2014-01-01

Biofuel Feedstock Optimization Considering Different Land Cover Scenarios and Watershed Impacts

Rodney Wayne Vance
University of Texas at El Paso, rvance@miners.utep.edu

BIOFUEL FEEDSTOCK OPTIMIZATION CONSIDERING DIFFERENT LAND COVER SCENARIOS AND WATERSHED IMPACTS

RODNEY WAYNE VANCE

Department of Mechanical Engineering

APPROVED:

Jose F. Espiritu, Ph.D., Chair

Heidi A. Taboada, Ph.D.

W. Shane Walker, Ph.D.

Norman D. Love Jr., Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

BIOFUEL FEEDSTOCK OPTIMIZATION CONSIDERING DIFFERENT LAND COVER SCENARIOS AND WATERSHED IMPACTS

by

RODNEY WAYNE VANCE, B.S. in Biology

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Mechanical Engineering

THE UNIVERSITY OF TEXAS AT EL PASO

December 2014

Acknowledgements

First and foremost, I would like to thank Dr. Jose F. Espiritu for his support and guidance during my graduate studies at UTEP. I would also like to thank Dr. Heidi Taboada, Dr. Norman Love, and Dr. Shane Walker for their participation on my dissertation committee. Additionally, I would like to thank Oswaldo Aguirre for his help and collaboration. Finally, I would like to thank the Department of Mechanical Engineering and the United States Department of Agriculture for providing me with financial assistance.

Abstract

With an increased demand for renewable energy production, especially the conversion of biomass to biofuels, perennial grasses are gaining interest as a renewable source of biofuel feedstocks. Identifying the trade-offs between bioenergy crop cultivation and nutrient runoff, erosion, and water requirements will be important as the demand for these crops continues to grow. The primary objective of this study is develop an integrated optimal control model that estimates the potential effects on water quality and demand and soil erosion from cultivating switchgrass and other perennial grasses instead of conventional crops at the watershed scale. The Soil and Water Assessment Tool is used to model these land use changes. In this research, we developed an optimization method based on genetic algorithms to evaluate different land cover change scenarios and their effect at the watershed level by coupling the SWAT model with a multi-objective genetic algorithm, that takes into consideration the minimization of nutrient loading, sediment yield due to erosion at the watershed outlet, the effects on regional water resources, while maximizing biomass production. The optimal control model will help further the understanding of the environmental impacts of cultivating biofuel feedstocks and is intended to aid policy makers and stakeholders when making decisions to increase feedstock production.

Table of Contents

Acknowledgements iv

Abstract v

Table of Contents vi

List of Tables viii

List of Figures ix

Chapter 1: Introduction 1

1.1 Introduction 1

1.2 Problem Statement 3

1.3 Related Research 4

1.4 Statement of Purpose 6

1.5 Objectives 7

1.6 Contributions and Broader Impacts 8

Chapter 2: Land Use, Biofuels, & Feedstocks 10

2.1 Land Use, Land Use Change, Feedstock Cultivation 10

2.2 Biofuels and Sustainability 13

2.3 Biomass and Perennial Grasses 14

Chapter 3: Hydrologic Modeling and SWAT 17

3.1 Hydrologic Modeling 17

3.2 Soil and Water Assessment Tool (SWAT) 20

3.3 SWAT Physical Processes 22

Chapter 4: Optimization 36

4.1 Optimization Methods 37

4.2 Multi-Objective Optimization 38

4.3 Mathematical Methods 40

4.4 Meta-heuristic Methods 42

4.5 Multi-objective Evolutionary Algorithms 47

Chapter 5: Methodology 50
Chapter 6: Integrated Modeling Approach 56
6.1 SWAT Setup 57
6.2 Problem Formulation 63
6.3 Solution Methodology 65
Chapter 7: Multiple Objective Evolutionary Algorithm for OCM 66
7.1 Initialization 67
7.2 Evaluation 69
7.3 Selection..70
7.4 Reproduction..74
Chapter 8: Results 77
8.1 Lake Fork Watershed Application 77
8.2 San Juan Watershed Application 81
Chapter 9: Conclusion..89
Chapter 10: Contributions and Recommendations 92
10.1 Contributions...92
10.2 Recommendations..93
References..98
Vita ……………………………………………………………………………107

List of Tables

Table 6.1: Possible Land Covers and Fertilization Applications for Each HRU......................... 64

Table 7.1: Example Fitness Values for Interval.. 71

Table 7.2: Example Normalized Objectives 72

Table 7.3: Example Euclidean Distances... 72

Table 7.4: Example Intervals for Fitness Metric #2 73

List of Figures

Figure 5.1: Flowchart of Genetic Algorithm 53

Figure 5.2: Integrated Modeling Approach 54

Figure 6.1: Flowchart of Integrated Modeling Approach Using SWAT and MOEA 57

Figure 6.2: Digital Elevation Map 58

Figure 6.3: Land Use/Land Cover Map 59

Figure 6.4: Slope 59

Figure 6.5: Soil Distribution Map 60

Figure 6.6: Delineated Watershed with Subbasins and Stream Network 61

Figure 6.5: Example Management Input File 62

Figure 6.8: Selection from Example Output File 63

Figure 7.1: Flowchart of Genetic Algorithm 67

Figure 7.2: Example Chromosome Representation of Individual Landscape Scenario 68

Figure 7.3: Example of an Initial Population 68

Figure 7.4: Example of Evaluated and Nondominated Solutions 69

Figure 7.5: Example of Fitness Metric #1 Evaluation 71

Figure 7.6: Example of Fitness Metric #2 Evaluation 73

Figure 7.7: Example of Aggregated Fitness Metric 74

Figure 7.8: Example Nondominated Solutions Selected for Elitism and Crossover 74

Figure 7.9: Example of Subsystem Rotation Crossover 75

Figure 7.10: Example of Two Point Mutation 76

Figure 8.1: Lake Fork Watershed 78

Figure 8.2: Final Delineated Lake Fork Watershed with Stream Network 79

Figure 8.3: Pareto Solutions for Lakefork Watershed. 80

Figure 8.4: Decision Alternatives for Lakefork Watershed 80

Figure 8.5: Outputs to Watershed Outlet for Lakefork Example 81

Figure 8.6: San Juan Watershed 82

Figure 8.7: Final Delineated San Juan Watershed with Stream Network 83

Figure 8.8: Pareto Solutions for San Juan Watershed. 84

Figure 8.9: Decision Alternatives for San Juan Watershed 85

Figure 8.10: Outputs to Watershed Outlet for San Juan Example 85

Figure 8.11: Example Landscape Scenario from Pareto Solutions for Lakefork Watershed. 87

Figure 8.12: Example Landscape Scenario from Pareto Solutions for San Juan Watershed. 88

Figure 10.1. Graphical User Interface (GUI) for Optimal Control Model 97

Chapter 1: Introduction

1.1 Introduction

It is clear that there is an increased demand for renewable energy production, especially the conversion of biomass to biofuels. The Energy Independence and Security Act of 2007 (EISA), has pushed for the expansion of biofuel volumes extended target dates to 2022. The new Renewable Fuel Standard required the use and production of 9 billion gallons of biofuels for 2008, and a target of 36 billion gallons in 2022. Of these 36 billion gallons at least 16 billion should be developed from cellulosic biofuels, and no more than 15 billion gallons derived from corn ethanol. Additionally, it is becoming more of a policy priority to identify sustainable approaches for the production of bioenergies. The EISA also requests that federal agencies begin to identify and report environmental concerns linked to biofuel production, for example, water and soil quality and land productivity (US Congress, 2007).

With these mandates in mind, it becomes a challenge to select suitable feedstocks and locations for cultivation. Biofuel crop selection will not be uniform and will be based on regional factors, productivity, and sustainability. In order to address sustainability of feedstock production, assessment of environmental impacts is required. Priorities for developing a sustainable biofuel industry include maximizing bioenergy crop production while reducing negative environmental impacts of land use change. To meet production demands and sustainability criteria it becomes necessary to develop innovative strategies and tools to assess production amounts and impacts.

In order to fill this biofuel feedstock demand, perennial grasses like switchgrass and miscanthus are gaining interest as an alternative to first-generation biofuel feedstocks like corn

and sugar cane. Even though corn ethanol is available, second- generation biofuels are becoming more popular. This is mainly due to the fact that second-generation biofuels do not directly compete with food production. Furthermore, second-generation biofuels can be cultivated on more marginal lands and using less intensive agriculture practices. Unlike corn, which must be replanted every year, perennial grasses can preserve and increase soil quality. Once established, perennial grasses return annually without need for replanting. Using miscanthus and switchgrass as feedstocks for ethanol production in the U.S. could considerably reduce the amount of agricultural land reassigned to bioenergy production while still aiming to meet mandated production goals. (Miguez and Villamil, 2008; Monti et al, 2007).

Agricultural management practices, like heavy use of nitrogen fertilizers, increased tillage, irrigation methods, and crop selection all have effects on watersheds and the environment. Second generation biofuel feedstock, like switchgrass and miscanthus, have the potential to reduce erosion and nutrient losses within the watershed (Blanco-Canqui, 2010; USDA, 2010). On agricultural lands, soil weathering and erosion processes along with increased fertilization are the primary means of movement of pollutants. Excessive amounts of non-point source pollutants can contribute to eutrophication of the receiving water bodies and impair water quality.

The production of bioenergy crops may alter the hydrology and ecosystem services of a particular region, but the impacts may not always be the same. Impacts may be negative or neutral depending on crop selection and management practices, while others may offer positive benefits and improve water quality. It is crucial to identify scenarios that represent better land use and management practices in order to limit the environmental impacts that occur from agricultural watersheds.

1.2 Problem Statement

It is important to identify the trade-offs between bioenergy crop cultivation, environmental impacts, and ecosystem services. Contemporary, intensive agriculture systems tend to concern only a single ecosystem service, either the production of ample food, fiber, feed, or biomass. However, in order to promote a more sustainable approach to cropping systems, it is important to manage these systems using an ecological approach where management decisions are based on complete ecosystems and all environmental costs and benefits are assessed. Rather than just one ecosystem service, agricultural landscapes can be managed for multiple services that have the potential to minimize the negative environmental impacts of modern agriculture (Robertson and Swinton, 2005). Non-point source pollution contributes heavily to the negative impacts of environmental and water quality and is a major source of pollutants to impaired water bodies, which include sediment and nutrient run off from agricultural sources. The approach must consider both environmental goals like pollution reduction and the production of agricultural commodities, in this case, biofuel feedstocks. Better management of agricultural watersheds is important and in order to analyze these land use changes, scenarios that represent different land use/land cover arrangements and management practices need to be simulated. However, it is time consuming and tedious to manually test all possible scenarios of land cover change. This leaves a requirement for an integrated modeling approach. The development of an integrated modeling system to evaluate sediment, nutrient, water requirements, and crop yields as a result of applied decision alternatives, like cultivating perennial grasses in place of conventional crops is necessary to acquire a better understanding of the impacts at watershed levels. This approach requires a hydrologic and environmental model that can predict the impacts

at watershed levels, and a heuristic multiple objective evolutionary algorithm that can search and identify all possible scenarios that meet specified objectives.

1.3 Related Research

Recent articles have addressed the importance of developing decision support tools that include ecosystem services in order to minimize the adverse environmental impacts associated with agricultural systems and biofuel feedstock cultivation (Tallis and Polasky, 2009; Engel et al, 2010). However, there has not been that much research or information that quantifies water and soil quality, and other ecosystem services with the expected increases in biofuel feedstock production. Related research has not identified the environmental impacts associated with the conversion and cultivation to biofuel feedstock crops and has not investigated the topic from the approach of managing agricultural landscapes. Additionally, very few have coupled USDA's Soil and Water Assessment Tool (SWAT) with evolutionary optimization algorithms.

Those that have coupled SWAT with genetic algorithms have created optimal control models mainly for best management practices and detention basin locations to reduce pollutant loads and pesticide control. For example, Kaini and Artita have developed a variety of methods for evaluating cost-effective, optimum combination of detention ponds, parallel terraces, filter strips, and other best management practices to reduce pollutant loads by coupling a genetic algorithm with SWAT. The optimal control models where able to identify least cost combinations dependent on size, type, and location of best management practices (Kaini & Artita 2007, 2008, 2009, 2010). Maringanti et al. (2008 and 2011) developed a multi-objective tool for the selection and placement of best management practices for pesticide control by combining a genetic algorithm with a distributed parameter watershed model. According to Maringanti, other optimization models that had a dynamic linkage with water quality models, could only analyze

smaller watersheds due to increased computation time. Instead of having a direct interface with the watershed model, Maringanti, developed a database of pollution and cost information of different BMPs under consideration, allowing them to apply model at a much larger scale. Gitau et al. (2004 and 2006) also utilized a BMP database to optimize BMP placement and cost with a genetic algorithm and SWAT. Additionally, Muleta and Nicklow (2002) developed an integrative modeling approach to simultaneously limit sediment yield and maximize farm-level profit by coupling SWAT with a multi-objective evolutionary algorithm known as Strength Pareto Evolutionary Algorithm (SPEA). However, the integrated model simulated only at the approximate spatial scale of a farm, without addressing the importance of watershed management practices.

Other studies have utilized optimization techniques for calibration and validation of the SWAT model for their particular case studies. Parameter calibration studies have been conducted by Zhang et al, by applying single-objective and multi-objective optimization genetic algorithms to SWAT to optimize the parameters of the model using observed stream flow data (Zhang et al, 2008 and 2010).

Very little land use change research using SWAT to model environmental impacts have been performed. These projects only evaluated a few scenarios rather than the multiple combinations that can be evaluated by coupling SWAT with an evolutionary algorithm. For example Ling et al used SWAT to estimate the potential effects on riparian nitrate loads when cultivating *Miscanthus x giganteus* in place of conventional crops in the Salt Creek watershed in East-Central Illinois. However, only four scenarios of 0%, 10%, 25%, and 50% land use changed to miscanthus production were modeled (Ling et al, 2010).

1.4 Statement of Purpose

In order to sustainably cultivate biofuel feedstocks while meeting production demands and simultaneously fostering healthy ecosystems, assessment of environmental impacts in response to agricultural landscape changes is required. Suitable management options can be evaluated and implemented at the watershed scale to help minimize negative environmental consequences to natural resources and potentially increase ecosystem services.

The primary objective of this research was to develop an integrated modeling approach that can identify optimal agricultural landscapes in regards to bioenergy feedstock production and ecosystem services. The resulting optimal control model (OCM) ensures sustainable biofuel production systems for the future by estimating the potential effects on water quality and soil erosion from cultivating miscanthus, switchgrass, and other perennial grasses in place of conventional crops and pastures at the watershed scale. The OCM evaluates multi-objective management of land use change and identifies potential tradeoffs that may occur. The integrated model can assess multiple land covers and determine optimal land use and management scenarios that have the potential to improve the production of biofuel feedstocks while minimizing environmental impacts. The model effectively identifies optimal agricultural landscapes and quantifies the tradeoffs between the conflicting objectives of improving watershed health and quality while maximizing feedstock production. Thereby, it also serves as a decision support tool for watershed management that can aid in identifying suitable agricultural landscapes for the production of bioenergy feedstocks at a watershed scale.

1.5 Objectives

Objective 1: Develop an optimal control model by integrating a hydrologic and environmental model with a multiple objective evolutionary algorithm that can search and identify optimal landscape scenarios for feedstock production.

Objective 2: Quantify and identify the environmental impacts and tradeoffs of land use changes and management practices on water quality and ecosystem services when converting traditional land uses to biofuel feedstock production.

Objective 3: Optimize selection and location of biofuel feedstock crops that will promote healthier ecosystems and agricultural landscapes, while maximizing biomass production yields.

In order to complete the above objectives, the following tasks were completed:

1. Develop code in MATLAB that integrates the Soil and Water Assessment Tool (SWAT) with a multi-objective evolutionary algorithm (MATLAB).
2. Parameterize SWAT for second-generation biofuel feedstock crops.
3. Integrate a multi-objective genetic algorithm with SWAT.
4. Identify and quantify the trade-offs (environmental impacts) associated with increased biofuel feedstock cultivation.
5. Optimize the selection and placement of various biofuel feedstock crops and management practices under multi-objective functions.

The developed integrated modeling system has the ability to assess management impacts and land use changes on watershed quality and soil erosion. To illustrate its broad application, it

was applied to two different example watersheds, Lake Fork watershed in Texas, that has a total drainage area of 487 km^2, and the San Juan watershed in Mexico that spans 4136 km^2, to evaluate nitrogen, phosphorous, and sediment yields to the watershed outlet while maximizing feedstock production.

1.6 Contributions and Broader Impacts

The long-term goals of this research is to aid science based decisions for biofuel feedstock production in the context of promoting ecosystem service management while increasing water quality and healthy ecosystems at the watershed scale. The research will provide significant contributions to the fields of environmental and water management, and biofuel feedstock production. More specifically the research outlined above will:

- Develop a comprehensive modeling system for the management of environmental objectives in a watershed that has the capacity to evaluate agricultural production of biofuel feedstocks.
- Apply a hydrologic simulation model and a multi-objective evolutionary genetic algorithm to identify optimal agricultural landscape scenarios.
- Will help further the understanding of the environmental impacts of cultivating second-generation biofuel feedstocks and is intended to aid policy makers, water resource engineers and planners, when making decisions to increase feedstock production.
- Identifying the spatial distributions of biofuel feedstocks within a watershed will result in maximum biomass production, mitigate negative environmental impacts, and facilitate the displacement of biofuels derived from corn, reducing direct and indirect land use changes.

Identifying the tradeoffs between the competing objectives will help to improve water quality while meeting the demand of renewable energy production, when cultivating biofuel feedstocks. With minimal modification, users can apply additional or alternative objectives and constraints to the integrated modeling system to identify optimal land use scenarios in any watershed.

Chapter 2: Land Use, Biofuels, & Feedstocks

2.1 Land Use, Land Use Change, Feedstock Cultivation

Humans alter the Earth's system in a variety of ways, but the transformation and appropriation of land to produce goods and services produces the most impacts. Land use change altars ecosystem structures and performance, and effects biological system interactions with surrounding areas (Vitousek et al, 1997).

According to the most recent inventory of U.S. major land uses, cropland represents 18 percent of the United States total land area with 408 million acres in 2007. The USDA's Economic Research Service identifies cropland as an aggregate of five types of acreage. Three types of cropland are devoted to crop production: cropland harvested, crop failure, and cultivated summer fallow. The two types not directly used for crop production are cropland pasture and idle cropland but may be used for production in following years. Eighty two percent of the 408 million acres was cropland used for crops. The remaining eighteen percent was used for pasture or was idle in 2007. Cropland use for pasture varies in quantity and quality depending on the region. Pasture in the Corn Belt is generally classified as good quality cropland, whereas the pasture in the Plains regions and much of the South has higher proportions of marginal cropland (Nickerson et al, 2011). These marginal croplands may be more appropriate for cultivating second-generation biofuel feedstocks.

The use of crops to produce renewable energy has seen an increase over the last 30 years, mainly due to energy policies and congressional mandates requiring increased amounts of renewable fuels in the United States energy supply (Nickerson et al, 2011). Among these renewable energies are biofuels, a low carbon alternative to petroleum. Energy policies, such as the Renewable Fuel Standard and the Energy Independence and Security Act of 2007, have set

an increase of biofuel use and production, from 9 billion gallons in 2008 to 36 billion gallons in 2022. This increases the demand for existing commodities and also creates new markets for perennial energy crops, which ultimately changes agricultural landscapes (Marshall et al, 2011; Biomass Research & Development Board, 2008). With recent legislation calling for more production of renewable biofuels to reduce dependence on oil imports and lower greenhouse gas emissions, the agricultural sector has met the challenge of providing food, feed, fiber, and now fuel needs. Production of biofuels has many benefits, including the creation of new biomass industries that would the revitalization of rural communities and the farming sector. Furthermore, a revitalization of agriculture and forestry would lead to social stability and an increased stimulus to the economy. Additionally, the conversion of biomass into energy would help decrease the requirement and use of fossil fuels while decreasing greenhouse gas emissions (Saga et al, 2008). However, as their popularity and production increase, so do concerns about the magnitude of environmental impacts from land use transformations.

USDA's ERS has identified areas of concern that arise from this increased demand for land to produce bioenergy feedstocks. The amount of feedstock required to meet projected biofuel demands, land requirements and locations, crop yields, and the productivity of these lands, should all be considered when determining land use impacts and biofuel policy (Marshall et al, 2011). Increased biofuel feedstock production can come from three areas to accommodate the demand: acreage not currently in production, acreage shifted from other crops, or increased productivity. The expected profit and environmental tradeoffs of alternative land uses will control how land is will be designated for bioenergy feedstock production, conventional crop cultivation, and non-crop uses like grazing and idling. Other factors like geography, soil type,

and agricultural management practices can create considerable variations in the environmental impacts among land use change (Biomass Research and Development Board, 2008).

The conversion of land to produce biofuel feedstocks can be classified as either a direct or an indirect land-use change. Converting land directly into feedstock production represents a direct land-use change. However, if feedstock production displaces conventional crop production, which results in the conversion of a grassland or forest for that crops production elsewhere, that would represent an indirect land-use impact (Marshall et al, 2011).

Land-use shifts have the potential to create negative and positive impacts on the environment. Shifting from one crop to another can affect the land as well as the availability of the displaced crop. These shifts affect prices and create different impacts to environment, changing the function of that land relative to goods and services. In order to achieve sustainability in biofuel production, it is important to identify the environmental tradeoffs of cultivating feedstocks (Biomass Research and Development Board, 2008). Other important factors in biofuel production include planting decisions, crop selection, land use, management practices such as irrigation and fertilization amounts and timing, all have effects on water quality, soil erodibility, and other environmental impact. The environmental impacts from agricultural shifts in production vary by region. As the nation continues to demand more biofuel production and markets emerge, the agricultural landscape will continue to change (Malcolm et al, 2009).

According to the IEA Executive Committee (2007), future biomass production on different land types can fall into one of three categories, energy farming on current agricultural land, energy farming on marginal and degraded land, and biomass waste and residue collection. Current agricultural land, arable and pasture, can contribute large amounts of biomass and at low

production costs, especially when cultivating perennial crops. Additionally, with lower productivity and higher costs, marginal and degraded lands can also contribute to the production of biomass. The higher costs are attributed to higher upfront investments for the regeneration of land, but at the same time there is less competition with other land uses and the added potential benefit of soil improvement. Lastly, organic wastes and residues can supply an additional yet smaller amount of biomass.

2.2 Biofuels and Sustainability

In order for biofuels to become fully integrated into the U.S. market, they must be environmentally, economically, and socially sustainable. Being able to cultivate feedstocks while ensuring that the food, feed, and fiber supplies remain available, and while maintaining a healthy environment are necessary for sustainable production. With an increase in feedstock production, water quality and supply may be a limiting factor in some locations. It is important to ensure the availability of water for society as well as irrigation and processing of biofuels. Advances in sustainable management and conservation practices will be necessary to optimize efficient water use and maintain a balanced water demand. Additionally, water quality enhancement is essential to the sustainable production of feedstocks. Sedimentation from erosion, nutrient input and flows, and pesticides runoff, all have negative impacts on water quality and aquatic habitats. However, these impacts will vary by region, depending on soil type, season, crop, and farm management practices. As feedstock production increases and becomes more geographically dispersed, it will become more important to assess the cumulative effects over broad spatial scales and in larger bodies of water (Biomass Research and Development Board, 2008)

There are many concerns rising over the environmental tradeoffs of producing biofuels.

