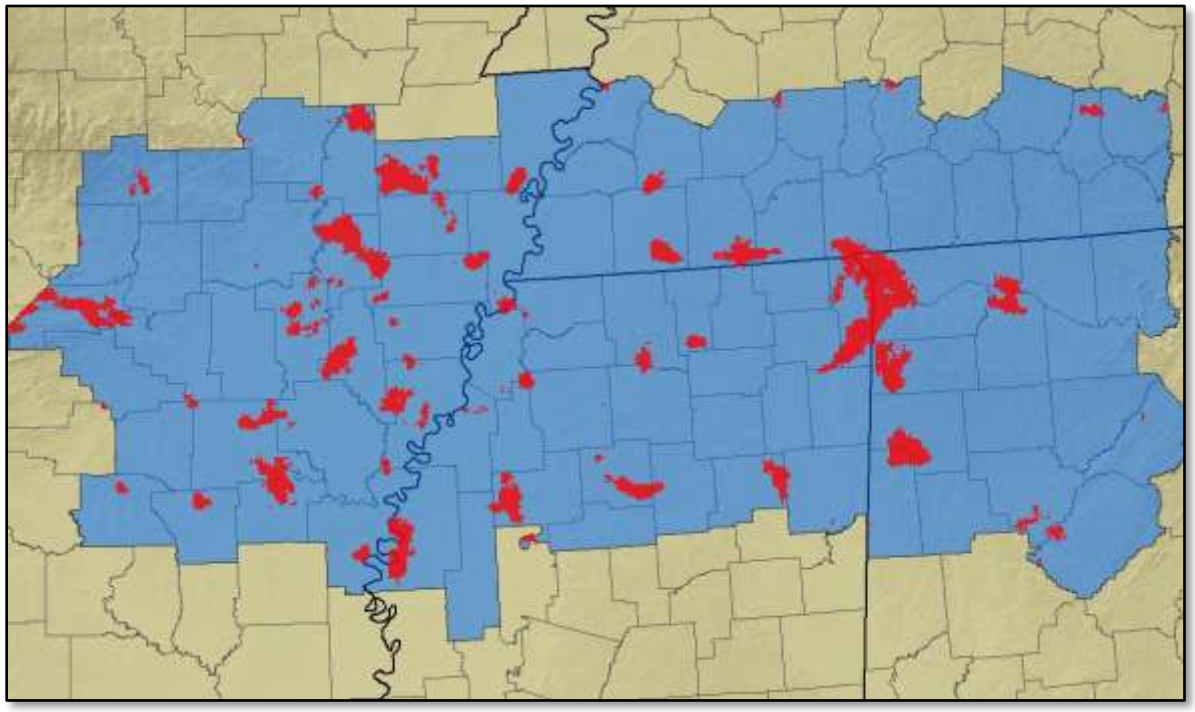


Figure 5.5 – Suitability map showing a shift to rural areas

the suitable areas shaded by the model were shifted to more rural locations that still met the road, railroad, and water criteria. A large shaded area that stands out on the map is situated on the tri-state borders of Tennessee, Mississippi, and Alabama with a number of other significant splotches appearing in eastern Arkansas and southern Mississippi.

With all other parameters the same as the initial scenario, the facility siting analysis ran, resulting in a biorefinery situated in central Alabama approximately 30 miles west of Birmingham. The feedstock supply is shown to come from 6 counties surrounding the plant (Figure 5.6). The siting report shows a total annual feedstock cost of $32.4 million with transportation costs making up over half of the total (Figure 5.7). This is not an unexpected result since the plant was sited in a more rural area where the road network is less dense and travel distances are longer. Cotton acreage, at nearly 60,000, was still

preferred for conversion to switchgrass although hay and wheat both play a larger role than in the previous run, totaling 17,300 and 1400 acres respectively.