Probably the most important concern to address is the development of strategies that reduce the amount of land allocated for feedstock production. Biofuel production per acre varies widely. For example, estimates have shown rapeseed producing 100 gallons per acre, while corn produces 400 gallons per acre, and sugarcane up to 660 gallons. However, cellulosic ethanol has seen even larger estimates of per acre ethanol yields, more than 1,000 gallons, which would greatly reduce land requirements (Coyle, 2010). This is why second generation biofuels, like cellulosic ethanol, are gaining more popularity. Second generation biofuels, those made from non-food feedstocks and wastes, have the potential to lessen the effects on food and feed markets, impacts on environment, and their production can be more geographically dispersed utilizing marginal croplands. Second generation biofuels utilize widely available biomass, crop residues, and perennial energy crops such as switchgrass and miscanthus (Coyle, 2007).

Cellulosic ethanol production offers many advantages over grain-based production of ethanol. Potential benefits include the ability to use a wide variety of fast growing, low value, perennial non-food crops as feedstocks and that produce higher ethanol yields. Additionally, these feedstocks require less management and resources, and can be grown on marginal lands. Overall, these benefits increase land use efficiency, cost effectiveness, and provide feedstock production pathways with less environmental impacts (Marshall et al, 2011).

2.3 Biomass and Perennial Grasses

Biomass has become one of the most pivotal bioenergy options to date (Hamelinck and Faaij, 2006) and according to IPCC (2007) will continue to hold this popularity for the next fifty years. Biofuel markets have increased along with production capacity, competition, and international trade flows. This is leading to much debate toward the sustainability of biofuel production (Demirbas, 2009). Biomass is unique from other forms of renewable energy in the

fact that it can be converted to multiple forms of energy storage such as liquids, gases and solids that can provide mechanical power generation and heat or be used for electricity. Biomass can by utilized in three potential ways. Biomass can simply be burned to generate heat and electricity, converted to gas-like fuels, or used to produce liquid fuels. Liquid fuels, or biofuels, generally consist of ethanol and methanol. The popularity of biofuels is increasing, due to the fact that one-third of the United States energy supply is required for transportation and biofuels have the potential to supply this demand (Tewfik, 2004). National policies of various countries are beginning to set increased targets and expectations for bioenergy production. These ambitious and challenging long-term energy scenarios require that there be a sufficient amount of biomass resources and biomass markets that can guarantee reliable and sustainable approaches (IEA, 2007). There is a vast amount of biomass feedstocks available, including but not limited to forest products, biorenewable wastes, energy crops, aquatic plants, and food and sugar crops. This research will focus on energy crops, in particular grasses.

Perennial grasses like switchgrass and miscanthus are gaining interest as an alternative to first-generation biofuel feedstocks like corn and sugar cane. Unlike corn, which must be replanted every year, perennial grasses can preserve and increase soil quality. Once established, perennial grasses return annually without need for replanting. Using miscanthus and switchgrass as feedstocks for ethanol production in the United States could considerably reduce the amount of land allocated to biofuel production while continuing to meet government mandated production goals (Miguez and Villamil, 2008; Monti and Fazio, 2007).

Switchgrass and miscanthus can be cultivated on marginal lands in dry regions. Switchgrass is a prairie grass that is native to many U.S. regions and well adapted to the Midwest, Southeast, and Great Plains. With many environmental benefits, including the

improvement of carbon balances, improved soil and water quality, switchgrass is becoming a promising biofuel crop. Additionally, perennial grasses can be grown across a wide variety of conditions and yields large amounts of biomass. Estimates of up to 500 gallons of ethanol per acre have been given for current varieties, and with continued improvements and innovation switchgrass can help meet biofuel goals (Biomass Research and Development Board, 2008). Both switchgrass and miscanthus are ideal biomass crops due to their efficient use of nitrogen, good pest and disease resistance, and potential to reduce nitrate leaching (Miguez and Villamil, 2008; Monti and Fazio, 2007). These are all desirable characteristics for sustainable production and improved environmental services.

Agricultural management practices, like heavy use of nitrogen fertilizers, increased tillage, irrigation methods, and crop selection all have effects on watersheds and the environment. The planting of perennial grasses has the potential to reduce erosion within the watershed; however, increased fertilization and irrigation may have negative impacts on water quality. On agricultural lands, soil weathering and erosion processes along with increased fertilization are the primary means of movement of pollutants. Excessive amounts of non-point source pollutants can contribute to eutrophication of the receiving water bodies impair water quality. It is crucial to identify better land use scenarios and management practices that limit the environmental impacts that can occur in agricultural watersheds from the increase of feedstock production.

Chapter 3: Hydrologic Modeling and SWAT

3.1 Hydrologic Modeling

Hydrologic modeling requires the simulation of physical, chemical, and biological processes occurring within a hydrologic system or watershed. Watershed models are used to simulate environmental processes that include the flow of water, sediment, chemicals, nutrients, and microbial organisms. Watershed models also quantify the effects that human activities have on these processes (Singh and Frevert, 2005). Being able to simulate these processes allows users to address many different water and environmentally related problems and topics in a much more simplified manner.

Almost all environmental and water management problems require the use of watershed models to identify appropriate solutions. Watershed models are becoming essential tools in addressing a multitude of environmental and water resource problems. Watershed models have been used to explore flooding, droughts, erosion, nonpoint source pollution, irrigation, and many other aspects of water resource planning, development, and design. At the field scale, watershed models can be used for many purposes, for example, management and conservation practices. These practices include the planning and evaluating of soil conservation, irrigation, wetland restoration, stream restoration, and water-table management. Additionally, at the large scale, models can be utilized to develop and evaluate flood protection projects, drought management, floodplain management, and water quality and supply issues (Singh and Woolhiser, 2002).

Hydrologic models can be classified in different ways. One particular way of distinguishing hydrologic models is by first classifying them as either symbolic or material (Singh, 1988). A material model represents a real physical system by another system that has similar properties, but is less challenging to work with. Material models are further classified as

scale or analog. A scale model is a system that resembles the original, just at a reduced scale. However, an analog model uses a system that has similar physical properties, but measures different responses. For example, the flow of water can be represented by the flow of electricity.

On the other hand, symbolic models, are models not represented by real physical systems, but are models of mathematical nature. These models use symbolic expressions in a logical sense, to represent an idealized situation that shares the properties of the original system (Singh, 1988). Mathematical models can be further defined as theoretical, empirical, or conceptual. A theoretical model or physically based model, also known as a white-box model, is based on basic physics' equations and primary laws governing hydrologic processes. Empirical models, also called black-box models, utilize approximate equations that contain parameters determined by data analysis. The parameters of such models generally do not have much physical significance placed on them and are determined from observations of input-output relationships. Due to this fact, they are also referred to as input-output models. An example of an empirical model is a time series model. Lastly, conceptual models, also called gray-box models, are a combination of theoretical and empirical models. Such models are based on highly simplified forms of governing physical laws and processes along with observed data (Wong and Koh, 2008).

Whether the mathematical model is theoretical, empirical or conceptual, it can also be categorized as linear or nonlinear. A linear model utilizes constant parameters that do not vary during the simulation. Alternatively, nonlinear models use dependent parameters that vary during the simulation. The choice to use a linear or nonlinear model depends on how significant the parameter value variances relate to other parameters. Additionally, a model can be considered time-invariant if its input-output relationships are not affected by time; otherwise, it is a time variant model. Since climate varies by season and affects physical processes in hydrologic

systems differently, most hydrologic systems are considered time-variant. Models can also be categorized as lumped or distributed. Models are considered lumped when the entire watershed is represented as one homogenous unit; whereas, distributed models take into consideration spatial variability. Lastly, models are classified as deterministic if no random variables are used, or stochastic if random values are generated for the model inputs (Singh, 1988).

This research involves the analysis of watershed responses for spatially and temporally varied land use, land cover, and management practice scenarios. The long-term environmental impacts of these scenarios will be evaluated, making it necessary to use a continuous-time model. A continuous-time hydrologic model generally can simulate many processes including, precipitation, evapotranspiration, soil moisture, and snowmelt, all while taking into account seasonal variances.

There is a vast assortment of watershed models available and their diversity is equally as expansive. This variety of models allows users to easily find an applicable model for almost any real world water management problem. Most models are rather comprehensive and can be suitable for a wide range of problems by simulating reasonable well the basic hydrologic processes in space and time. When selecting an appropriate hydrologic model, it is necessary to take into account the spatial and temporal scales the model was designed for, as well as the data requirements. In this case we found the Soil and Water Assessment Tool (SWAT) to be a suitable model. By hydrologic model classification, SWAT is a physically-based, semi-distributed, continuous-time hydrologic model that was developed to predict the long term impacts of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use, and management conditions. This model is effective in simulating pollutant loadings, such as sediment and nutrient yields to streams and

rivers as a result of spatially varied land use and land cover scenarios and management practices. The model requires specific information regarding weather, topography, soil properties, vegetation, and land management practices to simulate processes such as surface and subsurface flows, sediment transport, nutrient transport and cycling, and crop growth (Arnold et al, 1998; Neitsch et al, 2011). Included within the SWAT model is a United States weather generator model that is capable of generating climatic data using monthly average data summarized over a number of years, and useful for filling in missing gaps in measured data. SWAT is capable of operating on daily or sub-daily temporal scales depending on the data that is available for that time scale. An added benefit of the model is that it can be set up on an ArcGIS platform, which allows users to easily prepare inputs and visualize model outputs. Additionally, the watershed data required by SWAT is easily accessible from government agencies (Neitsch, 2011).

There has been considerable use of the SWAT model in the United States and is available for free in the public domain. The SWAT model provides an excellent foundation for the integration of a modeling tool and optimization technique to provide evaluation of biofuel feedstock and management selection and placement and evaluation of erosion and nutrient cycling ecosystem services as impacted by land use change and increase in feedstock cultivation.

3.2 Soil and Water Assessment Tool (SWAT)

When modeling with SWAT, the first step is to define a watershed area. SWAT requires a digital elevation model (DEM) in order to delineate the watershed of interest and subdivide it into subbasins. SWAT determines flow directions and stream networks are generated, flow accumulations, and the user can determine a critical source area in order to set the minimum drainage area that is required to form the origin of a stream. Finally, a watershed outlet is selected to complete the delineation process. The convergence of two tributaries defines an

outlet within a stream network, and results in a subdivision of the watershed creating different subbasins.

The SWAT model is classified as a semi-distributed model. It uses digital land use maps and soil maps to categorize the land uses and soil types in each subbasin of the watershed. A digital elevation model (DEM) is used to identify the drainage network and define the watershed boundaries and drainage morphology (such as area, slope, and length) for the main basins. In order to set up a watershed simulation, the watershed first needs to be delineated and partitioned into subunits or subbasins. The user can further define the subdivision of subbasins into smaller hydrologic response units (HRUs) based on unique combinations of land use, soil type, and management practices. An HRU is defined as a lumped area that has its own specific and unique land, soil, and management criteria. SWAT gives the user two options when subdividing subbasins into HRUs. One option is to represent the subbasin by its dominant land use and soil type, thus creating an equal number of HRUs as subbasins in the watershed. The second option allows users to define threshold levels for land use and soil types, which subdivides each subbasin into multiple HRUs. Being able to subdivide the subbasins into HRUs creates increased variability of model inputs and delivers more precise simulations of the physical processes occurring in the watershed. The creation of HRUs also helps to simplify a simulation run by grouping together all similar soil and land use areas into a single response unit rather than simulating many individual areas. Unfortunately, these multiple HRUs are identified without reference to their location. However, by using the first option, the HRUs will represent individual subbasins and ensure hydrologic connectivity between HRUs and their location can be determined.

3.3 SWAT Physical Processes

Since SWAT is a physically based model, it uses various algorithms to simulate the physical processes within a watershed. Instead of using regression equations to describe input and output relationships, SWAT requires specific data about weather, soil, topography, vegetation, and management practices for the watershed being studied. The physical processes are then directly modeled by SWAT using the input data. Significant to this study are hydrologic processes, sediment and nutrient transport mechanisms, along with plant dynamics and management practices. SWAT's Theoretical Documentation gives detailed descriptions of all processes modeled and the parameters associated with them (Neitsch, 2011).

3.1.1 Hydrologic Processes

Regardless of the type of problem being studied using SWAT, the main driving force underlying all that happens in the watershed is water balance. In order to accurately estimate processes like movement of pesticides, sediments, and nutrients, the hydrologic cycle must be consistent with what is happening in the watershed. Hydrology of a watershed can be divided into two components, the land phase and the routing phase. The land phase controls the loading amounts of water, sediment, nutrients, and pesticides to the main channel in each subbasin. The water or routing phase relates to the movement of water and sediments through the subbasin channels and to the outlet. SWAT simulates the hydrologic cycle by using the water balance equation:

$$SW_t = SW_O + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{gw}) \quad (3.1)$$

where SW_t is the final soil water content (mmH_2O), SW_O is the initial soil water content on day i (mmH_2O), t is the time (days), R_{day} is the amount of precipitation on day i (mmH_2O), Q_{surf} is the amount of surface runoff on day i (mmH_2O), E_a is the amount of evapotranspiration on day i (mmH_2O), w_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mmH_2O), and Q_{gw} is the amount of return flow on day i (mmH_2O) (Neitsch, 2011).

SWAT increases accuracy and quality of the physical description of the water balance by enabling the subdivision of the watershed into smaller units. This allows the model to consider different evapotranspiration rates for various crops and soils. Additionally, runoff is estimated for each individual HRU and then routed in order to calculate the total runoff for the entire watershed.

3.1.2 Surface Runoff

The flow that occurs along a sloping surface is known as surface runoff, or overland flow. Surface runoff occurs when the infiltration capacity of the soil is less than the intensity of precipitation. To calculate surface runoff volume, SWAT utilizes a procedure that incorporates a modified version of USDA's Soil Conservation Service Curve Number (Soil Conservation Science, 1972). SWAT estimates surface runoff volumes for each HRU using this empirical procedure that can be expressed as follows:

$$Q_{surf} = \frac{(R_{day} - 0.2S)^2}{R_{day} - 0.8S} \tag{3.2}$$

Where Q_{surf} is the accumulated runoff or rainfall excess ($mm\ H_2O$), R_{day} is the rainfall depth for the day (mm), and S, the retention parameter (mm), is a function of the Curve Number, CN. The retention parameter can be calculated by

$$S = 25.4\left(\frac{1000}{CN} - 10\right) \tag{3.3}$$

where CN is the Curve Number for the day, which is a function of soil permeability, land use and soil moisture conditions.

In order to predict sediment loss and nonpoint source pollution, it is necessary to identify the peak runoff rate, which is a good indicator of the erosive power of the runoff. To do this, SWAT uses a modified rational formula that is based on the assumption that the rate of runoff increases until the basin reaches full concentration, and the entire subbasin area is adding to the flow at the outlet (Williams, 1975). The formula is

$$q_{peak} = \frac{\alpha_{tc} \cdot Q_{surf} \cdot Area}{3.6 \cdot t_{conc}} \tag{3.4}$$

where q_{peak} is the peak runoff rate ($m^3\ s^{-1}$), α_{tc} is the fraction of daily rainfall that occurs during the time of concentration, Q_{surf} is the surface runoff ($mm\ H_2O$), $Area$ is the subbasin area (km^2), t_{conc} is the time of concentration for the subbasin (hr) and 3.6 is a unit conversion factor (Neitsch, 2011).

3.1.3 Erosion Process

Soil weathering and erosion includes the detachment, transport, and deposition of sediments, nutrients, and pesticides from land surfaces to streams and rivers. SWAT simulates the erosion process in order to determine sediment, nutrient and pesticide loadings.

SWAT uses the Modified Universal Soil Loss Equation (MUSLE) to compute the amount of erosion caused by rainfall and runoff (Williams, 1995). The MUSLE is a modified version of the Universal Soil Loss Equation (USLE) that was originally developed by Wischmeier and Smith (1965). USLE estimates erosion as a function of rainfall energy. However, in MUSLE, a runoff factor replaces the rainfall energy factor, which improves the sediment yield prediction. This runoff factor also eliminates the need for delivery ratios, which are required by USLE, due to the rainfall factor only accounting for energy used in detachment of sediments, whereas the runoff factor accounts for both energies used in detaching and transporting sediment. The MUSLE is expressed as

$$S_{yield} = 11.8 \text{ x } (Q_{surf} \text{ x } q_{peak} \text{ x } A_{hru})^{0.56} \text{ x } K \text{ x } C \text{ x } P \text{ x } (LS) \text{ x } F_c \qquad (3.5)$$

where S_{yield} is the sediment yield on a given day (metric tons), Q_{surf} is the surface runoff volume ($mm\ H_2O/ha$), q_{peak} is the peak runoff rate (m^3/s), A_{hru} is the area of the HRU (ha), K is the USLE soil erodibility factor, C is the USLE cover management factor, P is the USLE support practice factor, LS is the USLE topographic factor, and F_c is the coarse fragment factor (Williams, 1995)

3.1.4 Nutrient Transport

Soil weathering and erosion processes are the major causes of nutrient transport from land into water systems. Uncontrolled and excessive nutrient loadings to streams and rivers create water bodies that are not suitable for consumption and can accelerate eutrophication. SWAT models the movement of nutrients such as nitrate, organic nitrogen and phosphorous, and soluble phosphorous from land areas to stream networks.

When pH is at a normal level, most soil minerals are negatively charged, which in turn creates a repulsive interaction with anions such as nitrate. This interaction is labeled negative adsorption. Since cations have a preferred attraction to these mineral surfaces, anions are immediately excluded from the area. Transport of anions through soil is directly affected by this process (Jury et al, 1991). Nitrates can be transported by surface runoff, lateral flow, or percolation. In order to estimate the amount of nitrate lost from the soil layer and moved with water, the concentration of nitrate and the amount of mobile water in the layer is calculated. The concentration of nitrate in the mobile water is calculated using

$$conc_{NO3,mobile} = \frac{NO3_{ly} \cdot \left(1 - exp\left[\frac{-w_{mobile}}{(1-\theta_e) \cdot SAT_{ly}}\right]\right)}{w_{mobile}} \tag{3.6}$$

where $conc_{NO3,mobile}$ is the concentration of nitrate in the mobile water for a given layer (*kg N/mm* H_2O), $NO3_{ly}$ is the amount of nitrate in the layer (*kg N/ha*), w_{mobile} is the amount of mobile water in the layer (*mm* H_2O), θ_e is the fraction of porosity from which anions are excluded, and SAT_{ly} is the saturated water content of the soil layer (*mm* H_2O). Based on the depth of the soil layer, the amount of mobile water is

$$w_{mobile} = Q_{surf} + Q_{lat,ly} + w_{perc,ly} \qquad \text{for top 10mm} \qquad (3.7)$$

$$w_{mobile} = Q_{lat,ly} + w_{perc,ly} \qquad \text{for lower soil layers} \qquad (3.8)$$

where w_{mobile} is the amount of mobile water in the layer (*mm* H_2O), Q_{surf} is the surface runoff generated on a given day (*mm* H_2OI), $Q_{lat,ly}$ is the water discharged from the layer by lateral flow (*mm* H_2O), and $w_{perc,ly}$ is the amount of water percolating to the underlying soil layer on a given day (*mm* H_2O). Nutrients from the top 10mm of soil are allowed to interact with and be transported by surface runoff.

Nitrate removed by surface runoff is given as

$$NO3_{surf} = \beta_{NO3} \text{ x } conc_{NO3,omobile} \text{ x } Q_{surf} \qquad (3.9)$$

where $NO3_{surf}$ is the nitrate removed in surface runoff (*kg N/ha*), β_{NO3} is the nitrate percolation coefficient, $conc_{NO3,mobile}$ is the concentration fo nitrate in the mobile water for the top 10 mm of soil (*kg N/mm* H_2O), and Q_{surf} is the surface runoff generated on a given day (*mm* H_2O).

Nitrate removed in lateral flow is given as

$$NO3_{lat,ly} = \beta_{NO3} \text{ x } conc_{NO3,omobile} \text{ x } Q_{lat,ly} \qquad \text{for top 10mm} \qquad (3.10)$$

$$NO3_{lat,ly} = conc_{NO3,omobile} \text{ x } Q_{lat,ly} \qquad \text{for lower soil layers} \qquad (3.11)$$

where $NO3_{lat,ly}$ is the nitrate removed in lateral flow from a layer (*kg N/ha*), β_{NO3} is the nitrate percolation coefficient, $conc_{NO3,omobile}$ is the concentration of nitrate in the mobile water for the layer (*kg N/mm* H_2O), and $Q_{lat,ly}$ is the water discharged from the layer by lateral flow (*mm* H_2O).

Nitrate moved to the underlying layer by percolation is given as

$$NO3_{perc,ly} = conc_{NO3,omobile} \text{ x } w_{perc,ly} \tag{3.12}$$

where $NO3_{perc,ly}$ is the nitrate moved to the underlying layer by percolation (*kg N*/ha), $conc_{NO3,omobile}$ is the concentration of nitrate in the mobile water for the layer (*kg N/mm* H2O), and $w_{perc,ly}$ is the amount of water percolating to the underlying soil layer on a given day (*mm* H_2O).

Another form of nitrogen, organic nitrogen, may also be transported by surface runoff. Organic N attaches to soil particles more readily and, thus, this form of nitrogen loading is represented as a function of sediment loading from the HRU. To estimate the amount of organic nitrogen transported with the sediment to the stream network a loading function developed by McElroy et al. (1976) and later modified by Williams and Hann (1978) is used and given as

$$orgN_{surf} = 0.001 \cdot conc_{orgN} \cdot \frac{sed}{area_{hru}} \cdot \varepsilon_{N:sed} \tag{3.13}$$

where $orgN_{surf}$ is the amount of organic nitrogen transported to the main channel in surface runoff (*kg N/ha*),, *sed* is the sediment yield on a given day (*metric tons*), $area_{hru}$ is the HRU area (*ha*), $\varepsilon_{N:sed}$ is the nitrogen enrichment ration, and $conc_{orgN}$ is the concentration of organic nitrogen

in the top 10mm (*g N/metric ton soil*) and can be calculated by

$$conc_{orgN} = \frac{(orgN_{frsh,surf} + orgN_{sta,surf} + orgN_{act,surf}}{\rho_b \cdot depth_{surf}} \tag{3.14}$$

where $orgN_{frsh,surf}$ is nitrogen in the fresh organic pool in the top 10mm (*kg N/ha*), $orgN_{sta,surf}$ is nitrogen in the stable organic pool (*kg N/ha*), $orgN_{act,surf}$ is nitrogen in the active organic pool in the top 10mm (*kg N/ha*), ρ_b is the bulk density of the first soil layer (*Mg/m*^{3}), and $depth_{surf}$ is the depth of the soil surface layer (*10mm*).