The next task was to evaluate how the model would react to an increase in crop price; specifically in the case of this scenario, the price of cotton. With all other parameters the same (including the shift away from urban areas), the price of cotton was raised to $0.90 per bale from the default value of $0.57 and the facility siting model was run again. The results show a shift in plant location northwestward to Prentiss County Mississippi and the utilization of soybean acreage instead of cotton (Figure 5.8 and Figure 5.9). The average cost per ton is a bit higher as expected. Corn, in addition to hay and wheat acreage, begins to play a role in switchgrass production as an increase in cotton price makes other alternatives more attractive.

In the interest of evaluating the sensitivity of the model's adjustable hay availability parameter, the next run kept all previous input values the same but increased the amount

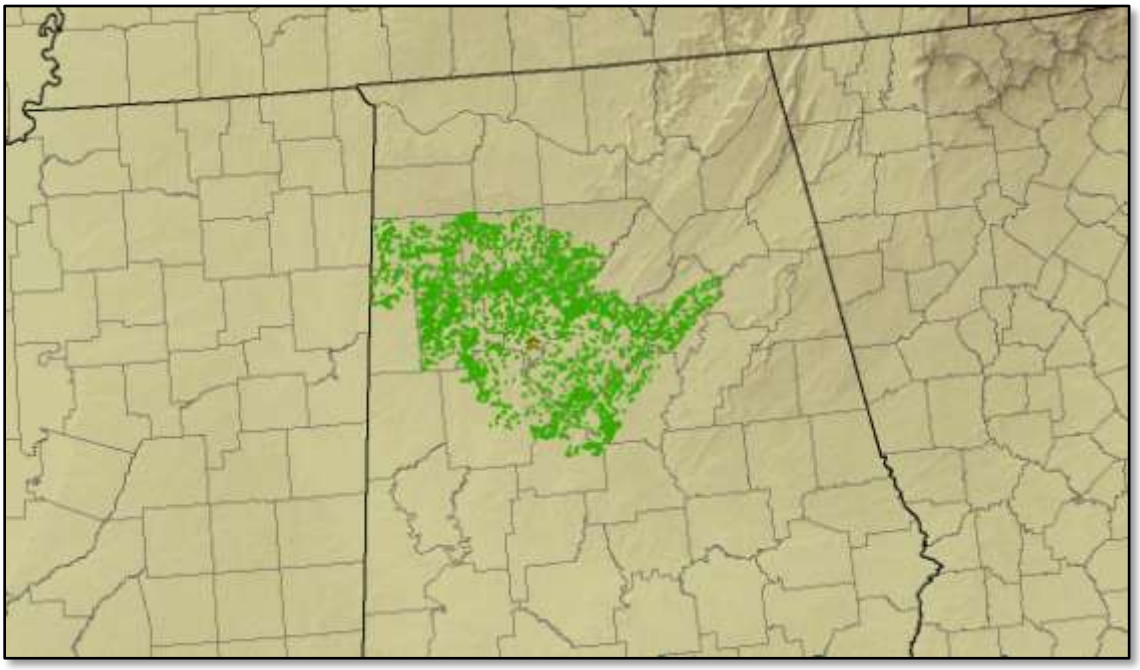


Figure 5.6 – Biorefinery location and feedstock supply map resulting from the facility siting analysis in which the modified suitability map was used.

BioRefinery Siting Report

Feedstock Summary

Total Feedstock Cost	$32,495,262
Total Farmgate Cost	$14,166,720
Total Transportation Cost	$18,328,539
Total Switchgrass Supply	700,390
Marginal Cost	$25.64
Average Feedstock Cost Per Ton	$46.40
Average Farmgate Cost Per Ton	$20.23
Average Transportation Cost Per Ton	$26.17

Cropland Converted To Switchgrass

Barley Acres	
Corn Acres	516
Cotton Acres	59,592
Hay Acres	17,321
Oats Acres	
Rice Acres	
Sorghum Acres	
Soybeans Acres	406
Wheat Acres	1,406

Figure 5.7 – Biorefinery report showing increased transportation costs.