The enrichment ration is defined as the ratio of the concentration of organic nitrogen transported with the sediment to the concentration in the soil surface layer. SWAT utilizes a logarithmic relationship presented by Menzel (1980) to determine the enrichment ratio. The equation used to determine $\varepsilon_{N:sed}$ is

$$\varepsilon_{N:sed} = 0.78 \text{ x } (conc_{sed,surq})^{-0.2468} \tag{3.15}$$

where $conc_{sed,surq}$ is the concentration of sediment in surface runoff ($Mg\ sed/m^3\ H_2O$). To calculate the concentration of sediment in surface runoff, the following is used:

$$conc_{sed,surq} = \frac{sed}{10 \cdot area_{hru} \cdot Q_{surf}} \tag{3.16}$$

where *sed* is the sediment yield on a given day (*metric tons*), $area_{hru}$ is the HRU area (*ha*), and Q_{surf} is the amount of surface runoff on a given day (*mm* H_2O).

The movement of phosphorous in soil primarily occurs by diffusion. Soluble phosphorous has low mobility and the amount transported in surface runoff is calculated by:

$$P_{surf} = \frac{P_{solution,surf} \cdot Q_{surf}}{\rho_b \cdot depth_{surf} \cdot k_{d,surf}} \tag{3.17}$$

where P_{surf} is the amount of soluble phosphorus lost in surface runoff (*kg P/ha*), $P_{soulution,surf}$ is the amount of phosphorus in solution in the top 10 mm (*kg P/ha*), Q_{surf} is the amount of surface runoff on a given day (*mm* H_2O), ρ_b is the bulk density of the top 10mm (Mg/m^3), $depth_{surf}$ is the depth of the surface layer (*10mm*), and $k_{d,surf}$ is the phosphorus soil partitioning coefficient (m^3/Mg), which is the ratio of soluble phosphorus concentration within the first 10 mm of soil to the concentration in the surface runoff (Neitsch, 2011).

Additional forms of phosphorus, organic and mineral, that are attached to soil particles may also be transported by surface runoff to streams. Consequently, they are also associated with sediment loading from the HRU. The loading function to calculate phosphorus transported with sediment is the same as the function for determining organic nitrogen. The concentration of phosphorous attached to sediment is calculated using the concentration equation for organic nitrogen, and the phosphorus enrichment ratio is calculated the same as well.

3.1.5 Plant Growth

An important factor to plant growth is temperature. All plants have an ideal range of temperatures they prefer and can withstand. SWAT models plant growth based on this concept.

In order for plant growth to take place, a base temperature must be met. At levels above this base temperature, plant growth is more rapid. However, only up to an optimum temperature is met, beyond this point plant growth slows until eventually ceasing at a maximum temperature. This concept of heat unit theory states that plants have temperature requirements related to their life cycle that can be calculated and quantified. A heat unit represents the portion of average daily temperature that is above the plant's base temperature, which in turn contributes to the plant's growth. Plant growth will only occur on the days when the average daily temperature is higher than the base temperature. In order to calculate the heat unit accumulation for a particular day, the following is used:

$$HU = \overline{T}_{\mathrm{av}} - T_{base} \qquad \text{when } \overline{T}_{av} > T_{base} \tag{3.18}$$

where HU is the number of heat units accumulated on a given day (*heat units*), $\overline{T}_{\mathrm{av}}$ is the mean daily temperature (°*C*), and T_{base} is the plant's base or minimum temperature for growth (°*C*). In order to calculate the total number of heat units a plant requires to reach maturity the following is used:

$$PHU = \sum_{d=1}^{m} HU \tag{3.19}$$

where PHU is the total heat units, or potential heat units, required for plant maturity (*heat units*), HU is the number of heat units accumulated on day d where $d = 1$ on the day of planting and m is the number of days required for a plant to reach maturity. When using the above

equation, it is necessary to know the number of days for a particular crop to reach maturity. Luckily, these values have already been quantified for most crops and are easily accessible (Neitsch, 2011).

The heat unit theory's reliability as a predictor of harvest dates, has led to its adaptation for the prediction of other plant development stages (Cross and Zuber, 1972). The success of this adaption has also led to the use of heat units to schedule other crop management operations. SWAT allows users to either schedule management operations by specific dates or by fractions of potential heat units. When using heat units for scheduling, SWAT schedules operations based on temperature. This method of timing is useful in large watersheds where climatic differences may occur throughout the watershed and have an effect on the timing of operations, or in areas where climate varies from year to year.

3.1.6 Management Operations

SWAT allows users to specify various agricultural management operations and water management practices to be simulated in order to predict the impacts of land use and land management. SWAT's agricultural management operations include planting and harvesting operations, harvesting and killing operations, grazing operations, tillage operations, and fertilizer and pesticide operations.

The planting operation identifies the initiation of plant growth, and requires specific information about the plant type, its planting date, and the total number of heat units required to reach maturity. A harvest operation designates the removal of plant biomass from an HRU without killing the plant. This operation is commonly used when harvesting grasses like switchgrass or *miscanthus*. This operation only requires that the user select a harvesting date. Ultimately, plant growth can also be completely terminated by using a harvest and kill operation.

Based upon the harvest index, a fraction of the plants' biomass is removed and the remaining is converted to residue on the soil surface. Again the only information needed for this operation is the date of harvest (Neitsch, 2011). Application of fertilizers and pesticides are also modeled in SWAT and can be scheduled along with the amount used according to crop requirements. The combinations of these agricultural management operations, which control plant growth and harvest yield, are used to determine the environmental impacts that may arise when increasing the cultivation of biofuel feedstocks.

3.1.7 Crop Yield

SWAT also simulates crop yield from an HRU, which is the portion of plant material that is accumulated above ground and collected on the day of harvest. When harvested, the plant material is removed entirely from the system along with the nutrients contained in the yield and will not get added to future residual and organic nutrient pools in the soil. SWAT uses a harvest index to identify the amount of above ground plant biomass that is removed as dry yield. For most crops, this value will range between 0.0 and 1.0, with the exception of plants whose roots are also harvested and have a harvest index greater than 1.0.

SWAT determines the harvest index daily for plants during the growing season using the following equation:

$$HI = HI_{opt} \cdot \frac{100 \cdot fr_{PHU}}{(100 \cdot fr_{PHU} + \exp[11.1 - 10 \cdot fr_{PHU}])} \tag{3.20}$$

where HI is the potential harvest index for a given day, HI_{opt} is the potential harvest index for the plant at maturity given ideal growing conditions, and fr_{PHU} is the fraction of potential heat units accumulated for the plant on a given day in the growing season (Neitsch, 2011).

3.1.8 Energy and Temperature

In order to accurately simulate water balance, accurate energy inputs are necessary. When water is introduced to a system in the form of precipitation, available energy in the form of solar radiation is a major driving force in movement of water on land. Temperature and solar radiation effect snow fall and melt, and evaporation.

SWAT incorporates various sun-earth relationship concepts to make solar radiation calculations. Distance between the sun and earth during its elliptical orbit, solar declination, solar noon, sunrise, sunset, and day length, and solar radiation are calculated in the model.

Additionally, temperature affects many physical, chemical, and biological processes, especially plant production. Daily air temperature values can be manually input to SWAT from observed records or input with the built in generator. These daily air temperature values are used to calculate soil and water temperatures. Daily maximum and minimum air temperatures are required by SWAT to model processes, and it is recommended that accurate observed data within watershed be used in order to improve output results. However, if air temperature data is not available, SWAT has a built in generator that is capable of generating daily air temperature values (Neitsch, 2011).

3.1.9 Weather Generator

SWAT requires various input data related to weather that can be read from observed data files or generated using the built in weather generator. Precipitation, maximum and minimum

daily temperatures, solar radiation, humidity, and wind speed, can be generated using SWAT's WXGEN weather generator model (Sharpley and Williams, 1990), which was developed for the contiguous United States. The weather generator incorporates multiple procedures for calculating precipitation, solar radiation and temperature, relative humidity, and wind speed. To generate precipitation, SWAT utilizes a Markov chain model to determine wet and dry days and multiple probability distribution functions to calculate precipitation amounts. It also includes a stationary generating process to calculate residuals for temperatures and solar radiation using serial correlations and correlation coefficients. Relative humidity is required by SWAT to determine evapotranspiration rates and vapor pressure. The weather generator calculates daily relative humidity using a triangular distribution method and wind speed using a modified exponential equation. Full details on how the WXGEN weather generator model functions can be found in SWAT's Theoretical Documentation (Neitsch, 2011).

Chapter 4: Optimization

In mathematics and computer science, the process of optimization consists of constructing systems and decisions as most effective and functional as possible subject to available resources and limiting constraints. However, in complex systems and designs, it is not practical and sometimes not possible to explore all possible designs, so optimization techniques are required. The purpose of optimization is to determine values for parameters that aim to minimize or maximize objective functions, while taking into consideration specific constraints. Feasible solutions consist of parameter values that meet all constraints, and feasible solutions with the best objective values are considered optimal solutions. Mathematically, a minimization problem can be expressed as

$$\text{Minimize } \mathrm{F} = f(X_t, U_t, t) \tag{4.1}$$

$$\text{Subject to } \; X_t \leq X_t \leq X_t \tag{4.2}$$

$$U_t \leq \; U_t \leq U_t \tag{4.3}$$

$$W(X_t, U_t) = 0 \tag{4.4}$$

where F is the objective function, X is the independent variable, U is the dependent variable, and the feasible solution space is represented and satisfies the *subject to* constraints. A variety of optimization methods can be used to obtain an optimal solution for this problem. The selection of optimization methods should consider the type and number of objective functions in the problem. Optimization methods range from more traditional methods like exhaustive search to sophisticated heuristic search methods like evolutionary algorithms.

4.1 Optimization Methods

Of the traditional optimization methods, one of the simplest techniques is exhaustive search. Exhaustive search, also known as brute-force search, is a very general algorithm that assesses all possible solutions with in the search space until a global optimum solution is determined. Although it is a simple method to implement, and an optimal solution is guaranteed, it can involve very large numbers of candidate solutions to evaluate, which requires considerable computational demand. The exhaustive search method can be modified to sample the search space systematically in order to decrease the amount of candidate solutions. However, there is a chance that the global optimum might be overlooked using this refinement, due to under sampling. Since many real world problems involve identifying a solution from extensive search spaces, exhaustive search is not usually an appropriate optimization method.

Another general approach is gradient-based optimization methods. Gradient based optimization methods utilize fundamentals of calculus to find minimums and maximums of objective functions. Gradient based methods can be grouped into two main types, direct and indirect. Direct methods search for local optima by starting with an arbitrary solution and move in a direction that is determined by the local gradient of the objective function. This method is known as hill climbing and can be seen applied in linear programming problems. On the other hand, indirect methods search for local optima by setting the gradient equal to zero and then solving the nonlinear equations (Pike, 1986). Gradient-based methods are limited in the fact that they are only local in scope and that the objective functions must have derivatives. These limitations hinder the practical application of these methods for solving many real world optimization problems.

On the other hand, metaheuristic search methods, which include evolutionary algorithms, can solve more complex problems. Evolutionary algorithms are more efficient when applied to solving problems that have discontinuous, non-differentiable, multi-modal, and noisy response surfaces. They can even be applied in situations that have dynamic objectives and constraints that are changing with time (Schwefel, 2000). Evolutionary algorithms are applied in the proposed research to solve the feedstock production and land use change problem. Considering the fact that the indicators of environmental impact are related to physical processes occurring in the watershed and are complex and dynamic, and non-linear in nature makes evolutionary algorithms a fitting method for solving this problem.

4.2 Multi-Objective Optimization

Optimization problems that involve multiple objectives introduce multiple potential solutions that compromise between competing objectives, compared to a single optimal solution. It is possible to apply single objective optimization methods to these problems by combining multiple objectives. This can be done one of two ways; either by optimizing only one objective while using the others as constraints or by applying weights to each objective. However, the assignment of weights is very subjective, and in order to accurately identify trade offs among different objectives, multiple model runs using different weighting schemes are required. When using a single objective approach, crucial details about trade off characteristics may be lost and the search space may be limited (Singh and Frevert, 2005). Consequently, multi-objective optimization methods are far more suitable and increasing in popularity for the evaluation of multi objective tradeoffs. By using the concept of Pareto dominance and optimality, these algorithms apply a population-based approach to identify multiple Pareto optimal solutions in a single model run (Deb and Beyer, 2001). This set of solutions is referred to as the Pareto optimal

front. A solution is determined to be Pareto optimal, or nondominated, if none of the other competed objectives can be improved without reducing some of the other objective values. All Pareto optimal solutions are equally good solutions, and can be used to quantify the tradeoffs of competing objectives. General multi-objective optimization problems and the concept of Pareto dominance can be expressed as follows

$$\text{Minimize } f(x) = (f_1(x), f_2(x), \ldots, f_n(x)) \tag{4.5}$$

$$\text{subject to: } g(x) = (g_1(x), g_2(x), \ldots, g_n(x)) \leq 0 \tag{4.6}$$

$$x \in D \subseteq x \tag{4.7}$$

where $f(x)$ represents the objective functions, and $g(x)$ represents the constraints, D is the feasible region of solutions, and X is the decision variable space. Considering two decision variables a and b, a is said to dominate b

$$\textit{iff } \forall i \in \{1,2,\ldots,n\}: f_i(a) \leq f_i(b) \text{ and } \exists i \in \{1,2,\ldots,n\}: f_i(a) < f_i(b) \tag{4.8}$$

Simply put, decision variable a dominates decision variable b if and only if it performs no worse than b in all n objectives and strictly better than b in at least one objective.

Several methodologies have been developed to solve multi-objective optimization problems. These methods fall into one of two main categories; mathematical programming methods and meta-heuristic methods.

4.3 Mathematical Methods

4.3.1 Goal Programming

Developed by Charnes et al, goal programming is a linear programming optimization based method that was designed specifically to handle multi-objective optimization problems (1955). In order to optimize multiple objectives, an aspiration level, or contribution coefficient, is assigned to each one. Goal programming works by minimizing the deviations from the prescribed aspiration levels of each objective, or goal. Consequently, the objective function is the sum of all the contribution coefficient and the constraints are simply the objectives. The simplicity of formulation and application makes goal programming one of the most widely used methods for multiple objective optimization. Unfortunately, the objectives are not simultaneously optimized and in more complex problems a feasible solution may not be obtained.

4.3.2 Weighted Sum Method

The weighted sum model, also referred to as multi-criteria decision making method, is based on the goal programming method and is capable of evaluating multiple alternatives in terms of multiple decision criteria (Fishburn, 1967). The decision maker assigns each objective function a different weight, or importance, and the set of alternatives are evaluated. The total importance of each alternative is compared with the other alternatives until a maximum or minimum sum of the weights is identified. This method is useful when an objective is more important that the others. However, this leads to having similar disadvantages as goal programming in that the assigning of weights is highly subjective and ambiguous. Multiple authors have addressed these concerns. For example, theoretical selection of weights does not always result in an actual optimal solution and

the decision maker may need to reevaluate the weights used (Messac, 1996). Additionally, there is difficulty in establishing an even distribution of Pareto optimal points that accurately and completely represent the Pareto front (Das and Dennis 1997).

4.3.3 Lexicographic Method

Proposed by Fishburn (1974) and similar to preceding methods, the lexicographic method functions by assigning importance to the objective functions that are being optimized. After being arranged in order of importance they are individually solved one by one. After the first, or most important objective is optimized, the second objective is optimized subject to the initial constraints along with a new constraint that ensures the previous objective remains optimal, and continuing until all objectives have been optimized. Although easy to implement, again there is ambiguity when selecting the order of importance. Like the goal programming method, it also does not optimize the objectives simultaneously. Additionally, there is a tendency for the process to stop prematurely before the less important objectives are optimized, due to lack of feasible solutions. Some modifications that have been proposed to counter these drawbacks include the relaxation and variation of constraints and the initial objective function importance values (Osyczka, 1984). Even with its disadvantages, the lexicographic method still provides Pareto optimal solutions efficiently and is relatively simple to use.

4.4.4 Utility Theory

Goal programming methods and multiple objective mathematical programming utilize Utility function as a base for optimization. Utility function is a measure of how valuable, or desirable a possible choice is, and is an appropriate method for solving multiple objective problems when a specific value function is known. Multi-attribute utility theory models the

decision maker's preferences and is expressed by a utility or value function. The value function outlines the ordering of the search space, and generates a Pareto optimal solution. The real challenge arises when trying to determine the value function. Nevertheless, if the decision maker is able to determine the value function, the method can be easily applied and effective in solving multiple criteria optimization problems (Georgy et al, 2005).

4.4 Meta-heuristic Methods

The aforementioned mathematical methods have been used to solve multiple objective optimization problems but not without considerable drawbacks. When dealing with larger complex problems that have significantly larger search spaces, it becomes impractical to apply mathematical methods. Alternatively, meta-heuristic methods are becoming more popular and overall more useful approaches for solving multiple objective problems. Meta-heuristic methods do not guarantee a global optimal solution, instead they use stochastic optimization to search large feasible solution spaces and arrive at estimated optimal solutions. Meta-heuristics find quality solutions that are good approximations to the true Pareto frontier and demand less computational effort.

4.4.1 Tabu-Search

First presented by Glover (1986), Tabu-search uses local search methods to find optimal solutions. Local search methods work by evaluating the immediate neighbors (similar solutions with only minor changes) of a potential solution hoping to find a better solution. Tabu-search strengthens the local search method by incorporating a memory system that keeps information about the characteristics of the previous solution visited. The information in the memory structures

guides the movement from one solution to another, and forms a 'tabu list'. This list is a set of rules and forbidden solutions that determine which solutions are acceptable to the search space. The versatility of this memory system creates an effective search path for multi-objective optimization problems (Hertz et al, 1995). Tabu-search has been applied to many different problems and has proven to be an efficient meta-heuristic technique.

4.4.2 Simulated Annealing

Simulated annealing is a metaheuritstic used for the global optimization of multiple objectives in a large search space. The name is derived from annealing in metallurgy. The algorithm simulates this physical process by slowly decreasing the probability of accepting inferior solutions when exploring the search space. During each iteration of the simulated annealing process, the current solution is randomly replaced with a nearby solution, which are determined by designated acceptance probability functions. Like annealing, these probabilities lead the algorithm to states of lower energy with each iteration, and are repeated until reaching a user defined termination point. The simulated annealing method has its advantages and disadvantages. The method effectively handles nonlinear, chaotic, and noisy models with multiple constraints and is easy to implement. It has the capability to elude local optimums and ultimately approach global optimal solutions. However, it requires a higher computational demand to find higher quality solutions, due to the random search. Hence, the optimal solution is not always guaranteed.

4.4.3 Swarm Intelligence

Swarm intelligence is a class of optimization methods that are inspired by natural biological systems and model the behavior of ant colonies and bird flocking. Swarm intelligence systems are made up of agents that interact locally together and with the environment. They follow simple rules and behave somewhat random and local, leading to an intelligent-type of global behavior. Ant colony optimization was originally proposed by Dorigo et al. (1996), and is based on the actions of ants. 'Ants', or simulation agents, explore solutions to locate optimal solutions by traveling through a parameter space that represents all possible solutions. These 'ants' record their positions and the quality of their solutions in a memory system, which aids in selecting future paths of better quality. Particle swarm optimization is a global optimization method that simulates the predictable movements of a bird flock. Developed by Eberhart and Kennedy (1995), particle swarm optimization can effectively be applied to nonlinear functions where a point in an n-dimensional space can represent the best solution. Particle swarm optimization works by first placing a number of simple entities, 'particles', in the search place to move throughout the space by using a combination of their own best solution histories and those of one or more members of the swarm. After multiple iterations of this process, eventually the swarm will move closer and closer to an optimum of the fitness function (Poli et al. 2007). Other examples of swarm intelligence include glowworm swarm optimization, artificial immune systems, and stochastic diffusion search.

4.4.4 Evolutionary Algorithms

For the multi-objective optimization carried out in this research, a genetic evolutionary algorithm will be used. Evolutionary algorithms are population-based metahueristic optimization techniques that are based on the process of evolution and are suitable for complex,

nonlinear problems (Haupt, R. and Haupt, S, 1998). These algorithms utilize populations of randomly generated initial possible solutions and iteratively apply operations of evaluation, selection, crossover, and mutation to evolve the population to optimal or near optimal solutions.

Throughout the years, various types of evolutionary algorithms have emerged and applied to solve different optimization problems within various disciplines. However, the following underlying mechanisms remain the same for all evolutionary algorithms: decision alternatives are represented as a population of individuals, the solution space is explored in two ways; random mutation and recombination of individuals, and they all assign a fitness measure for each individual, which allows potential solutions to be compared with others. They differ in the methods of evaluate fitness and reproduce. Some utilize mutation, others use a combination of elitism and crossover, and others use tournament selection when selecting the best individuals for a new population.

Overall, the basic process of an evolutionary algorithm is as follows:

1) The initial population is generated, usually randomly. The size of this population is assigned by the decision maker and depends on the nature of the problem. The fitness of each individual is then evaluated.
2) A proportion of the population is selected and undergoes reproduction to generate the new population. An individual solution from the population can be selected randomly or in most cases is selected based upon its fitness as a best solution.
3) The generational process continues until stopping criteria has been reached. Examples of terminating conditions are: fixed number of generations or computation time, minimum criteria for solution is met, or the fitness values have reached a plateau.

Evolutionary algorithms begin by initializing, or generating, the first population of individuals. These individuals represent proposed solutions for the optimization problem and can be represented by chromosomes. A chromosome, usually a simple string, represents a set of parameters that define a potential solution. A chromosome is further subdivided into genes that contain important information about the solution. Chromosomes can be encoded in various ways; binary coded, permutation encoding, and value coding. One of the simplest and widely used methods for coding chromosomes is binary coded. Binary encoding simply uses 0's and 1's to represent active or disconnected components in each gene. Another method of encoding is permutation encoding, where integers are used in each gene to represent a position in a sequence or a number of components in a particular position. Finally, value encoding creates chromosomes as a string of different values. Values can be anything connected to the problem like numbers, characters, or even more complicated objects.

One of the most important points in the evolutionary process is the selection of the better individuals that will create the next generation. A fitness function is generally used to assign fitness value to each individual and then a selection strategy is implemented. Some of the more popular strategies are roulette wheel selection, tournament selection, and rank selection. In roulette wheel selection, solutions are grouped in relation to their fitness values, a probability of selection is associated with each individual, and a proportion of the wheel is assigned to each solution. Potential solutions that have a higher fitness value will be less likely to be eliminated, but still have a chance of elimination. Another method is tournament selection, which involves multiple competitions of randomly chosen individuals. The individual with the best fitness wins the tournament and is selected for crossover. The size of the tournament can be adjusted to create different selection pressures, allowing for either more or less weak individuals to have the

chance of selection. Lastly, rank selection simply assigns a rank to each individual based on their fitness and those with better ranks are selected.

Another equally important operator used in evolutionary algorithms is crossover. Crossover is used to vary the population after each iteration, creating successive generations with better individuals. Simulating reproduction, crossover takes two parent solutions and produces a child using a combination of their genes. There are a multiple of crossover techniques available and each have their efficiencies at creating a diversity of quality children and effectively exploring the entire search space. Two of the most well-known crossover strategies are single point crossover and double point crossover. Single crossover functions by randomly selecting a crossover point on the chromosome of both parents and all data beyond that point is traded between the two, resulting in two children. Similarly, double point crossover randomly selects two points on the chromosome and everything in between these two points is exchanged between the two parents, also resulting in two children.

4.5 Multi-objective Evolutionary Algorithms

Evolutionary algorithms can be applied effectively to both single and multi-objective optimization problems. Examples of evolutionary algorithms that have been specifically designed for multiple objective optimization problems include: Strength Pareto Evolutionary Algorithm (SPEA) by Zitzler and Theile (1999), Strength Pareto Evolutionary Algorithm 2 (SPEA 2) by Zitzler et al (2001), Pareto Archived Evolution Strategy (PAES) by Knowles and Corne (2000), and Non-dominated Sorting Genetic Algorithm (NSGA II) by Deb et al (2002). The next section will briefly describe each of these.

4.5.1 Strength Pareto Evolutionary Algorithms (SPEA and SPEA2)

Developed by Zitzler and Thiele (1999), SPEA is a multi-objective optimization algorithm that is based on the concept of Pareto dominance for selection and evolution of potential solutions. It also utilizes the concept of elitism. Elitism ensures that the best, nondominated solutions are stored in an external archive for each population at every generation. At each generation the current population and the stored population are combined and fitness is evaluated by calculating the number of solutions each individual dominates. After fitness is evaluated, all nondominated solutions from the population and from the archived population are used for the next population. SPEA2 improved on SPEA by modifying the fitness function. SPEA2 incorporates a fine-grained fitness assignment, a density estimation, and improved archiving methods (Ziztler, 2001).

4.5.2 Pareto Archived Evolutionary Strategy (PAES)

Developed by Knowles and Corne (2000), PAES is a modified MOEA. The main difference of PAES from other MOEAs is the crossover operator. A parent undergoes mutation to create an offspring, and if this offspring dominates the parent then it is accepted as a new parent and continues the process. However, if it is dominated by the parent, it will be discarded and mutation is used again to generate a new child. The nondominated solutions are stored and new offspring are compared to this archive to ensure diversity. If a new individual dominates any member of the archive then that solution is eliminated and the new offspring replaces it. The authors claim that PAES is possibly one of the simplest algorithms capable of producing a diverse set of solutions that contribute to the Pareto optimal set.

4.5.3 Non-dominated Sorting Genetic Algorithm (NSGA-II)

Srinivas and Deb (1994) proposed a modified version of NSGA in order to address weaknesses of the previous version. It is population-based algorithm that begins with a randomly generated parent population of solutions. NSGA incorporates tournament selection, recombination, and mutation when creating the offspring population. Additionally, elitism is used to combine current and offspring populations for the successive generation. This population is then sorted using a non-dominating sorting approach and incorporates a crowded-comparison operator to allow diversity while also reducing population size.

Chapter 5: Methodology

Objective 1: Develop an optimal control model by integrating a hydrologic and environmental model with a multiple objective evolutionary algorithm that can search and identify optimal landscape scenarios for feedstock production.

The SWAT model provides an excellent foundation for the integration of a modeling tool and optimization techniques to provide evaluation of biofuel feedstock placement and management practices. The tool can evaluate ecosystem services as impacted by land use change and the increase in feedstock cultivation.

The methodology of this research is based on an integrated modeling approach in which a genetic algorithm is interfaced with a hydrologic model. In order to meet the primary objective, an optimization method was developed based on evolutionary algorithms to evaluate different land cover change scenarios and their effect at the watershed level by coupling a multi-objective genetic algorithm with the Soil and Water Assessment Tool (SWAT) to model these land use changes.

The integration of the two components was developed in MATLAB. MATLAB, developed by MathWorks, is a multi-paradigm numerical computing environment that uses a fourth-generation programming language.

Objective 2: Quantify and identify the environmental impacts and tradeoffs of land use changes and management practices on water quality and ecosystem services when converting traditional land uses to biofuel feedstock production.

To evaluate the impacts of different agricultural practices on soil and water quality at watershed scales, SWAT has been used effectively and efficiently. Models that have been calibrated and validated with measured runoff, along with sediment and nutrient losses from watersheds have been used to assist decision makers in the selection of conservation practices and allocation of resources (Singh and Frevert, 2005). By linking a multiobjective genetic algorithm with SWAT, many possible agricultural landscapes can be evaluated to find optimal solutions that minimize nutrient loading, and sediment yield due to erosion at the watershed outlet, while also minimizing the effects on regional water resources and maximizing crop yields.

SWAT, version 2012, was utilized to quantify the watershed scale impacts of biofuel feedstock production and to optimize selection and placement of biofuel feedstock crops. The ArcSWAT interface for SWAT was used to setup watershed parameters and create model input files. The interface includes a series of tools to delineate watershed boundaries, define subbasins and hydrologic response units (HRUs), and create management files. The Lake Fork and San Juan watershed were divided into subbasins using the automatic delineation tool in ArcSWAT. Digital elevation maps were used to identify the drainage network and define the watershed boundary and drainage morphology (such as area, slope, and length) for the main basins. The subbasins were further sub-dived into hydrologic response units (HRUs) identifying unique combinations of land use and soils within each subbasin.

Objective 3: Optimize selection and location of biofuel feedstock crops that will promote healthier ecosystems and agricultural landscapes, while maximizing biomass production yields.

Optimization techniques are used to make systems and decisions as advantageous as possible given specific constraints and resources. Optimization will determine parameter values for specified independent variables that minimize or maximize the given objective functions. After searching the solution space of values that meet all constraints, feasible solutions with the best objective function values can be considered optimal solutions.

For the multiobjective optimization carried out in this research, a genetic evolutionary algorithm was used. Evolutionary algorithms are based on the process of evolution and are suitable for complex, nonlinear problems. Evolutionary algorithms are population-based search methods that model evolutionary principles such as natural selection, reproduction, and random variation (Haupt, R. and Haupt, S, 1998). These algorithms utilize populations of possible solutions and iteratively apply operations of selection, crossover, and mutation to evolve the population to optimal or near optimal solutions. The genetic algorithm used in this project is a heuristic search algorithm that can solve difficult nonlinear optimization problems. The multi-objective optimization ensures that multiple factors such as nitrogen, phosphorous, and sediment yields, and biomass production are all taken into consideration when searching for optimal solutions.

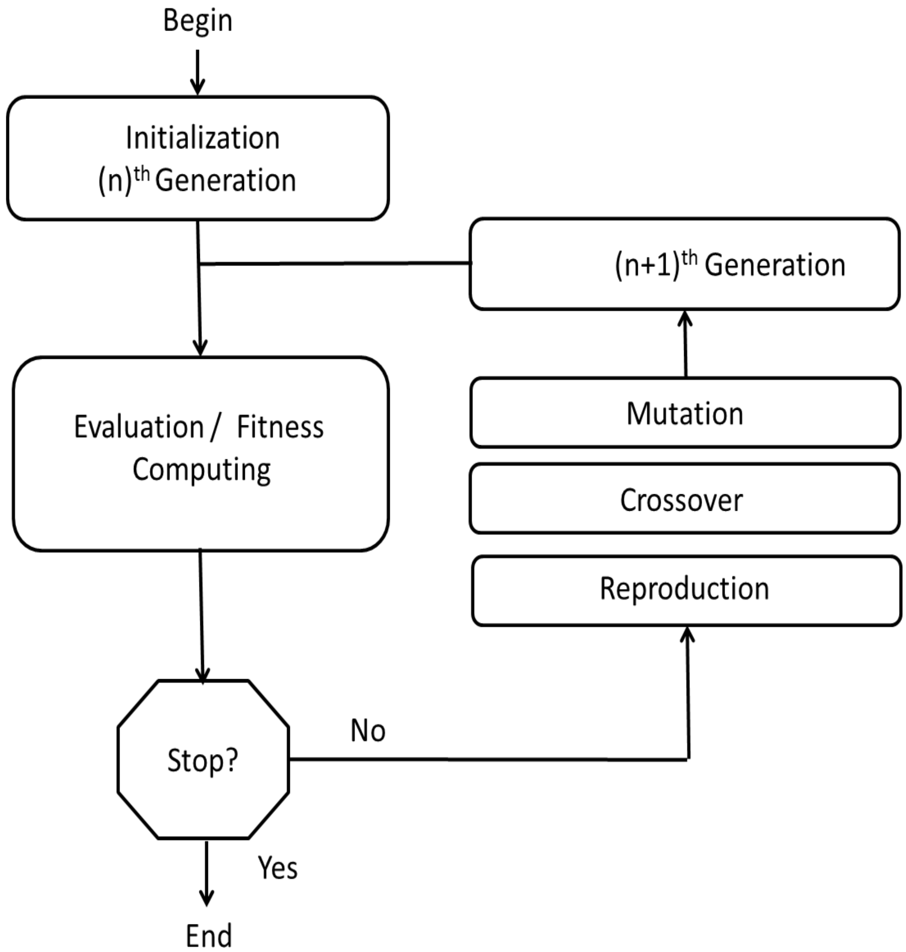

Figure 5.1: Flowchart of Genetic Algorithm

The approach taken in this research required that the SWAT model be interfaced with a multi-objective genetic algorithm. The genetic algorithm will dictate a specific land use/land cover scenario and management practice for SWAT to simulate and evaluate the watershed response. This output information is used by the genetic algorithm to create new scenarios until an optimal or near optimal scenario is identified that will meet the objective functions specified. Figure 5.2 outlines the integrated modeling approach. Application of the integrated modeling system as a

decision support tool is demonstrated using two example watersheds, but can be easily modified for any watershed of interest.

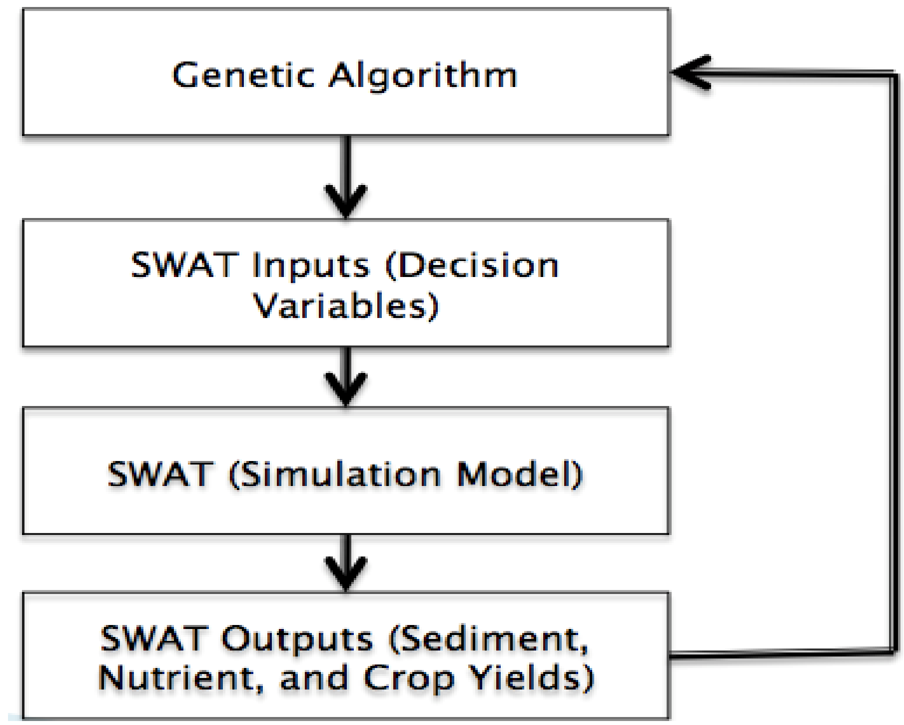

Figure 5.2: Integrated Modeling Approach

A multiobjective optimization problem incorporates multiple conflicting objectives and will ultimately identify a set of Pareto optimal solutions. After the algorithm evaluates many possible scenarios and assesses the fitness of each, a set of Pareto optimal solutions will be generated. Multiobjective evolutionary algorithms work by evolving the population of solution sets and are able to approximate the Pareto optimal set of solutions for the given criteria (Zhou et al, 2011). In this case, our population of solution sets consists of the many possible scenarios of land use and land cover arrangements along with underlying management practices for a specific

watershed. In the area of designing suitable agricultural landscapes that meet environmental objectives lies a multiobjective problem. However, the solution will not be unique, because, in order to satisfy several conflicting objectives there will be a set of multiple potential solutions, where none will be best for all objectives (Blasco et al, 2008). This Pareto front allows us to identify and quantify the trade-offs that occur when cultivating biofuel feedstocks in place of conventional crops.

Chapter 6: Integrated Modeling Approach

An integrated modeling approach involves interfacing a simulation model that evaluates system responses with a search algorithm that is capable of selecting optimal or near-optimal decision alternatives to achieve prescribed goals. The approach has been increasingly popular in various fields of study including large-scale water resources management. In this regard, its application has been demonstrated on the Lake Fork and San Juan watersheds to identify optimal landscape scenarios of biofuel feedstock production.

The integrated modeling approach is particularly advantageous in developing decision support systems that involve coupling of a simulation model and optimization algorithms. The approach does not require further simplification of the problem physics beyond those represented by the simulation model. Furthermore, the complexity of the overall optimization problem is decreased since system dynamics are simulated implicitly to the search algorithm (Nicklow, 2000). In this approach, the simulation model and the search algorithm are separate entities that are loosely connected, allowing easier updating of either the simulation model or the search algorithm to newer versions.

In the context of this research, the development of a computational, optimal control model (OCM) required a hydrologic and environmental simulation model be interfaced with a multi-objective search algorithm. The simulation model evaluates watershed responses resulting from various land management practices each time the search algorithm requires the information and the search algorithm identifies optimal or near-optimal land management practices that can achieve prescribed goals. Below is a flowchart of how the optimal control model functions.

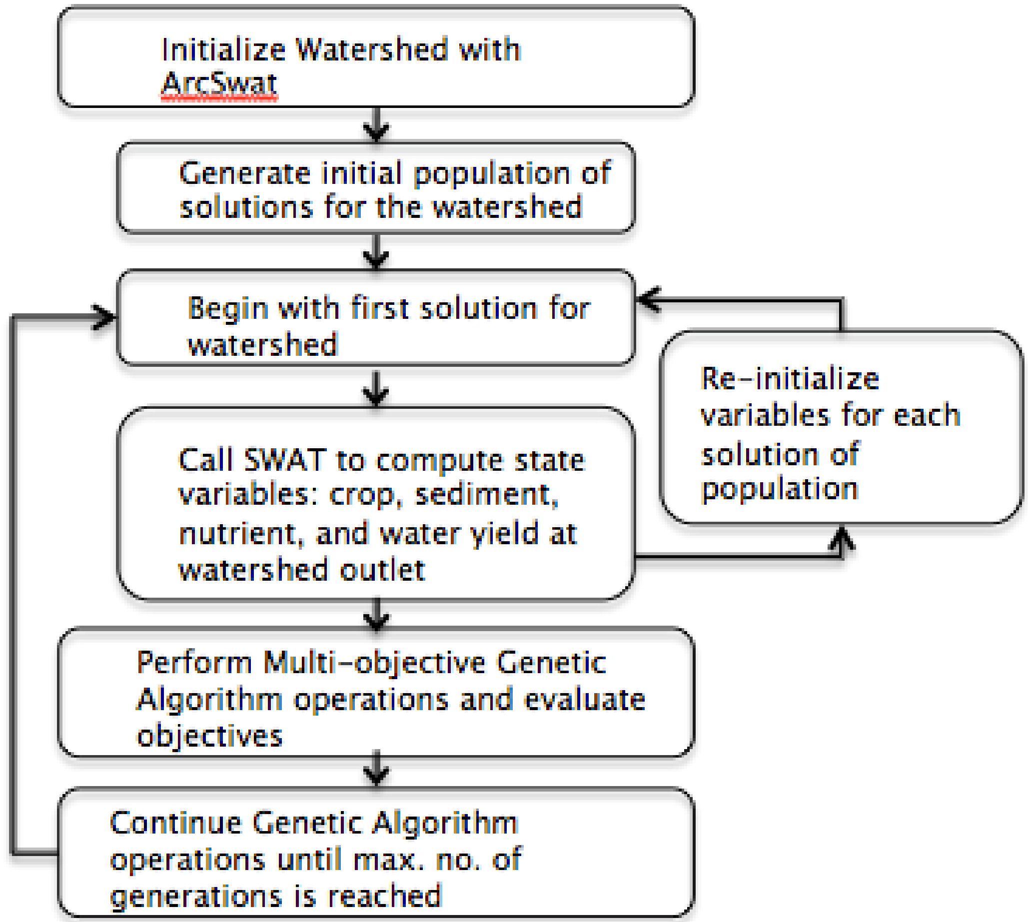

Figure 6.1: Flowchart of Integrated Modeling Approach Using SWAT and MOEA

6.1 SWAT Setup

Prior to running the OCM, SWAT requires initial setup of the watershed. SWAT requires specific information about weather, soil properties, topography, vegetation, and land management practices occurring in the watershed. The physical processes associated with water movement, sediment movement, crop growth, nutrient cycling, are directly modeled by SWAT using this input data.

The ArcSWAT interface was used to initialize watershed parameters and create model input files. The interface includes a series of tools to delineate watershed boundaries, define subbasins, and create management files. Figures 6.2 through 6.5 represent the shapefiles required by the interface to establish parameters and create the necessary data files for the simulation runs. The interface will overlay these shape files to delineate the watershed into subbasins and create a stream network.

Hydrologic response units (HRUs) are defined in SWAT model as lumped land areas with unique land use and soil types. While modeling with SWAT, the dominant land use and soils option of HRU distribution is employed. This results in an equal number of subbasins and HRUs, hence, subbasins and HRUs are equivalent spatial units. This option helps to know the exact location of the HRUs.

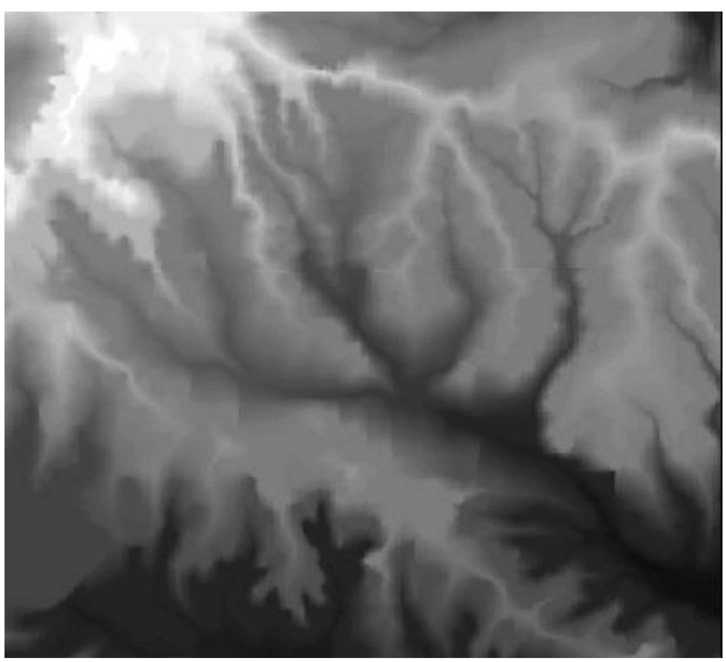

Figure 6.2: Digital Elevation Map

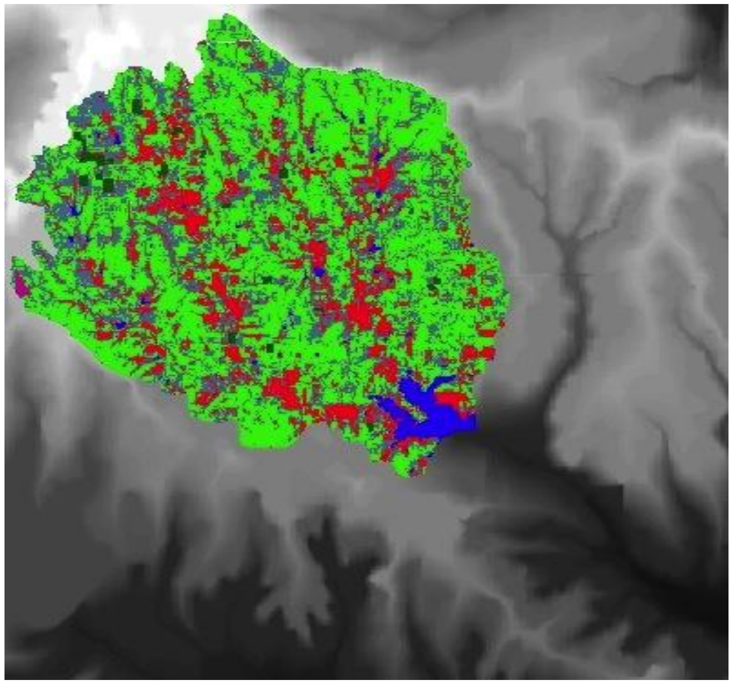

Figure 6.3: Land Use/Land Cover Map

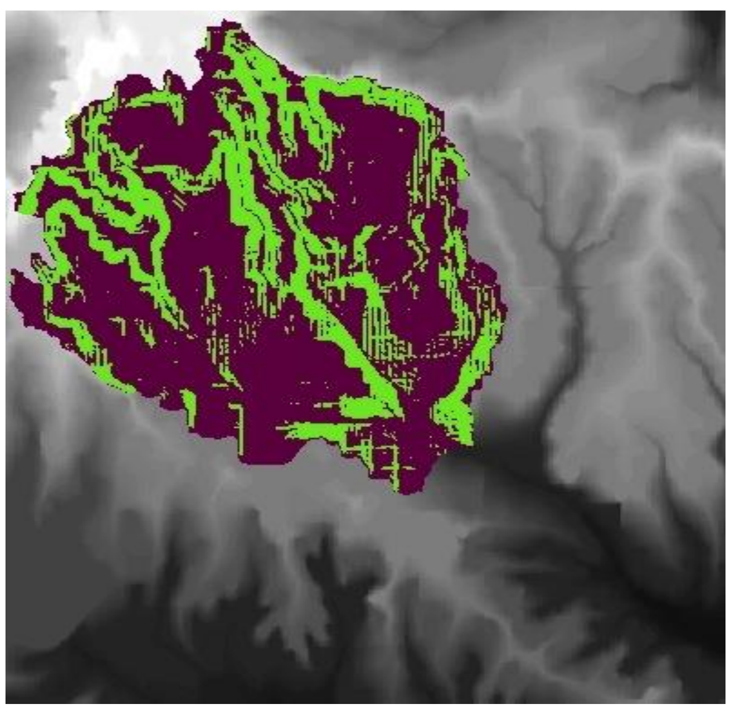

Figure 6.4: Slope

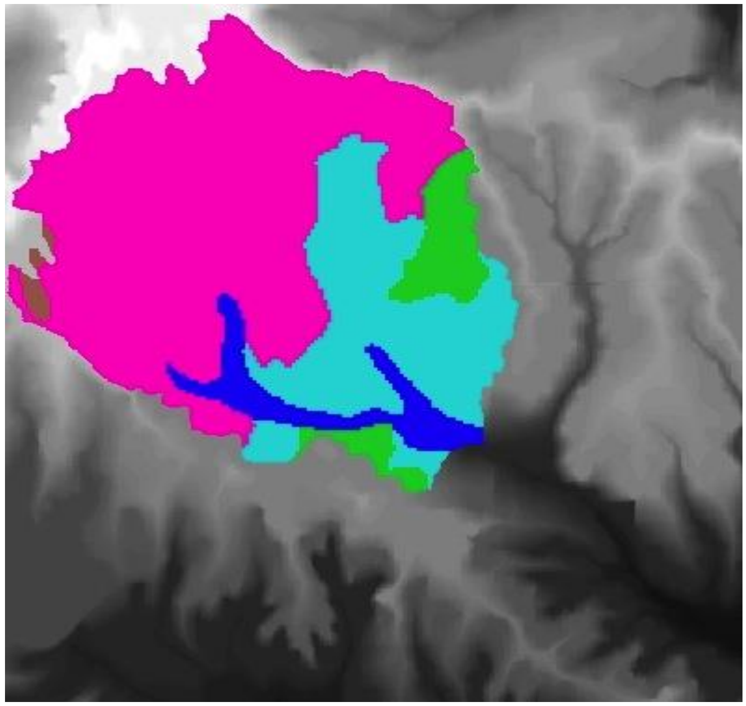

Figure 6.5: Soil Distribution Map

Additionally, SWAT requires that many other parameters be defined in order to run a simulation. These parameters can be manually input, modified, or set to SWAT default values. These parameters include climate data, water flow data, plant growth data, along with management data that includes tillage practices, fertilizer applications, and pesticide use. SWAT includes databases with default parameter values to select or modify.