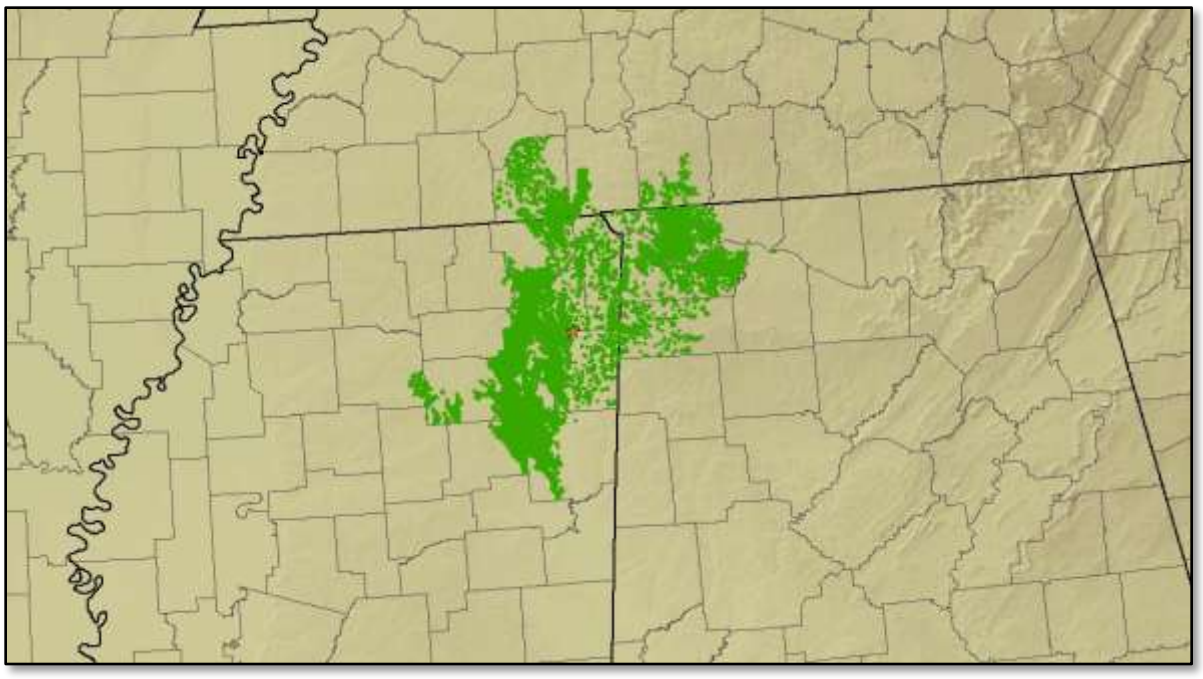


Figure 5.8 – Biorefinery location and feedstock supply map showing a shift away from cotton acreage and toward soybean acreage for switchgrass production.

BioRefinery Siting Report

Feedstock Summary

Total Feedstock Cost	$31,438,850
Total Farmgate Cost	$21,260,589
Total Transportation Cost	$10,178,272
Total Switchgrass Supply	701,001
Marginal Cost	$39.06
Average Feedstock Cost Per Ton	$44.85
Average Farmgate Cost Per Ton	$30.33
Average Transportation Cost Per Ton	$14.52

Cropland Converted To Switchgrass

Barley Acres	
Corn Acres	6,000
Cotton Acres	206
Hay Acres	6,265
Oats Acres	
Rice Acres	
Sorghum Acres	338
Soybeans Acres	67,310
Wheat Acres	1,213

Figure 5.9 – Biorefinery siting report showing a shift to soybean acreage for switchgrassproduction.

of hay acreage considered for conversion to 100%. The output of this scenario showed a shift in plant location eastward to an area just outside of Muscle Shoals Alabama (Figure 5.10). The feedstock availability map shows a dense circle of former hay land in 4 counties surrounding the biorefinery supplying it with switchgrass. The total feedstock cost and average cost per ton of feedstock drop considerably with over 76,000 acres of hay acreage converted (Figure 5.11). Allowing this much hay to be considered for conversion at the state's current hay price might not be realistic, but this exercise illustrates the model's ability to allow the user to adjust hay acreage availability to better match the conditions of their specific scenario.