After all input parameters have been determined, the interface writes all necessary input files for SWAT to use during simulation. These files will be stored and can be overwritten manually or by the interface for different simulations. Figure 6.6 illustrates an initialized watershed with stream networks and individual subbasins that have defined management and environmental parameters associated to each one stored in corresponding files and accessed by

SWAT during simulation. In the context of this example, the SWAT *management* files were our main input file of interest. Each subbasin has a management file associated with it that specifies important information such as crop variety, planting and harvesting dates, irrigation schedules, and fertilization application schedules and amounts.

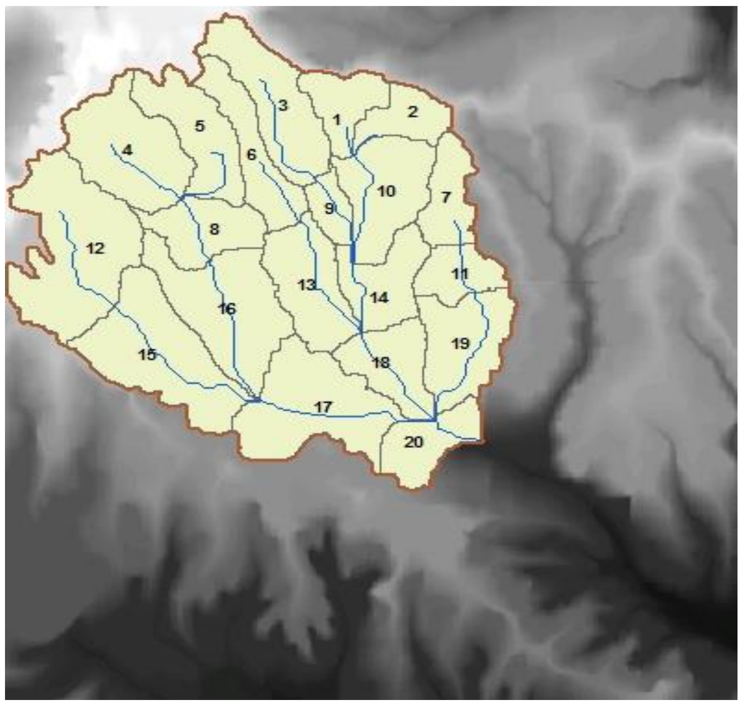

Figure 6.6: Delineated Watershed with Subbasins and Stream Network

```
.mgt file Watershed HRU:26 Subbasin:26 HRU:1 Luse:PAST Soil: TX633 Slope: 0-1 4/29/2013 1:
              0     | NMGT:Management code
Initial Plant Growth Parameters
              0     | IGRO: Land cover status: 0-none growing; 1-growing
              0     | PLANT_ID: Land cover ID number (IGRO = 1)
           0.00     | LAI_INIT: Initial leaf are index (IGRO = 1)
           0.00     | BIO_INIT: Initial biomass (kg/ha) (IGRO = 1)
           0.00     | PHU_PLT: Number of heat units to bring plant to maturity (IGRO = 1)
General Management Parameters
           0.20     | BIOMIX: Biological mixing efficiency
          84.00     | CN2: Initial SCS CN II value
           1.00     | USLE_P: USLE support practice factor
           0.00     | BIO_MIN: Minimum biomass for grazing (kg/ha)
          0.000     | FILTERW: width of edge of field filter strip (m)
Urban Management Parameters
              0     | IURBAN: urban simulation code, 0-none, 1-USGS, 2-buildup/washoff
              0     | URBLU: urban land type
Irrigation Management Parameters
              0     | IRRSC: irrigation code
              0     | IRRNO: irrigation source location
          0.000     | FLOWMIN: min in-stream flow for irr diversions (m^3/s)
          0.000     | DIVMAX: max irrigation diversion from reach (+mm/-10^4m^3)
          0.000     | FLOWFR: : fraction of flow allowed to be pulled for irr
Tile Drain Management Parameters
          0.000     | DDRAIN: depth to subsurface tile drain (mm)
          0.000     | TDRAIN: time to drain soil to field capacity (hr)
          0.000     | GDRAIN: drain tile lag time (hr)
Management Operations:
              1     | NROT: number of years of rotation
Operation Schedule:
         0.150  1  119          2016.00000   0.00     0.00000 0.00   0.00  0.00
         1.200  5                  0.00000
         0.150  3   30            40.00000   0.00
         0.150 10    1   0         0.00000   0.00     0.00000 0.00
```

Figure 6.5: Example Management Input File

After defining simulation parameters (dates, time steps, etc) SWAT will run the simulation, then write and store watershed output data in a separate file for review. SWAT outputs include crop yield, nitrogen and phosphorous yields, sediment yields, water demands, and many others.

```
1
    SWAT Sept 7 2012    VER 2012/Rev 573                                                                0/ 0/   0      0: 0: 0

    General Input/Output section (file.cio):
    4/29/2013 12:00:00 AM ARCGIS-SWAT interface AV

         Number of years in run:     6
         Area of watershed:       486.830 km2
1
    SWAT Sept 7 2012    VER 2012/Rev 573

    General Input/Output section (file.cio):
    4/29/2013 12:00:00 AM ARCGIS-SWAT interface AV

Annual Summary for Watershed in year    1 of simulation

UNIT                                     PERCO   TILE                            WATER    SED     NO3     NO3     NO3     NO3        N       P       P
TIME    PREC     SURQ    LATQ      GWQ   LATE      Q      SW      ET      PET    YIELD  YIELD    SURQ    LATQ    PERC    CROP  ORGANIC SOLUBLE ORGANIC TILENO3
        (mm)     (mm)    (mm)     (mm)   (mm)   (mm)    (mm)    (mm)    (mm)     (mm)   (mm)    -------------------(kg nutrient/ha)-------------------- (kg/ha)
   1   35.60     0.00    0.05     0.00   0.00   0.00  133.89   34.24  125.10     0.05   0.00    0.00    0.00    0.00    0.00     0.00    0.00    0.00    0.00
   2  115.22    25.00    0.32     0.00   0.00   0.00  171.46   51.42  132.96     8.65   0.01    0.04    0.01    0.00    0.00     0.04    0.00    0.00    0.00
   3  124.16    22.93    0.46     0.13   1.84   0.00  216.60   54.14   96.15    16.88   0.03    0.07    0.01    1.33    0.00     0.12    0.00    0.01    0.00
   4   46.47     4.45    0.10     0.71   0.03   0.00  189.81   72.77  175.18    11.81   0.02    0.04    0.00    0.02   36.32     0.09    0.00    0.01    0.00
   5   94.29    16.29    0.26     0.50   0.00   0.00   30.81  241.04  297.94    14.53   0.01    0.04    0.00    0.00   27.33     0.05    0.00    0.01    0.00
   6  115.97     0.39    0.30     0.18   0.00   0.00   27.00  122.77  252.88     9.23   0.00    0.02    0.00    0.00    1.25     0.02    0.00    0.00    0.00
   7   74.83     0.00    0.24     0.07   0.00   0.00    6.28  100.34  259.66     5.09   0.00    0.01    0.00    0.00    1.01     0.01    0.00    0.00    0.00
   8   66.00     0.01    0.18     0.03   0.00   0.00   16.40   60.56  228.74     2.78   0.00    0.01    0.00    0.00    0.57     0.01    0.00    0.00    0.00
   9   97.06     8.06    0.36     0.01   0.00   0.00   56.46   50.59  217.09     4.50   0.00    0.01    0.00    0.00    0.05     0.00    0.00    0.00    0.00
  10   81.49     2.26    0.29     0.00   0.00   0.00   88.76   49.01  144.81     4.61   0.00    0.01    0.00    0.00    0.00     0.00    0.00    0.00    0.00
  11  119.01    13.29    0.48     0.00   0.00   0.00  153.22   40.07   78.77     6.64   0.00    0.02    0.01    0.00    0.00     0.01    0.00    0.00    0.00
  12  191.35    68.39    0.45     0.01   1.44   0.00  223.18   43.95  100.75    21.27   0.01    0.06    0.00    0.13    0.00     0.02    0.00    0.00    0.00

 1999 1161.45  161.08    3.47     1.65   3.30   0.00  223.18  920.90 2110.03   106.03   0.08    0.33    0.04    1.48   66.53     0.36    0.02    0.04    0.00
```

Figure 6.7: Selection from Example Output File

```
SWAT Sept 7 2012    VER 2012/Rev 573

General Input/Output section (file.cio):
4/29/2013 12:00:00 AM ARCGIS-SWAT interface AV

                                          Average Plant Values (kg/ha)

HRU       1 SUB   1 SWCH  Yld =  5011.8 BIOM =   7163.9
HRU       2 SUB   2 SWCH  Yld =  5184.9 BIOM =   7407.0
HRU       3 SUB   3 SWCH  Yld =  5148.0 BIOM =   7307.1
HRU       4 SUB   4 SWCH  Yld =  4785.7 BIOM =   6794.7
HRU       5 SUB   5 MISC  Yld =  5762.3 BIOM =   7317.2
HRU       6 SUB   6 MISC  Yld =  4008.9 BIOM =   5105.3
HRU       7 SUB   7 PAST  Yld =  2496.7 BIOM =   3557.7
HRU       8 SUB   8 SWCH  Yld =  5443.0 BIOM =   7725.4
HRU       9 SUB   9 SWCH  Yld =  4802.4 BIOM =   6872.3
HRU      10 SUB  10 SWCH  Yld =  4779.7 BIOM =   6844.3
HRU      11 SUB  11 SWCH  Yld =  4802.3 BIOM =   6872.1
HRU      12 SUB  12 MISC  Yld =  5757.9 BIOM =   7311.4
HRU      13 SUB  13 MISC  Yld =  4285.0 BIOM =   5456.9
HRU      14 SUB  14 SWCH  Yld =  3641.4 BIOM =   5174.3
HRU      15 SUB  15 MISC  Yld =  5865.6 BIOM =   7472.4
HRU      16 SUB  16 PAST  Yld =  2628.7 BIOM =   3742.4
HRU      17 SUB  17 SWCH  Yld =  4076.9 BIOM =   5785.1
HRU      18 SUB  18 MISC  Yld =  3733.5 BIOM =   4754.5
```

Figure 6.8: Selection from Example Output File

6.2 Problem Formulation

For the optimal control model, a chromosome or potential solution consists of two decision alternatives, which are the feedstock selection and fertilization amounts to be applied during the decision period. Land cover and fertilization amount combinations are utilized to generate potential decision alternative scenarios in search for optimal agricultural landscapes. Table 6.1 lists the combinations that will be used in this application. The land covers include switchgrass, miscanthus, and original conventional land covers for the demonstration watersheds and fertilization applications of 90, 30, and 0 kg ha^{-1}. Other crops can also be included in the scenarios if desired.

Table 6.1: Possible Land Covers and Fertilization Applications for Each HRU

Integer Code	Land Cover and Fertilization Options
1	Conv. Crop/Original Land Cover – No Change
2	Switchgrass – 90 kg ha^{-1} Nitrogen application
3	Switchgrass – 30 kg ha^{-1} Nitrogen application
4	Switchgrass – 0 kg ha^{-1} Nitrogen application
5	Miscanthus – 90 kg ha^{-1} Nitrogen application
6	Miscanthus – 30 kg ha^{-1} Nitrogen application
7	Miscanthus – 0 kg ha^{-1} Nitrogen application

A possible scenario to be evaluated by SWAT is represented by a chromosome that consists of the land cover and fertilization options to be randomly assigned for each subbasin of the watershed. The chromosome string created for the optimization problem consists of genes equal to the number of subbasins in the watershed to be evaluated. Each alternative land cover and management scenario is represented by the chromosome as a series of particular decision parameters or genes by the integer codes in Table 6.1. The following response variables will be evaluated at the watershed outlet for each chromosome or possible scenario: average annual sediment yield, nitrogen and phosphorous yields, and biomass accumulation. The current objective function for this problem is to minimize sediment and nutrient yields to the watershed,

while maximizing biomass production. The multi-objective optimization problem, is formulated as follows:

$$\text{Minimize} \quad [\, SY + PY + NY + BY^{-1} \,] \tag{6.1}$$

$$\text{where } SY = \sum_{i=n} S_i \;, PY = \sum_{i=n} P_i \;, NY = \sum_{i=n} N_i \;, BY = \sum_{i=n} B_i \tag{6.2}$$

where SY is average annual sediment yield, PY is average phosphorous yield, NY is average annual nitrogen yield, BY is average annual biomass yield, for each HRU, i. The objective function to be evaluated by the genetic algorithm for each possible scenario is (Eq. 6.1), the sum of all components for each i, where n is the total number of HRUs or subbasins.

6.3 Solution Methodology

The optimal control model was developed by interfacing SWAT with an MOEA and operates at a spatial scale of HRU or subbasin. The OCM begins by creating a population or set of landscape scenarios with random land covers and management practices that satisfy the crop management constraints for each of the subbasins in the entire watershed. This is followed by scheduling management operations required for watershed simulation and used in the decision process. Thus, management files containing these scheduled operations are prepared for each of the HRUs. The information contained in these files includes dates of planting and harvesting; required heat units; Curve Numbers for the land cover that account for tillage and soil types; and times and dosages of fertilizer application. SWAT is then called to simulate the physical responses of each subbasin or HRU, including the average annual sediment, phosphorus and nitrogen yields, and crop yield for each year of the decision period.

Chapter 7: Multiple Objective Evolutionary Algorithm for OCM

In order to develop an optimal control model for the management of ecosystems services by evolving agricultural landscapes when converting traditional agriculture to biofuel feedstock cultivation a MOEA was be required. This research implemented a customized MOEA developed by Taboada and Coit (2008), which incorporates aspects from various metaheuristic methods with modifications. This MOEA was implemented and customized for the proposed problem and provided quality approximations to global optimal solutions. The modifications this algorithm includes from other MOEAs are the utilization of three different fitness functions and a subsystem rotation crossover during the reproduction stage. The multi-objective optimization ensures that multiple factors such as nitrogen, phosphorous, and sediment yields, and biomass production are all taken into consideration when searching for optimal solutions. The MOEA used follows the same procedure of initialization, evaluation, selection, and reproduction. The following figure shows a general flow chart of the algorithm.

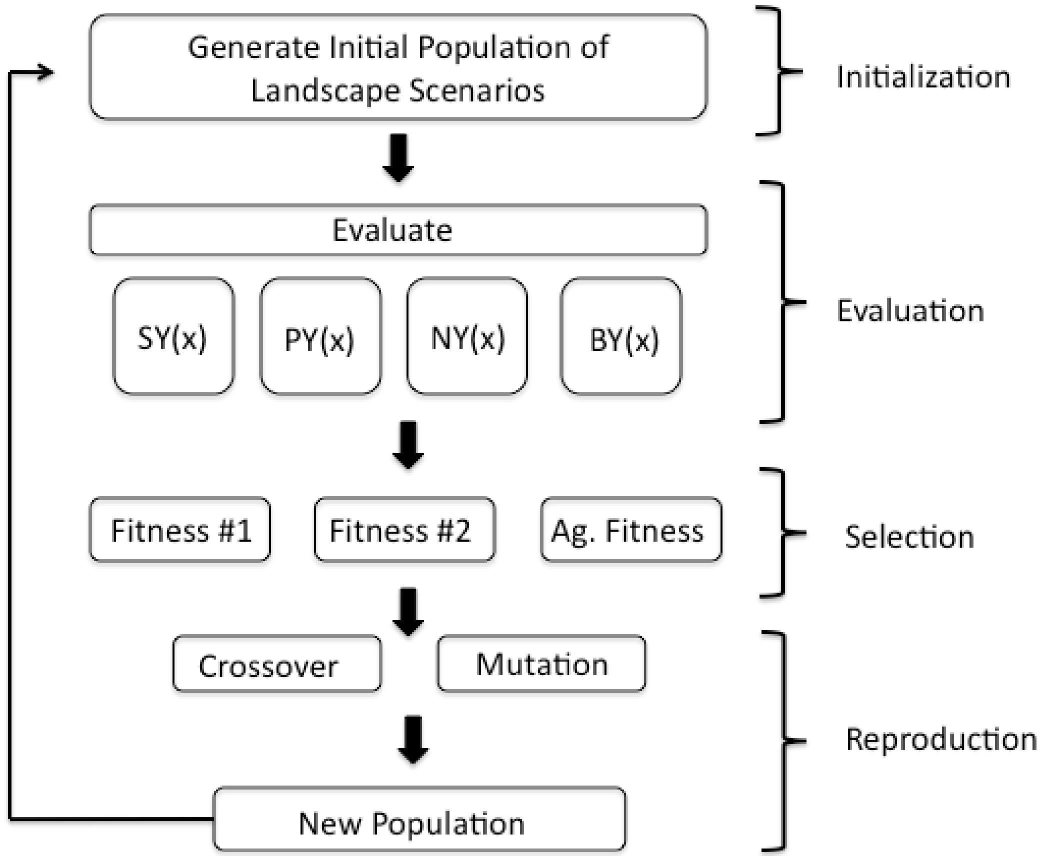

Figure 7.1: Flowchart of Genetic Algorithm

7.1 Initialization

The algorithm begins by creating an initial set of possible solutions. Vital to the quality of the future generations, this initial set is created randomly to ensure a diverse set of solutions. This variety of genetic material ensures that the search space is thoroughly and efficiently explored. In the proposed research, the algorithm's population will consist of a specific number

of possible landscape scenarios for the watershed. Each landscape scenario consists of a combination of land covers, including biofuel feedstocks, and fertilization amounts for each HRU. These combinations are utilized to generate potential decision alternative scenarios in search for optimal agricultural landscapes. These combinations were listed in Table 6.1.

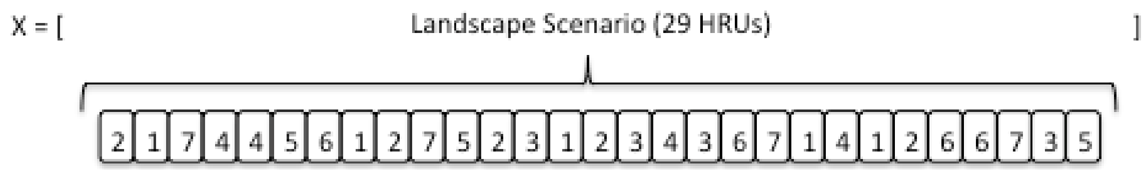

Figure 7.2: Example Chromosome Representation of Individual Landscape Scenario

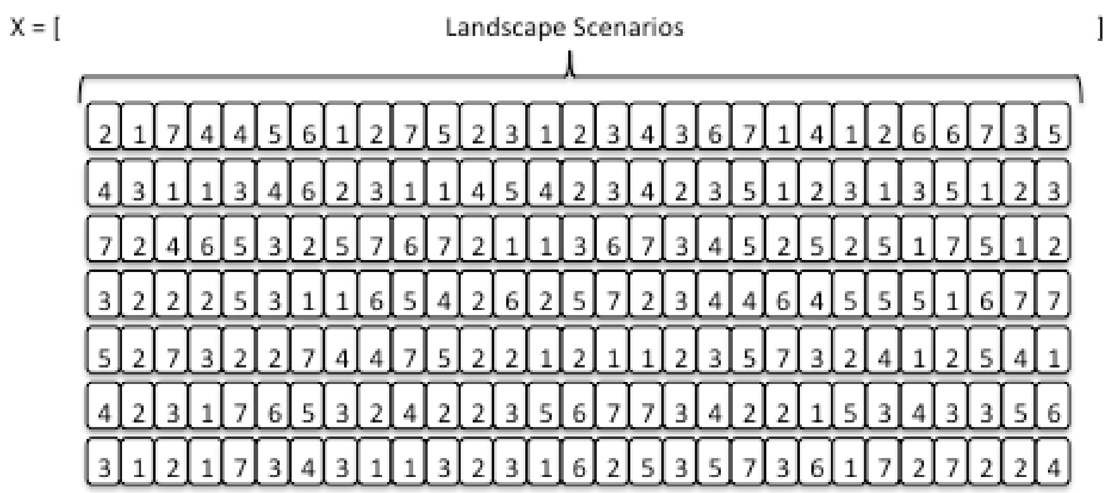

Figure 7.3: Example of an Initial Population

7.2 Evaluation

The evaluation stage follows and the four objectives considered in the proposed problem are evaluated for each possible scenario of feedstock production. The algorithm incorporates the concept of Pareto dominance. After each objective is evaluated for every possible solution, solutions that are dominated by others are removed from the population and only nondominated solutions continue on the next stage of the algorithm.