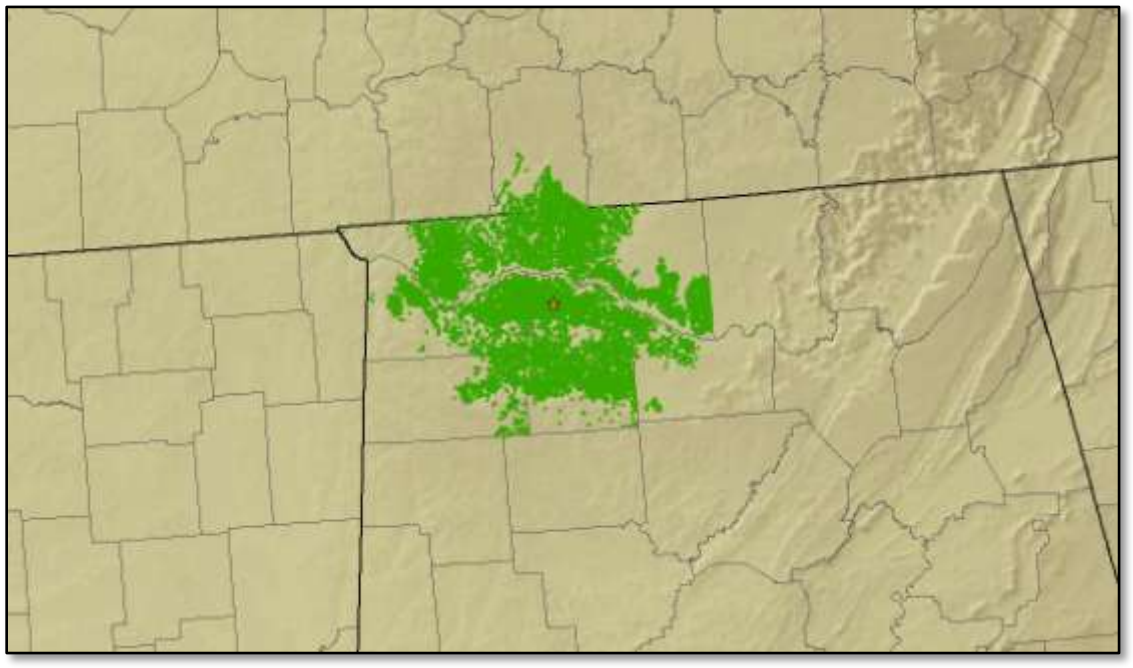
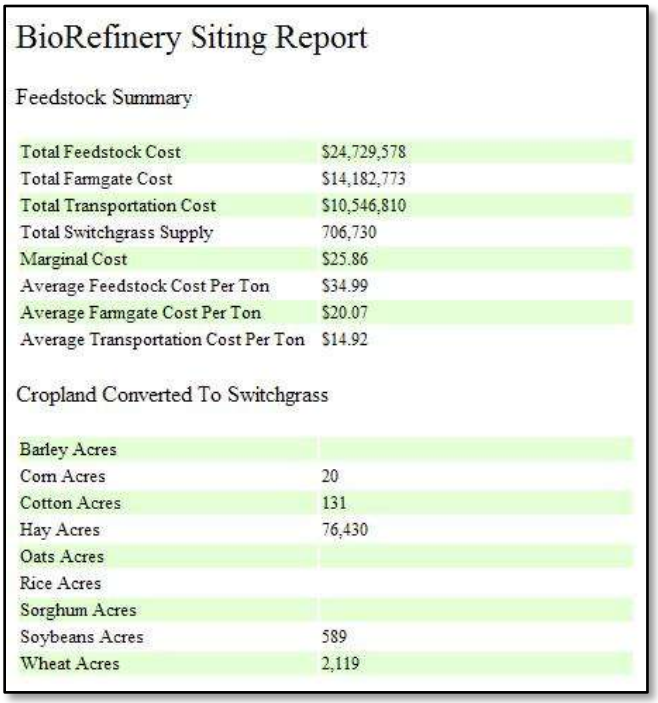


Figure 5.10 – Biorefinery location and feedstock supply map resulting from the change to allow for more hay acreage to be considered for switchgrass production.



BioRefinery Siting Report

Feedstock Summary

Total Feedstock Cost	$24,729,578
Total Farmgate Cost	$14,182,773
Total Transportation Cost	$10,546,810
Total Switchgrass Supply	706,730
Marginal Cost	$25.86
Average Feedstock Cost Per Ton	$34.99
Average Farmgate Cost Per Ton	$20.07
Average Transportation Cost Per Ton	$14.92

Cropland Converted To Switchgrass

Barley Acres	
Corn Acres	20
Cotton Acres	131
Hay Acres	76,430
Oats Acres	
Rice Acres	
Sorghum Acres	
Soybeans Acres	589
Wheat Acres	2,119

Figure 5.11 – Biorefinery siting report showing a shift to hay acreage for switchgrass production and a decrease in feedstock costs.

CHAPTER 6 – CONCLUSIONS

6.1 – Summary

With renewable transportation fuels becoming a priority in recent years finding ideal locations for biorefineries has become an important issue for the emerging ethanol industry. Investors, policy makers, and researchers are interested in evaluating feedstock availability across regions and determining where these facilities can be placed to best exploit the available resources in a cost effective manner. While supply of feedstock is certainly a dominant factor in biorefinery location, other information must be accessed to find an ideal site in the real world and the criteria varies among those looking at the issue. GIS is well suited for analyzing these types of decisions and has been used in various capacities to solve similar problems for years. At the start of this study it was thought that by investigating previous siting model efforts, a new more comprehensive and robust model could be created to better address the concerns of the emerging industry.

The primary objective of this study was to develop a comprehensive GIS model for siting cellulosic ethanol plants using switchgrass as a feedstock. The BIOFLAME software achieves this goal by providing users the ability to construct custom scenarios that not only address fundamental feedstock supply and transportation questions but also incorporate unique siting determinates to address specific needs of investors. Using techniques to optimize performance, BIOFLAME is able to provide near instant results and give the user the freedom to adjust parameters and see the consequences of the changes on the fly. The

modular nature of the software allows for relatively easy updating of input data and a simple way for new siting determinates to be plugged as they are realized. The product of this study not only improves on and extends existing similar modeling efforts but also serves as a foundation of a modeling framework that can be extended in the future in terms of spatial resolution, geographic scale, transportation modes, feedstocks, and facility types.

6.2 – Limitations

The model developed during the course of this study is not without its limitations. Currently only switchgrass is being utilized as a feedstock for cellulosic ethanol plants when in reality a variety of materials such as forest and crop residues could be included. It is quite possible that as biorefineries are constructed, they will convert multiple feedstocks and therefore the regional availability of these materials would be a dominant factor in facility location. Because switchgrass is typically baled and hauled on trucks, roads were the only transportation mode included in this model. It would be useful to have the ability to incorporate rail and barge networks as well in the event an investor or policy maker would like to consider such transportation alternatives. Using the suitability analysis the user can let ethanol demand play a role in plant location, by for example letting major cities be of high importance, but the model by nature focuses heavily on the supply side. A more comprehensive facility location model might evaluate both supply and demand when siting facilities. Such a model would certainly benefit from incorporating multi-modal

transportation not only for getting feedstock to the plant, but also getting the ethanol product to consumers.