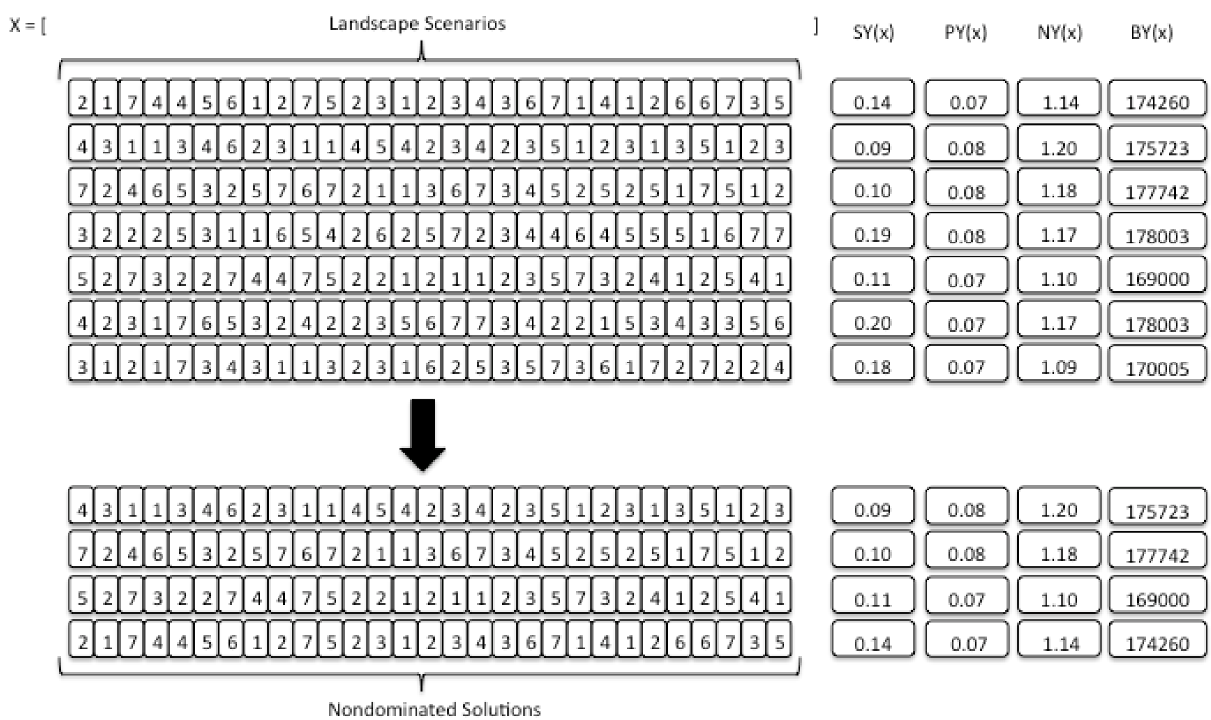

Figure 7.4: Example of Evaluated and Nondominated Solutions

7.3 Selection

The next stage of the algorithm is selection. In this critical stage, parent solutions are selected to create new solutions for the next generation. The best solutions from the set of nondominated solutions are selected by three fitness functions. Evolutionary algorithms use fitness functions to measure the quality of the evaluated solutions. To ensure a quality solution three different fitness functions are used, one to select the most dominating individuals and achieve proximity and the other to select individuals with greater distances from other solutions to ensure population diversity. Finally an aggregated fitness metric is also used.

7.3.1 Fitness Metric #1

Fitness metric, $f_1(x)$, is a dominance count-based metric. By selecting individuals that are more dominating it intends to find solutions that are near the true Pareto front. Evaluating every possible solution would generate a true Pareto or global Pareto optimal solution set. However, evaluating every possible solution is not practical and the true Pareto front is not known. In order to estimate the true Pareto front, this fitness metric selects solutions that dominate more solutions than others, with the idea that those solutions generally lie closest to the true Pareto front. In other words, solutions that dominate more solutions receive a higher, or better, fitness value.

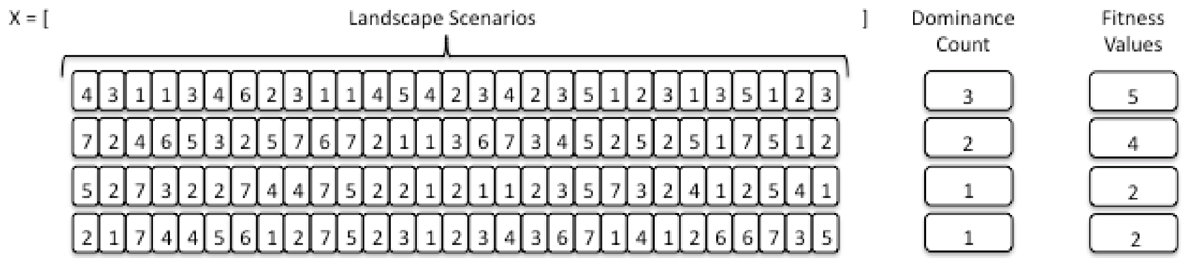

Figure 7.5: Example of Fitness Metric #1 Evaluation

Fitness Metric #1 identifies solutions of better fitness by evaluating their dominance count. For example, the nondominated solutions in Figure 7 have a range of dominance counts from one to three. The maximum dominance count is three, and we will use an interval range of five to assign a fitness value for each solution. The fitness values are calculated by using Table 7.1.

Table 7.1: Example Fitness Values for Interval

Fitness Value	Intervals
1	0 ≤ Dominance count < 0.6
2	0.6 ≤ Dominance count < 1.2
3	1.2 ≤ Dominance count < 1.8
4	1.8 ≤ Dominance count < 2.4
5	2.4 ≤ Dominance count ≤ 3.0

7.3.2 Fitness Metric #2

The second fitness metric, $f_2(x)$, aims to create diversity in the solution set and cover the entire search space. This fitness metric is distance based and selects solutions that are further away from other solutions, in order to resist the possibility of converging on local optimal rather

than finding global optimal solutions. To prevent this possibility the fitness metric assigns a better value to solutions with greater distances from others. The underlying assumption to this fitness metric is that distant solutions will also generate distant solutions and improve the diversity of continuing generations.

Considering the nondominated solutions from Figure 7.4, Fitness Metric #2 first normalizes the objective values.

Table 7.2: Example Normalized Objectives

Nondominated Solution	SY(x)	PY(x)	NY(x)	BY(x)
1	-1.069	1	1.172	0.476
2	-0.535	1	0.651	1.100
3	0	-1	-1.432	-1.601
4	1.604	-1	-0.391	0.024

Fitness Metric #2 proceeds by calculating the Euclidean distance of each solution from the others. The sum of the distances from each objective of each solution is also obtained.

Table 7.3: Example Euclidean Distances

Individual	1	2	3	4
1	0	0.9100	4.3246	4.0030
2	0.9100	0	3.6010	3.2773
3	4.3246	3.6010	0	2.1774
4	4.0030	3.2773	2.1774	0
Sum	9.2376	7.7883	10.1030	9.4577
Min	7.7883			
Max	10.1030			

Similar to Fitness Metric #1, intervals are defined and ranges are calculated from the minimum and maximum sums of distances. Figure 7.6 shows the nondominated solutions and their respective fitness values.

Table 7.4: Example Intervals for Fitness Metric #2

Fitness Value	Intervals
1	7.7883 ≤ Dominance count < 8.2512
2	8.2512 ≤ Dominance count < 8.7142
3	8.7142 ≤ Dominance count < 9.1771
4	9.1771 ≤ Dominance count < 9.6401
5	9.6401 ≤ Dominance count ≤ 10.1030

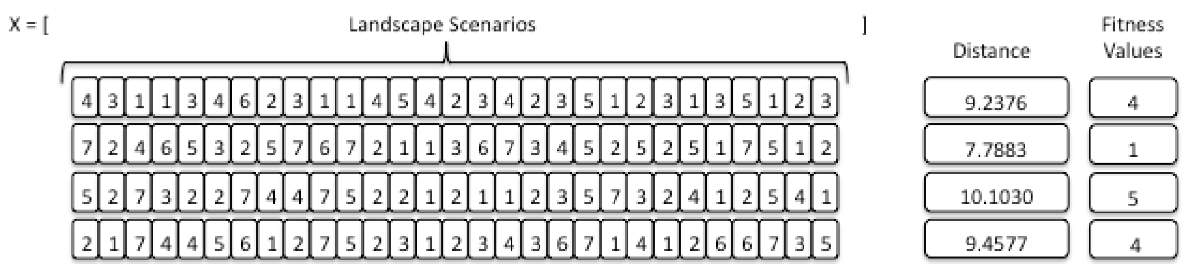

Figure 7.6: Example of Fitness Metric #2 Evaluation

7.3.3 Fitness Metric #3

Lastly, the algorithm utilizes a third fitness metric, $f_3(x)$, which is basically an aggregate of fitness metrics one and two. Both metrics are equally weighted and the solutions are ranked. These solutions are presumed to be close to the true Pareto front as well as diverse. Solutions with the highest aggregated values will be more likely to undergo crossover and create the next generation of solutions.

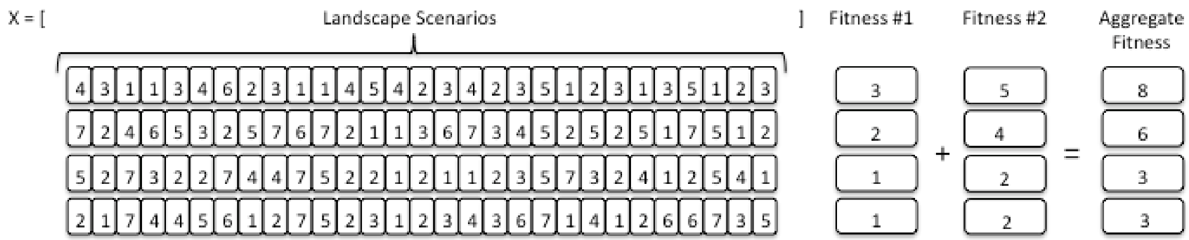

Figure 7.7: Example of Aggregated Fitness Metric

7.4 Reproduction

After the entire population has been evaluated and the solutions are selected and ranked the reproduction step generates new individuals, or possible solutions, for the next generation. Reproduction incorporates three operators, elitism, crossover, and mutation.

Elitism simply ensures that the best solutions form each generation are not lost and selected to survive into the next generation. Twenty five percent of the nondominated solutions are selected and continue into the next generation for evaluation while the entire group is used for crossover.

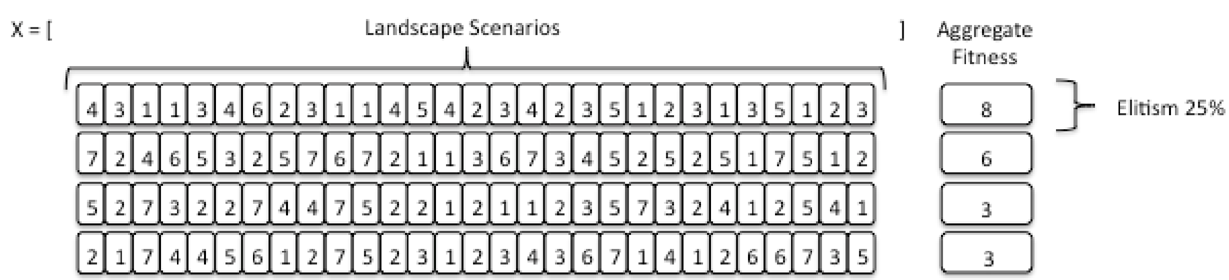

Figure 7.8: Example Nondominated Solutions Selected for Elitism and Crossover

Subsystem rotation crossover (SURC) is used as the crossover operator in this algorithm. Subsystem rotation crossover operates by dividing the chromosome into several subsystems. Of the subsystems, one is used to rotate and exchange information among solutions while the others remain fixed. This rotation continues until the subsystem returns to its original position. This method of crossover creates a greater number of offspring in the mating pool and supplies a greater amount of diverse solutions to choose from.

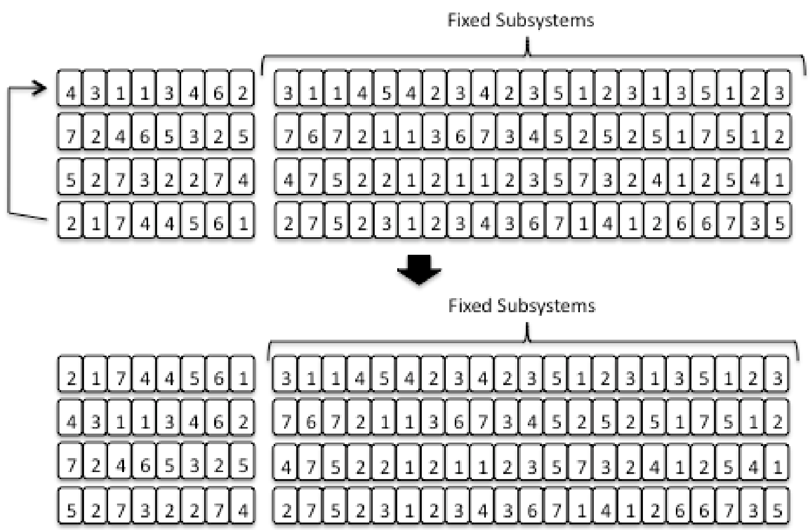

Figure 7.9: Example of Subsystem Rotation Crossover

Two point mutation with a probability of 0.01 was used in this algorithm. One percent of the population will undergo mutation by randomly selecting two points on the chromosome that will switch their values.

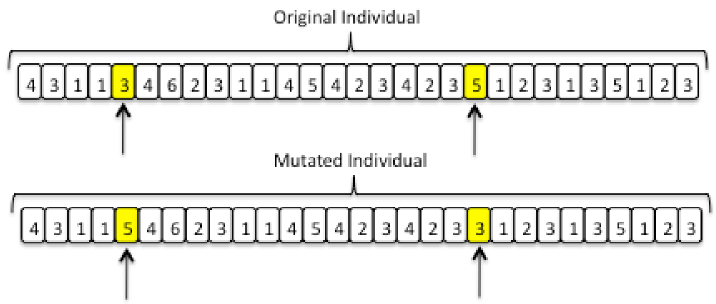

Figure 7.10: Example of Two Point Mutation

Ultimately, the new generation will consist of 25% elite parents and 75% children randomly selected from the mating pool. This new population will reenter the evaluation stage and repeat the entire process until the specified number of generations is met. Once this stopping criterion is met, the nondominated solutions from the last generation represent the near optimal solutions to the problem and potential landscapes for biofuel feedstock production.

Chapter 8: Results

To demonstrate the integrated modeling approach and the functionality of the optimal control model, it was applied to two different example watersheds. In order to illustrate the optimal control models versatility and ease of application the watersheds selected were of different spatial scales and located in different geographic regions with varying original land covers and watershed attributes. SWAT's built in weather generator was used for the temperature, precipitation, humidity, and wind inputs.

8.1 Lake Fork Watershed Application

First, the optimal control model was used to evaluate a portion of the Lake Fork watershed. The area of interest was a 487 km^2, agriculturally dominated basin consisting primarily of pasture and row crops, within the Lake Fork watershed located in the Texas-Gulf Region. The watershed was divided into 32 subbasins and the OCM began by evaluating an initial population of 40 individuals and reiterating the multiobjective evolutionary algorithm for 40 generations.

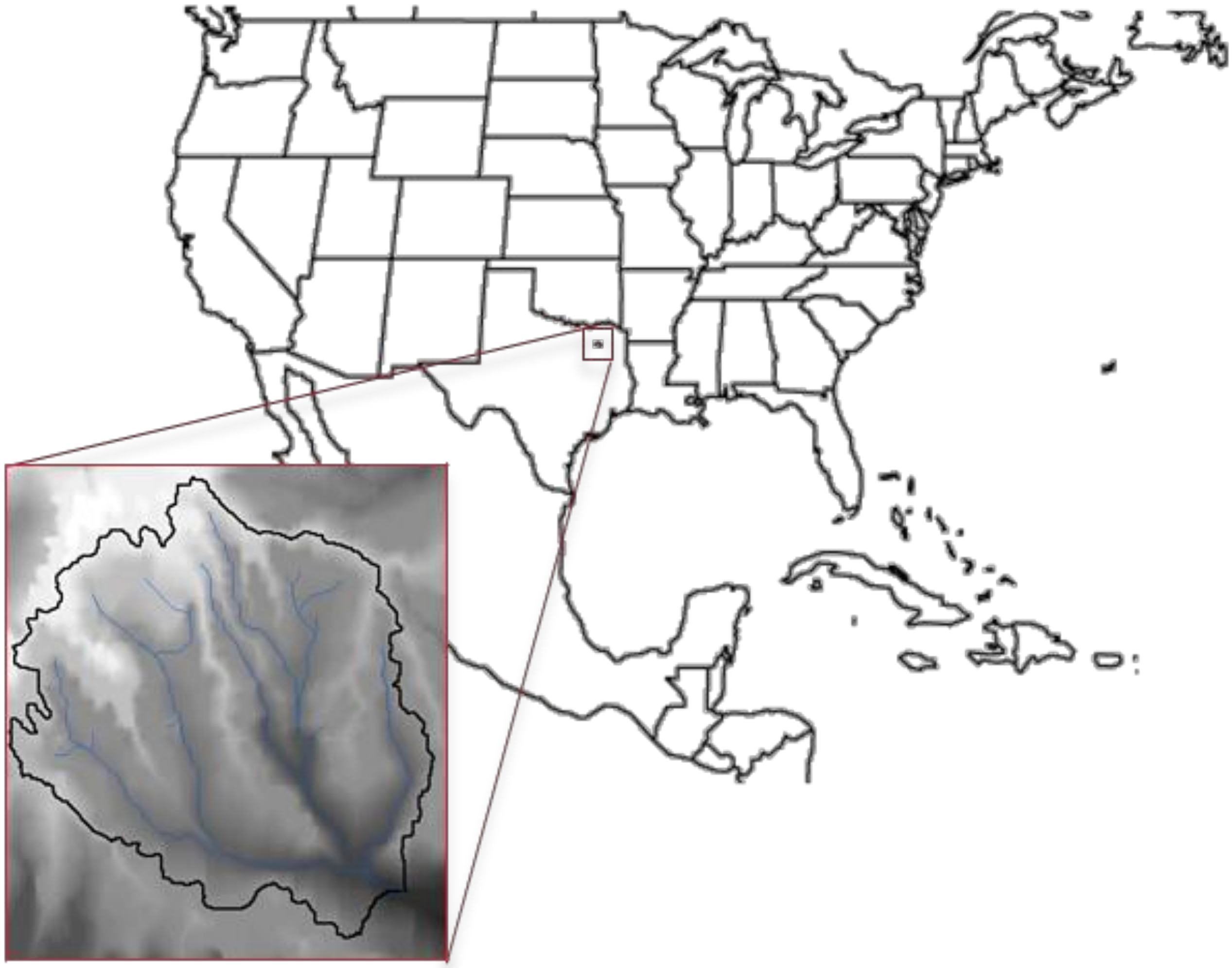

Figure 8.1: Lake Fork Watershed

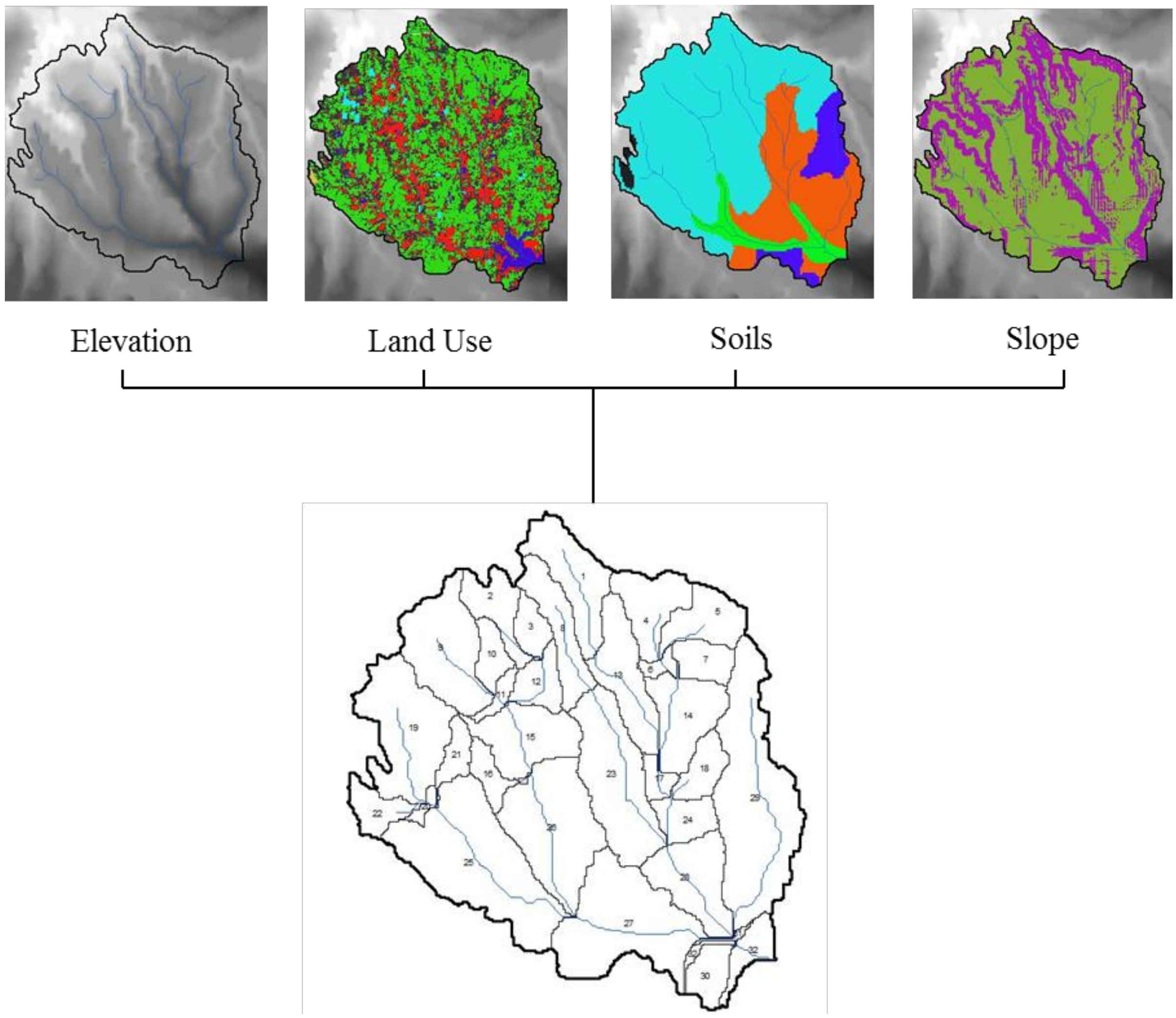

Figure 8.2: Final Delineated Lake Fork Watershed with Stream Network

Figure 8.3 shows the Pareto solutions and tradeoffs with respect to feedstock yield, sediment yield, and nitrate runoff to the watershed outlet in the Lake Fork example.

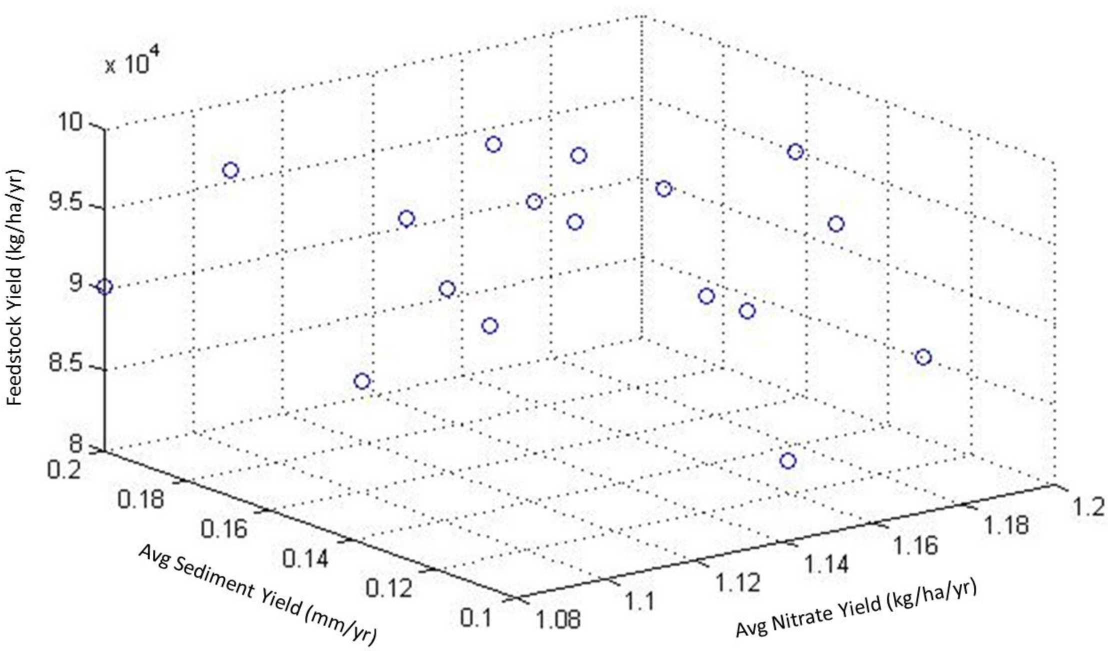

Figure 8.3: Pareto Solutions for Lakefork Watershed.

Pareto Solutions	SUB 1	SUB 2	SUB 3	SUB 4	SUB 5	SUB 6	SUB 7	SUB 8	SUB 9	SUB 10	SUB 11	SUB 12	SUB 13	SUB 14	SUB 15	SUB 16	SUB 17	SUB 18	SUB 19	SUB 20	SUB 21	SUB 22	SUB 23	SUB 24	SUB 25	SUB 26	SUB 27	SUB 28	SUB 29
1	4	7	1	4	1	7	1	6	6	7	1	3	2	6	4	7	2	2	2	1	7	5	4	2	6	5	3	4	3
2	3	5	3	6	6	2	5	5	4	7	2	3	1	7	5	4	5	4	5	4	6	4	7	2	1	1	1	3	4
3	6	2	4	5	1	5	3	1	4	2	1	2	2	2	1	5	2	4	5	4	4	4	1	4	6	7	2	2	4
4	4	5	1	6	4	6	3	4	1	2	5	3	7	1	7	4	5	7	3	3	4	6	6	1	2	3	1	4	3
5	1	6	7	6	1	3	5	6	2	2	5	4	5	2	2	6	6	7	1	2	1	4	2	7	1	1	4	6	3
6	6	3	5	7	4	5	6	4	5	1	7	2	2	6	4	6	3	2	1	5	4	4	5	1	3	6	5	1	1
7	7	5	7	4	4	6	2	4	7	5	6	6	5	2	5	1	5	5	6	7	7	6	5	7	5	1	1	7	4
8	2	3	5	2	3	5	2	6	7	6	3	5	1	7	7	6	2	5	1	3	3	2	2	3	1	5	4	5	5
9	5	3	5	5	1	2	2	5	6	3	6	5	1	5	3	7	1	4	3	4	6	3	6	4	1	2	6	4	2
10	2	5	2	6	7	5	6	6	3	7	7	3	6	5	4	7	1	4	5	2	3	2	6	1	4	1	4	2	2
11	6	7	1	7	5	1	2	4	7	7	2	7	7	4	6	1	3	7	6	7	5	1	6	7	5	6	6	3	5
12	5	7	2	2	3	1	5	3	7	3	5	2	3	2	6	7	3	5	3	4	6	5	3	3	4	3	3	4	6
13	2	3	2	4	7	5	1	3	1	4	4	7	6	4	7	1	3	7	7	5	5	3	7	1	6	5	6	3	6
14	3	7	3	5	5	4	5	5	2	1	7	2	1	4	7	5	2	3	4	7	2	6	5	3	2	3	4	1	5
15	1	2	5	6	5	4	4	3	6	2	5	2	3	5	6	1	7	6	4	4	4	3	4	4	6	6	5	3	6
16	1	6	3	5	6	1	1	4	4	7	6	6	1	1	1	6	7	5	1	6	1	1	5	3	5	6	5	6	2
17	6	3	4	1	1	4	6	7	1	4	4	1	3	2	6	3	4	2	5	2	5	5	6	4	1	2	7	2	6

Figure 8.4: Decision Alternatives for Lakefork Watershed

Solution	Feedstock Yield (kg/ha/yr)	NO_3 Yield (kg/ha/yr)	P Yield (kg/ha/yr)	Sediment Yield (mm/yr)
1	91934	1.15	0.07	0.12
2	96733	1.13	0.07	0.14
3	83599	1.15	0.08	0.11
4	88166	1.18	0.08	0.11
5	99719	1.17	0.08	0.13
6	97135	1.13	0.07	0.15
7	98657	1.09	0.07	0.18
8	99786	1.13	0.07	0.16
9	99459	1.14	0.07	0.15
10	91975	1.15	0.07	0.13
11	90226	1.08	0.07	0.2
12	96129	1.17	0.07	0.12
13	85358	1.11	0.07	0.17
14	97687	1.15	0.07	0.14
15	94857	1.12	0.07	0.17
16	91428	1.12	0.07	0.16
17	89999	1.12	0.07	0.15

Figure 8.5: Outputs to Watershed Outlet for Lakefork Example

8.2 San Juan Watershed Application

Secondly, it was applied to the San Juan River watershed located in Hidalgo, Mexico. The San Juan River watershed is 4136 km^2, and consists of a mixture of grassland, cropland, and woodland shrubs. It was divided into 47 subbasins for evaluation with the OCM, beginning with an initial population of 20 and reiterating for 40 generations. The OCMs capacity to evaluate much larger watersheds and identify Pareto solutions is illustrated below.

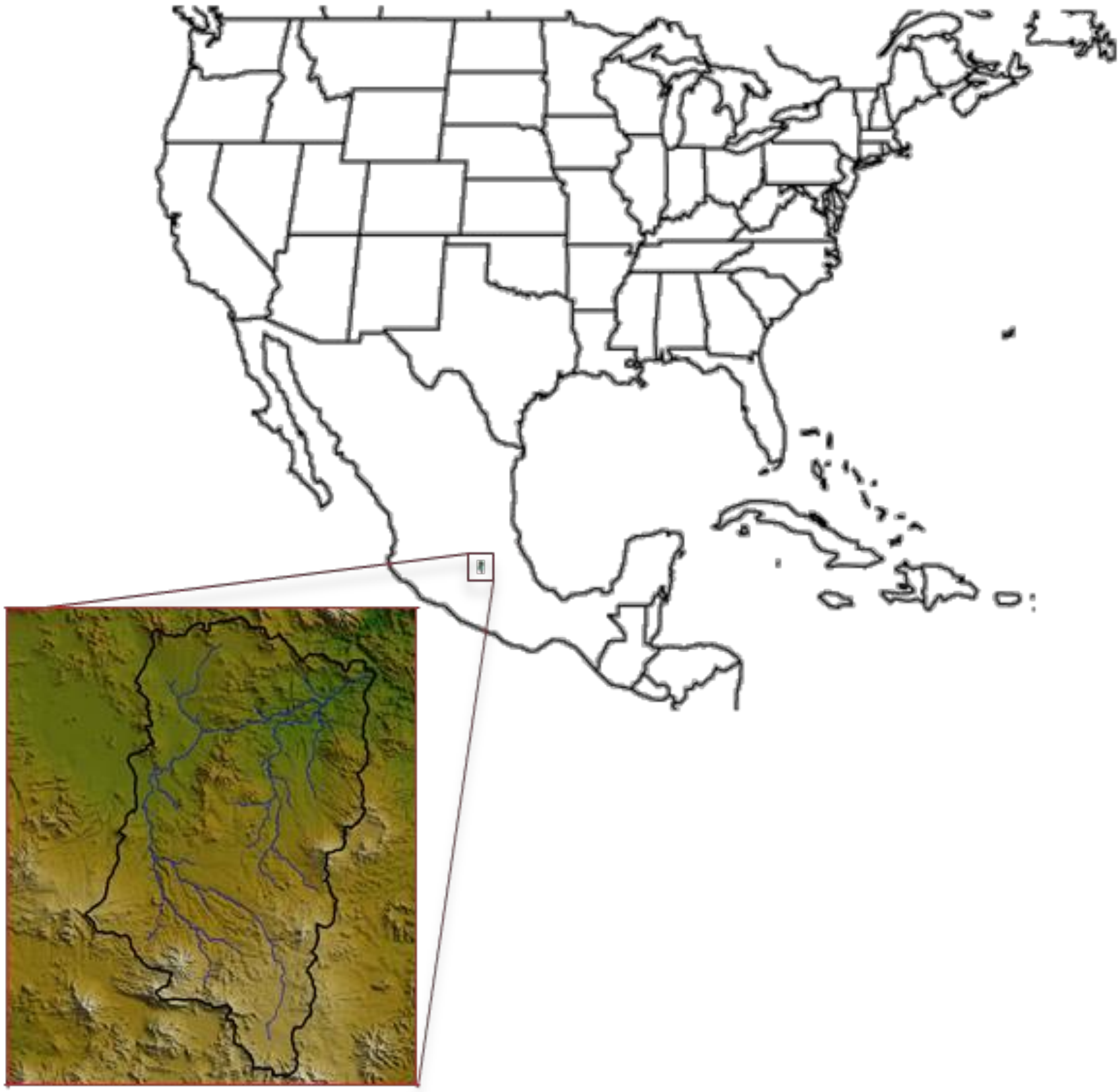

Figure 8.6: San Juan Watershed

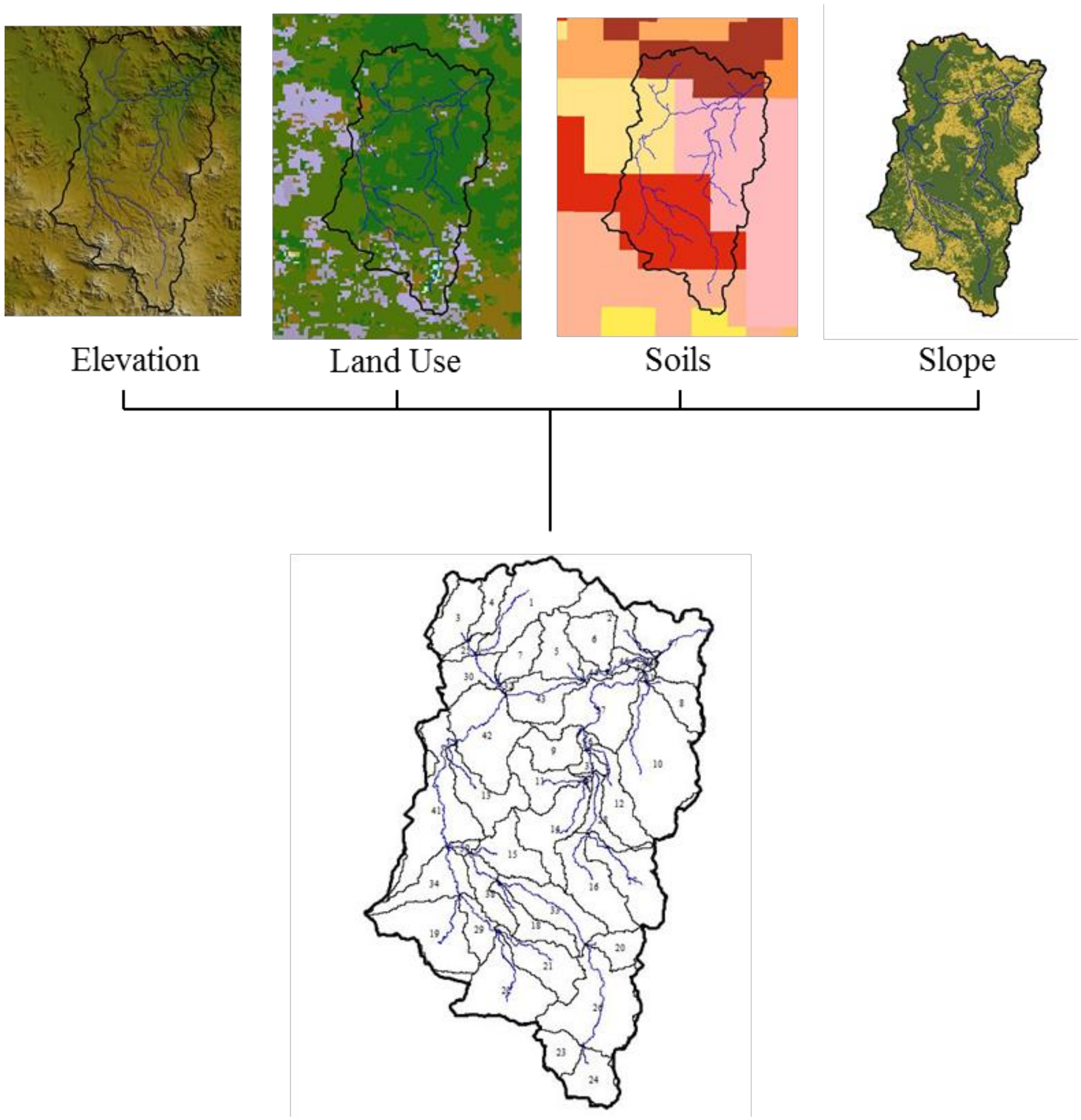

Figure 8.7: Final Delineated San Juan Watershed with Stream Network

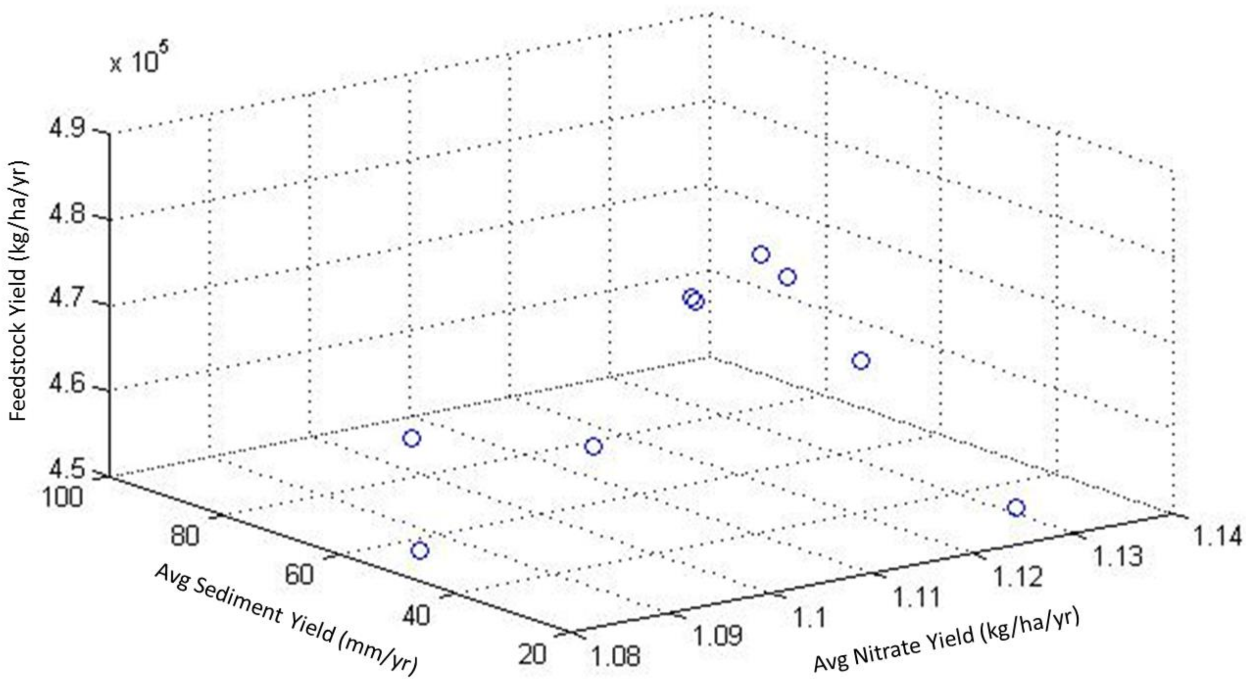

Figure 8.8: Pareto Solutions for San Juan Watershed.

Pareto Solutions	SUB 1	SUB 2	SUB 3	SUB 4	SUB 5	SUB 6	SUB 7	SUB 8	SUB 9	SUB 10	SUB 11	SUB 12	SUB 13	SUB 14	SUB 15	SUB 16	SUB 17	SUB 18	SUB 19	SUB 20	SUB 21	SUB 22	SUB 23	SUB 24
1	4	2	3	3	2	3	5	2	5	4	4	2	2	5	2	4	7	6	5	7	6	2	6	6
2	6	2	6	6	6	3	3	7	4	4	5	7	4	5	4	7	7	2	2	2	6	2	7	5
3	4	2	4	5	4	5	7	3	6	2	1	3	6	3	7	5	3	7	7	6	2	2	6	5
4	7	3	1	4	6	3	5	5	4	4	7	1	2	4	5	7	4	4	6	4	7	4	4	4
5	1	6	3	6	6	3	1	7	1	5	6	7	4	5	5	3	6	7	5	5	5	1	6	2
6	2	4	1	3	1	4	5	6	3	4	7	7	4	7	1	6	3	2	6	3	2	1	7	5
7	2	3	4	3	3	4	2	1	5	2	6	3	3	6	6	2	1	2	5	6	4	2	6	7
8	6	2	1	3	4	7	5	6	4	4	7	3	4	2	6	4	3	5	1	1	5	5	5	7
9	4	7	1	4	4	7	5	7	5	6	4	4	2	5	7	4	4	6	1	7	1	4	5	4

Pareto Solutions	SUB 25	SUB 26	SUB 27	SUB 28	SUB 29	SUB 30	SUB 31	SUB 32	SUB 33	SUB 34	SUB 35	SUB 36	SUB 37	SUB 38	SUB 39	SUB 40	SUB 41	SUB 42	SUB 43	SUB 44	SUB 45	SUB 46	SUB 47
1	6	3	3	7	4	4	5	7	4	5	4	7	7	2	2	2	6	2	7	5	2	7	7
2	2	7	7	4	2	3	3	2	3	5	2	5	4	4	2	2	5	2	4	7	6	5	7
3	4	7	4	7	4	4	6	3	4	1	6	4	5	6	2	7	7	3	6	3	2	2	5
4	2	1	1	7	2	7	5	7	7	1	3	7	5	7	4	1	5	2	4	6	5	5	4
5	3	3	3	1	1	2	5	5	7	7	3	4	2	4	6	7	5	1	4	6	2	4	2
6	7	7	5	7	2	7	7	3	1	6	1	6	4	5	2	4	5	1	4	7	1	4	5
7	2	7	2	6	7	4	4	7	1	4	5	6	2	1	1	6	7	3	7	3	1	3	4
8	5	4	2	1	6	2	2	2	4	4	6	7	6	7	2	7	1	2	1	5	7	5	2
9	4	3	5	7	5	7	1	4	1	4	5	2	6	3	7	3	1	2	1	7	1	1	1

Figure 8.9: Decision Alternatives for San Juan Watershed

Solution	Feedstock Yield (kg/ha/yr)	NO_3 Yield (kg/ha/yr)	P Yield (kg/ha/yr)	Sediment Yield (mm/yr)
1	481482	1.10	0.05	33.64
2	481173	1.10	0.05	33.21
3	482960	1.11	0.05	39.11
4	453658	1.08	0.04	45.94
5	450712	1.13	0.04	29.78
6	453810	1.10	0.04	82.22
7	468020	1.12	0.04	39.19
8	474896	1.12	0.04	52.00
9	453837	1.11	0.04	68.13

Figure 8.10: Outputs to Watershed Outlet for San Juan Example

The Pareto solutions indicated in Figures 8.3 and 8.8 depict decision alternatives or landscape scenarios that achieve the prescribed goals of maximizing feedstock production while aiming to minimize nutrient runoff and erosion. Although no one solution is better than the other, they are all near optimal scenarios that illustrate the tradeoffs and impacts of feedstock production. Users can apply additional selection methods and constraints to narrow these alternatives down even more and identify the most suitable landscape scenario for their desired objectives. Figures 8.11 and 8.12 represent example landscape scenarios selected from the Pareto solutions with land covers and fertilization options geographically referenced. The OCMs ability to automate the process of evaluating multiple decision alternatives while applying advanced optimization methods allows users to make better informed decisions when deciding to increase production of biofuel feedstocks and with minimal effort.

These examples should only be used to demonstrate the effectiveness and potential of the optimal control model. It is advised that the tool be calibrated, validated, and applied to a data rich watershed prior to utilizing for actual decision-making.

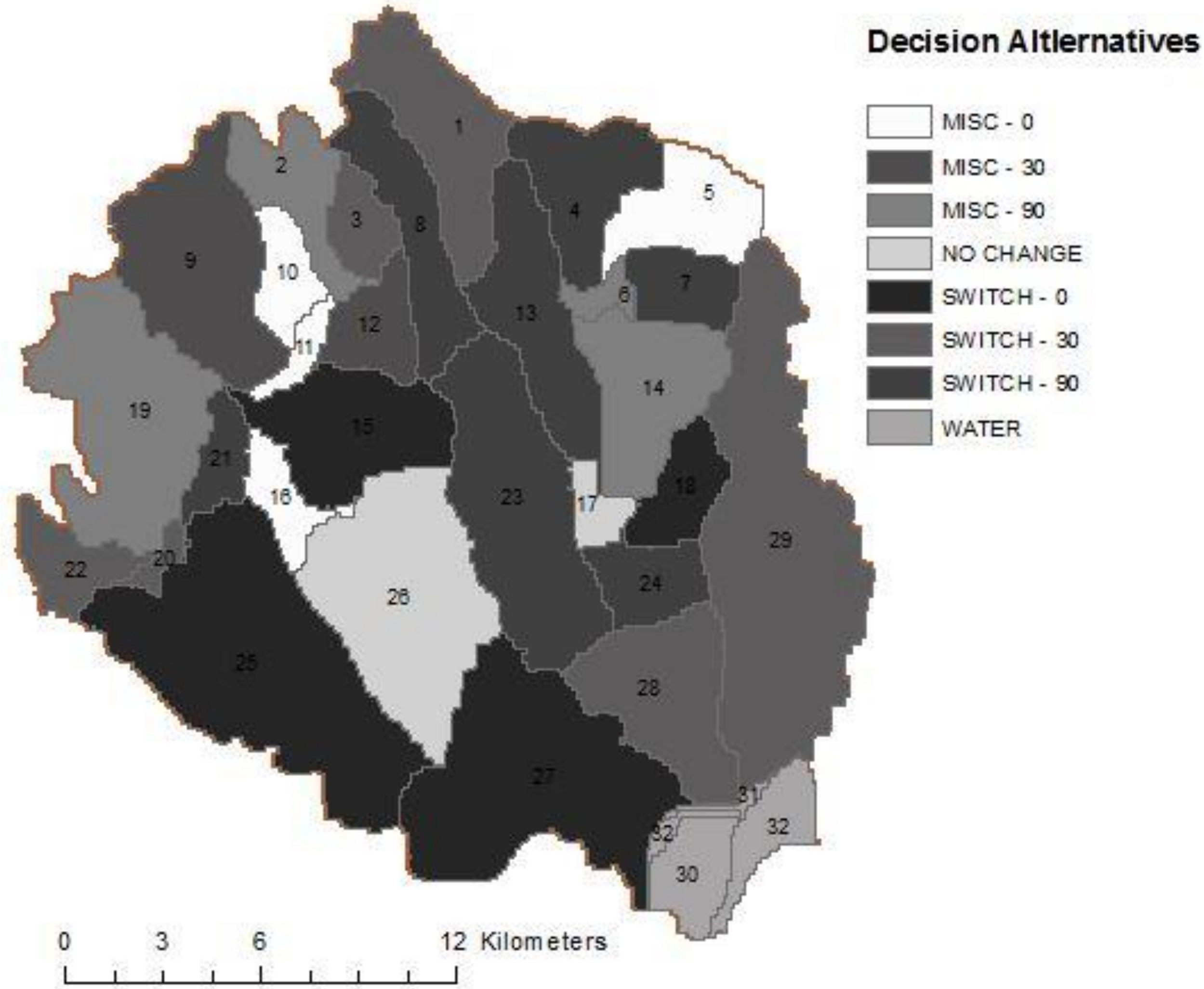

Figure 8.11: Example Landscape Scenario from Pareto Solutions for Lakefork Watershed.