Most of the model's results for the scenarios evaluated seem reasonable. One exception might be the occasional relatively low average farmgate cost per ton, which in some cases was estimated to be as low as $20. Some investigation into this found that the low average price came as a result of marginal lands around the region that had very low crop yields. The model by its nature will exploit the available data to find the absolute cheapest land available for switchgrass production, seeking out low yield cropland that might not actually be suitable for switchgrass in the real world. This is a limitation of the input data. Switchgrass yield data exists at the county level while traditional crop yield data is at the sub-county soil level. So if a particular crop zone is found to have a very low corn yield for example, but it's associated county is listed as having a relatively high switchgrass yield, that crop zone might be flagged as a very attractive area to convert to switchgrass production when in reality it might not be. If corn grows badly on that marginal land, switchgrass would likely not perform well either and data being fed into the model doesn't allow for this to be realized. Feeding higher resolution switchgrass yield data into the model or conversely using county level crop yield data would likely reduce the severity of these exploits.

6.3 – Future Development

The initial version of BIOFLAME builds on the ideas of similar models found in literature and provides a unique way for users to construct intricate real world scenarios for siting biorefineries and evaluating feedstock. The modular nature of the software allows for expansion in terms of scale, tabular and spatial data, and new suitability factors that may be important in the future. It is feasible that a future version could go beyond the southeastern U.S. and include the entire country. Crop zones, currently derived from generalized soil map units and county boundaries, could be brought down to the field level for a much more detailed geographic unit of analysis. Using crop growth models such as EPIC or ALMANAC, one could theoretically estimate crop yields including switchgrass yields at a very high resolution and make the data available for use in BIOFLAME. Other siting determinates relating to labor or economic polices could be added as suitability factors that influence plant location, each with adjustable levels of relative importance. Additional feedstocks beyond switchgrass and other modes for transportation would also benefit this model as would the ability to evaluate the ethanol demand influence on siting decisions.

Another area of expansion could be the incorporation of additional facility types such as pre-processing or storage facilities. An investor might be interested in collecting and storing switchgrass at various satellite facilities across a region and processing the feedstock into some form that could be packed more tightly into containers and shipped by rail or barge to a centralized biorefinery. This might allow the biorefinery to be situated

much farther away from the switchgrass fields than the 50 miles currently being considered by the model. Siting multiple facility types in this fashon using multi-modal transportation would require a substantially more complex siting algorithm but is within the realm of possibility with this modeling framework.

LITERATURE CITED

Agricultural Policy Analysis Center. 2009. POLYSYS Crop Budgets. Available online at http://agpolicy.org/polysys.html

Church, Richard L. 2002. Geographical information systems and location science. Computers & Operations Research 29: 541-562.

DuPont Danisco Cellulosic Ethanol LLC (DDCE). 2009. Tennessee Breaks Ground for its Pilot Biorefinery. News Release 10-14-2008. http://www.ddce.com/ Downloaded 3/5/2009.

Earth Resources Observation and Science Center. 2008. Digital Elevation Models. Available online at http://eros.usgs.gov/products/elevation/dem.php

Economics, Statistics, and Market Information System, USDA. 2009. Field and Miscellaneous Crop Price Per Unit. Available online at http://usda.mannlib.cornell.edu/MannUsda/homepage.do

English, Burton C, Daniel G. De La Torre Ugarte, Kim Jensen, Chad Hellwinckel, Jamey Menard, Brad Wilson, Roland Roberts, and Marie Walsh. 2006. 25% Renewable Energy for the United States By 2025: Agricultural and Economic Impacts. University of Tennessee, Biobased Energy Analysis Group.

Environmental Systems Research Institute. 2005. ESRI Data & Maps [DVD]. Available: Brock University Map Library Controlled Access G 3200 2005 E5.

Federal Emergency Management Agency. 2008. FEMA Hazard Maps. Available online at http://www.fema.gov/hazard/map/index.shtm

Kaylen, Michael, Donald L. Van Dyne, Youn-Sang Choi, Melvin Blasé. 1999. Economic Feasibility of Producing Ethanol From Lignocellulosic Feedstocks. Bioresource Technology 72, 2000, 19-32

Lambert, D.M., M. Wilcox, A. English, and L. Stewart. 2008. Ethanol Plant Location Determinants and County Comparative Advantage. Journal of Agricultural and Applied Economics April 2008.

Menard, Jamie. 2007. Personal Comment.

Miller, Harvey J. and Shih-Lung Shaw. Geographic Information Systems for Transportation. Oxford University Press, 2001.

National Agricultural Statistics Service. 2008. Data and Statistics. Available online at http://www.nass.usda.gov/Data_and_Statistics/index.asp

Noon, Charles, F. Benjamin Zhan, and Robin L. Graham. 2002. GIS-Based Analysis of Marginal Price Variation with an Application in the Identification of Candidate Ethanol Conversion Plant Locations. Networks and Spatial Economics, 2: 79-93.

Northeast Regional Biomass Program. 1994. Establishing an Ethanol Production Facility in the Northeast. Coalition of Northeast Governors Policy Research Center. http://www.nrbp.org/pdfs/pub03.pdf Downloaded 10/15/2007.

Ranta, Tapio. 2004. Logging residues from regeneration fellings for biofuel production-a GIS-based availability analysis in Finland. Biomass and Bioenergy 28: 171-182.

Renewable Energy Access. 2006. New York State Backs Cellulosic Ethanol. Article 5/12/2007. http://www.renewableenergyaccess.com/rea/news/story?id=44889 Downloaded 10/15/2007.

Reynolds, Robert E. 2002. Infrastructure Requirements for an Expanded Fuel Ethanol Industry. Downstream Alternatives Inc. Phase II Project Deliverable Report.

Smith, D. M. 1966. A Theoretical Framework for Geographical Studies of Industrial Location. Economic Geography, Vol. 42, No. 2: 95-113.

Soil Survey Staff, Natural Resources Conservation Service, USDA. Soil Survey Geographic (SSURGO) Database. Available online at http://soildatamart.nrcs.usda.gov accessed [month/day/year].

The National Commission on Energy Policy. 2004. Ending the Energy Stalemate: A Bipartisan Strategy to Meet America's Energy Challenges. www.energycommission.org Downloaded 10/15/2007.

Tiller, Kelly. 2007. UT Biofuels Initiative. Presentation for Public Hearing on the Niles Ferry Biorefinery Location, Vonore, TN. 8/16/2007.

U.S. Environmental Protection Agency. 2007. Renewable Fuel Standard Program. Article 8/30/2007. http://www.epa.gov/otaq/renewablefuels/ Downloaded 10/15/2007.

USGS Land Cover Institute. 2001. National Land Cover Database. Available online at http://landcover.usgs.gov/usgslandcover.php

Weber, Alfred [translated by Carl J. Friedrich from Weber's 1909 book]. Theory of the Location of Industries. Chicago: The University of Chicago Press, 1929.

Western Governors' Association. 2008. Strategic Assessment of Bioenergy Development in the West – Spatial Analysis and Supply Curve Development. http://www.westgov.org/wga/initiatives/transfuels/Task%203.pdf Downloaded 2-16-2009.

VITA

Bradly Scott Wilson was born in Paris, TN on February 16, 1976. He lived there until moving to Knoxville, TN to attend the University of Tennessee in 1994. In 1998 he received a Bachelor's Degree in Animal Science and began working full time for the Agricultural Policy Analysis Center as a computer programmer. In 2006 he entered the Master's program in Geography focusing on Geographical Information Science and application development.