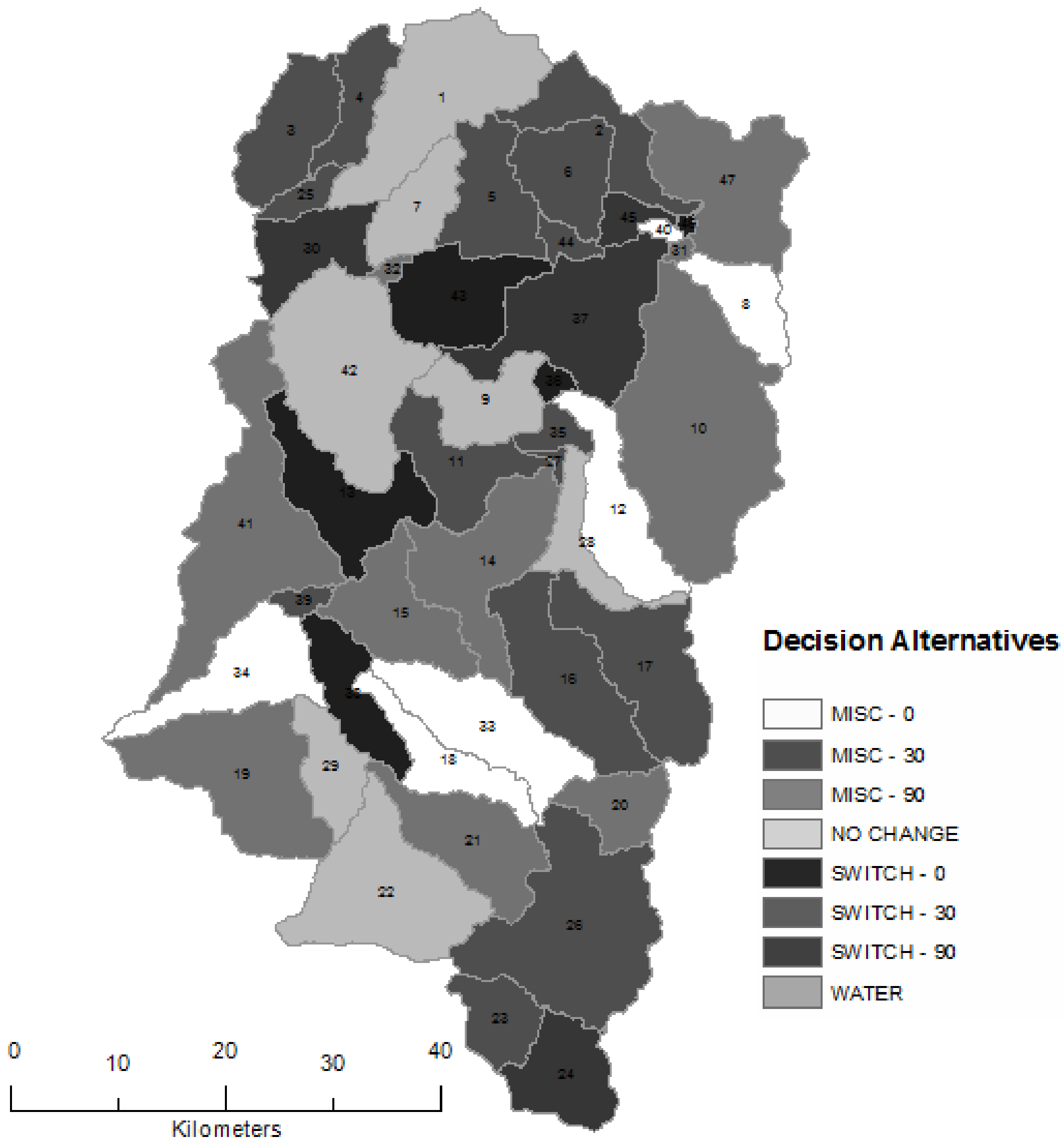

Figure 8.12: Example Landscape Scenario from Pareto Solutions for San Juan Watershed.

Chapter 9: Conclusion

It is already evident that the production and use of biofuels is increasing very rapidly and that the future supply and demand will depend largely on their competiveness with fossil fuels and global agricultural policies. Bioenergy markets have the potential to provide significant business opportunities, environmental services, and rural development. Biofuel feedstocks have the added advantage that they can be derived from a wide variety of biomass sources, on many types of lands, without threatening the global food and feed supply or biodiversity.

The future potential for biofuel feedstock production relies heavily on the availability of land. In addition, it is also important to consider the worldwide increasing demand for food, maintaining and protecting our environment, managing water reserves and soils, and many other sustainability requirements when choosing to increase production of biofuel feedstocks. Considering that the majority of this demand involves these complex related and often competing factors, it becomes difficult to assess and design a future biomass production strategies and scenarios. When exploring the potential long-term capacity of biomass resources at a global scale, many uncertainties and assumptions that can affect the availability of biomass must also be addressed. Crop yields and availability of land are dependent on many variables and related critical issues. These critical issues include, but are not limited to, competition for water resources, use of fertilizers and pesticides, effects of land-use, and competition with food and feed production. In order to thoroughly evaluate biofuel feedstock production on a national scale while taking into consideration the above factors, an extensive interdisciplinary research project would need to be conducted.

A less demanding scenario considers that bioenergy will contribute 20 to 50% of the world's future energy supply, and it is expected that about half would be derived from liquid biofuels (IEA, 2007). With these projections, the need for more available land is also going to be required. In order for these demands to be met, it will be highly important that developing countries develop bioenergy strategies and utilize perennial crops. The potential for developing countries to provide and export stable amounts of biomass-derived commodities creates a wealth of incentives and improved market access. This would help many rural communities in developing countries achieve greater socio-economic development. Many countries already have biomass resources available or have the potential to develop them, which could improve the accessibility and potential supply of biomass as an alternative energy option. Taking advantage of higher land use efficiencies by helping developing countries rationalize biomass production and agricultural techniques will be of key importance for biofuels in the future. Identifying and adapting biomass production and supply systems to regional conditions can further support increased demands.

With multiple projections of energy demands doubling and maybe even tripling in the next century, consequently it can be presumed that greenhouse gas concentrations in the atmosphere will also continue to rise rapidly mainly from CO2 emissions derived from fossil fuels. In order to mitigate the related impacts to climate change it is important to reduce greenhouse gas emissions. Sustainably producing energy from biomass can play a crucial role in reducing GHG emissions compared to fossil fuels. For example, biomass can be used to replace coal or co-fired with coal in power stations and has a high emission avoidance potential. However, these potentials vary depending on the efficiency of production and utilization, and

differ from country to country. With this in mind, it is important to carefully select strategies and policies in order to develop bioenergy systems optimally.

Chapter 10: Contributions and Recommendations

10.1 Contributions

The long-term goals of this research is to aid science based decisions for biofuel feedstock production in the context of promoting ecosystem service management and increasing water quality and healthy ecosystems at the watershed scale. This research provides significant contributions to the fields of environmental and water management, and biofuel feedstock production. More specifically the research has:

- Successfully developed a comprehensive optimal control model for the management of environmental objectives in a watershed with regards to biofuel feedstock production.
- Applied a hydrologic simulation model (SWAT) and a multi-objective evolutionary genetic algorithm to identify optimal agricultural landscape scenarios.
- Helped further the understanding of the environmental impacts of cultivating second-generation biofuel feedstocks and is intended to aid policy makers, water resource engineers and planners, when making decisions to increase feedstock production.
- Illustrated the potential to identify the spatial distributions of biofuel feedstock production within a watershed that will result in maximum biomass production, mitigate negative environmental impacts, and facilitate the displacement of biofuels derived from corn, reducing direct and indirect land use changes.

Identifying the trade-offs between the competing objectives will help to improve water quality and ecosystem services while meeting the demand of renewable energy production and cultivating biofuel feedstocks. With minimal modification, users can apply additional or

alternative objectives and constraints to the integrated modeling system to identify optimal land use scenarios in any watershed. An added benefit to the creation of this optimal control model is that its use does not have to be confined to biofuel feedstock simulations. It can be used to simulate watershed impacts and identify optimal landscapes with a variety of land covers and crops.

10.2 Recommendations

The integrated modeling system has proven to be a valuable decision support tool for evaluating biofuel feedstock production and watershed impacts. However, it can still be further expanded to include additional features and modeling functions. Potential additions are listed below:

For illustration purposes, the decision variables evaluated in this research included land covers of switchgrass and miscanthus, along with various fertilization application amounts. However, the optimal control model can be expanded to evaluate more decision variables such as other feedstocks and best management practices.

With a plethora of feedstocks to choose from, the optimal control model can be used to evaluate endless combinations of potential agricultural landscapes, from grains and grasses, woods, dedicated energy crops, to oilseeds. Popular established grains and grasses include corn, sugarcane, wheat, rice, and cassava. Less attention has gone to woods, mainly due to the fact that the technology for conversion is still developing but we may slowly begin to see potential in these feedstocks as fermentation systems improve. Popular feedstocks from the wood group include poplar, eucalyptus, willow, white pine, and yellow pine. The oilseed family consists of soybeans, oil palm, rapeseed, coconut, and castor. Novel energy feedstocks include jatropha, switchgrass sweet sorghum, camelina, and miscanthus. New research is continually surfacing on

the viability of these biofuel feedstocks and their potential for meeting energy demands while possibly creating greenhouse gas savings and carbon sequestration potentials.

SWAT incorporates an extensive plant database file. This database file includes plant growth parameters for many common species. However, it can easily be edited and modified to accommodate specific plant parameters or to include new plants not currently in the database. With the increased research in advanced biofuel crop production, these plant growth parameters are becoming more available through the literature.

Examples of land management practices that can be modeled are wetlands, buffer strips, detention ponds, along with various irrigation and harvesting practices, which all may have potential to alleviate non-point source pollution and promote water conservation.

The primary goal of this research was to develop a framework that can be further modified and built on in order to assess the impacts of biofuel feedstock production on a given system. Critical to this assessment is the implications of land management practices and cropping systems. SWAT is a text based model that utilizes multiple numerous text files that are read into the model to simulate these interactions and physical processes. The primary file that allows us to visualize and simulate potential land management scenarios is the HRU management file. The HRU management file is created for each HRU or subbasin and can be found in the Input file directory labeled with the HRU number and ending in .mgt. This file contains all the management information for that particular HRU and can be written to specify and schedule various planting, harvesting, irrigation, fertilizer and pesticide application, and tillage operations.

Additional modeling components can be integrated to evaluate more objectives. For example, the Policy Analysis System (POLYSYS) model can be integrated to include economic

simulations when optimizing land cover options. POLYSYS is a modular economic simulation modeling system that uses crop demands and market prices to evaluate agricultural planning decisions. By incorporating POLYSYS we can include cost of production and expected returns as objectives when deciding what crop systems to develop in a particular region. Another option would be to integrate the Farm Energy Analysis Tool (FEAT) or The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) model, in order to evaluate energy and greenhouse gas emissions. FEAT is a database model that can be used to estimate energy use and GHG emissions from different agricultural systems and crop production. The GREET model evaluates emission and primary resource consumption related to production and use of biofuels. Adding a Water Analysis Tool for Energy Resources (WATER) component to the optimal control model may also provide beneficial water footprint information. The WATER tool was developed to assess water use and quality specifically for the production of biofuels. The tool estimates water demand and availability throughout all biofuel production stages. The Daily Century Model (DAYCENT) is capable of simulating N and C fluxes from soil to atmosphere. DAYCENT uses similar input data as SWAT, daily weather, soil properties, and land management information when calculating N and C fluxes.

There are two ways to incorporate the previous components to the optimal control model using MATLAB. The first method would require the model be integrated similar to the approach taken in this research. Additional models can be used by the algorithm to evaluate the state variables and landscape scenarios for additional objectives of interest. After calling upon SWAT to simulate and identifying the outputs of interest, the algorithm will move to the next component and run it and extract the required outputs. Then the multi objective evolutionary algorithm can continue its selection and evaluation processes according to the prescribed objectives. Another

potential method for integration would be to develop an external database of values by using the aforementioned models individually and separate of the optimal control model. After manually running each tool separately for various scenarios and developing a general database of corresponding values, it can be referenced by the algorithm to identify values for that particular objective.

In order to realize the full potential as a decision support tool and confidently accept the suggested optimal landscape scenarios for feedstock production, it would be advised to apply this integrated modeling approach to a data-rich watershed. After identifying an appropriate watershed of interest, spatial sensitivity and feasibility analyses should be carried out. The sensitivity of SWAT to spatial scale can be determined by running SWAT with multiple discretization levels for watershed delineation. SWAT sets discretization levels by defining the critical source area (CSA). Critical source area is the minimum drainage area required to form the origin of a stream and determines the amount of detail in the stream network and the size and number of subbasins for the watershed. To identify the sensitivity of outputs, various discretization levels must be simulated and compared to observed data from gauging stations. A spatial sensitivity analysis will help to reveal the consequences of aggregating inputs and parameters on the simulated outputs. In order to provide accurate and improved model predictions, various discretization levels should be evaluated. Finally, it would be necessary to calibrate SWAT prior to running optimal control model to ensure reliability.

SWAT can be applied to any watershed with various physical characteristics. In order to match model behavior with the watershed being examined, model parameters must be adjusted accordingly to match model outputs with observed data.

Finally, the development of a user-friendly graphical user interface is in progress and will allow users to easily manipulate parameters and decision alternatives of the optimal control model.

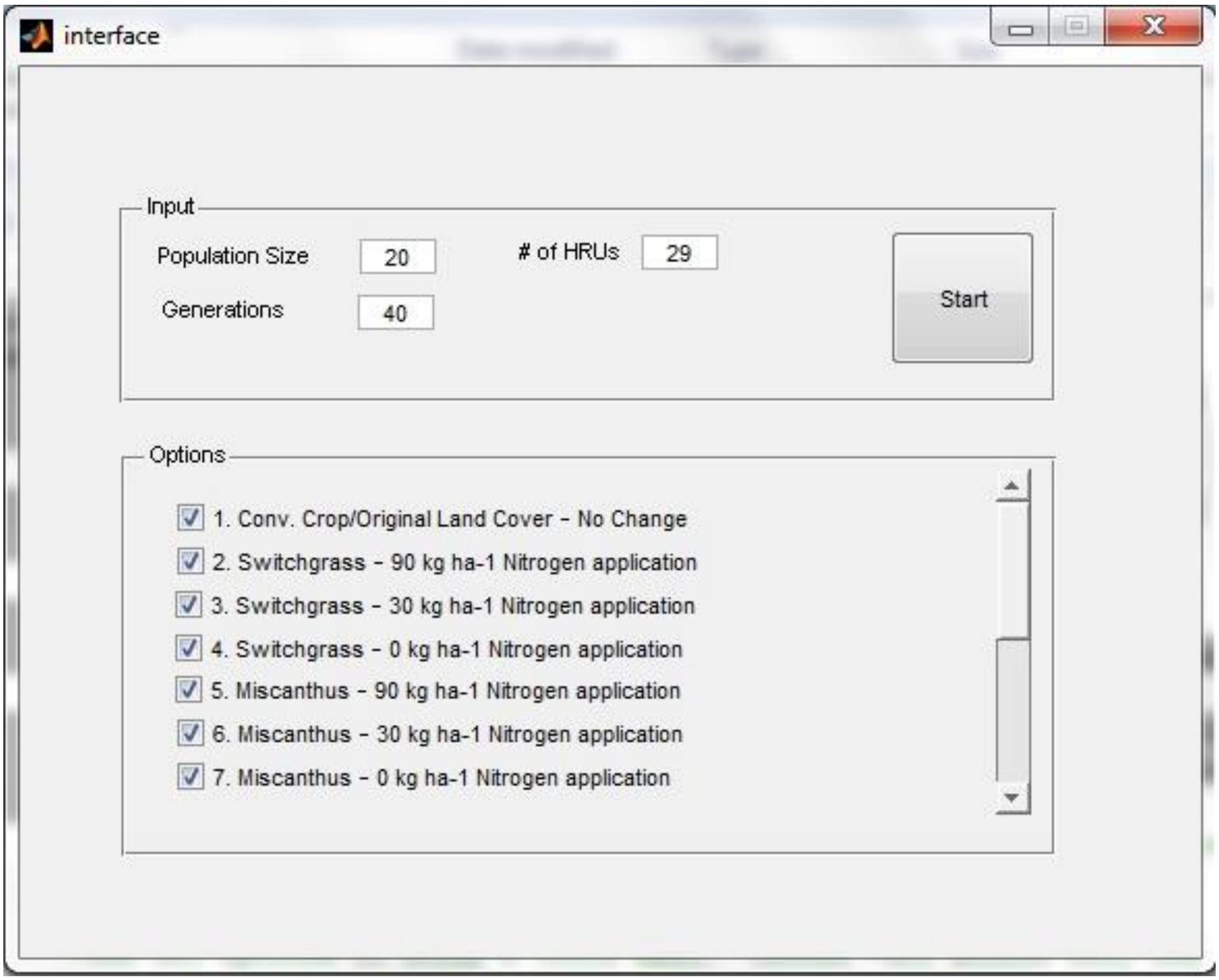

Figure 10.1. Graphical User Interface (GUI) for Optimal Control Model

References

Arnold, J. G., Srinivasan, R., Muttiah, R. S., Williams, J R. (1998). Large area hydrologic modeling and assessment - Part 1: Model development. *Journal of the American Water Resources Association 34(1),* 73-89.

Artita, K. S., Kaini, P., & Nicklow, J. W. (2007). Generating alternative watershed-scale BMP designs with evolutionary algorithms. In *Proc., 2007 World Environmental and Water Resources Congress*. Reston, Va.: ASCE.

Biomass Research and Development Board. (2008). Increasing Feedstock Production for Biofuels, Economic Drivers, Environmental Implications, and the Role of Research. U.S. Department of Agriculture.

Blanco-Canqui, H. (2010). Energy crops and their implications on soil and environment. *Agronomy Journal*, *102*(2), 403-419.

Blasco, X., Herrero, J. M., Sanchis, J., & Martínez, M. (2008). A new graphical visualization of n-dimensional Pareto front for decision-making in multiobjective optimization. *Information Sciences*, *178*(20), 3908-3924.

Charnes, A., Cooper, W. W., & Ferguson, R. O. (1955). Optimal estimation of executive compensation by linear programming. *Management science*, *1*(2), 138-151.

Coyle, W. (2007). The future of biofuels. *Economic Research Service, Washington, DC*.

Coyle, W. T. (2010). *Next-generation biofuels: near-term challenges and implications for agriculture*. DIANE Publishing.

Cross, H. Z., & Zuber, M. S. (1972). Prediction of flowering dates in maize based on different methods of estimating thermal units. *Agronomy journal*, *64*(3), 351-355.

Das, I., & Dennis, J. E. (1997). A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural optimization*, *14*(1), 63-69.

Deb, K., & Beyer, H. G. (2001). Self-adaptive genetic algorithms with simulated binary crossover. *Evolutionary Computation*, *9*(2), 197-221.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, *6*(2), 182-197.

Demirbas, A. (2009). Biofuels securing the planet's future energy needs. *Energy Conversion and Management*, *50*(9), 2239-2249.

Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *26*(1), 29-41.

Eberhart, R. C., and Kennedy, J. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, Nov 25–Dec 1*. *Edu Sci Technol 14*:1–19.

Engel, B., Chaubey, I., Thomas, M., Saraswat, D., Murphy, P., & Bhaduri, B. (2010). Biofuels and water quality: challenges and opportunities for simulation modeling. *Biofuels*, *1*(3), 463-477.

Fishburn, P. C. (1967). Letter to the Editor—Additive Utilities with Incomplete Product Sets: Application to Priorities and Assignments. *Operations Research*, *15*(3), 537-542.

Fishburn, P.C. (1974). Convex stochastic dominance with continuous distribution functions. *Journal of Economic Theory 7*, 143–158.

Gassman, P., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). *The soil and water assessment tool: historical development, applications, and future research directions*. Center for Agricultural and Rural Development, Iowa State University.

Gitau, M. W., Veith, T. L., & Gburek, W. J. (2004). Farm-level optimization of BMP placement for cost-effective pollution reduction. *Transactions of the ASAE*, *47*(6), 1923-1931.

Gitau, M. W., Veith, T. L., Gburek, W. J., & Jarrett, A. R. (2006). Watershed level best management practice selection and placement in the Town Brook watershed, New York. *JAWRA Journal of the American Water Resources Association*, *42*(6), 1565-1581.

Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers & Operations Research*, *13*(5), 533-549.

Hamelinck, C. N., & Faaij, A. P. (2006). Outlook for advanced biofuels. *Energy Policy*, *34*(17), 3268-3283.

Haupt, R.L. & Haupt, S.E. (1998). Practical Genetic Algorithms. John Wiley & Sons, New York.

Hertz, A., Taillard, E., & De Werra, D. (1995). A tutorial on tabu search. In *Proc. of Giornate di Lavoro AIRO* (Vol. 95, pp. 13-24).

IEA Bioenergy Executive Committee (2007). Potential Contribution of Bioenergy to the World's Future Energy Demand.

IPCC AR4 WG2 (2007), Parry, M.L.; Canziani, O.F.; Palutikof, J.P.; van der Linden, P.J.; and Hanson, C.E., ed., Climate Change 2007: Impacts, Adaptation and Vulnerability, Contribution of Working Group II to the *Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

Jury, W.A., Gardner, W.R., and Gardner, W.H. 1991. Soil Physics, 5th Edition. John Wiley & Sons, Inc. New York, N.Y.

Kaini, P., Artita, K. S., & Nicklow, J. W. (2008). Designing BMPs at a watershed scale using SWAT and a genetic algorithm. In *World Environmental and Water Resources Congress*.

Kaini, P., Artita, K., & Nicklow, J. W. (2007). Evaluating optimal detention pond locations at a watershed scale. In *World Environment and Water Res. Congress, Tampa, FL.*

Kaini, P., Artita, K., & Nicklow, J. W. (2009). Generating different scenarios of BMP designs in a watershed scale by combining NSGA-II with SWAT. In *World Environmental and Water Resources Congress 2009: Great Rivers* (pp. 1-9).

Kaini, P., Artita, K., & Nicklow, J. W. (2012). Optimizing structural best management practices using SWAT and genetic algorithm to improve water quality goals. *Water resources management*, *26*(7), 1827-1845.

Knowles, J. D., & Corne, D. W. (2000). Approximating the nondominated front using the Pareto archived evolution strategy. *Evolutionary computation*, *8*(2), 149-172.

Ling, T., Eheart, W., Cai, X., Miguez, F. (2010). Modeling miscanthus in the Soil and Water Assessment Tool (SWAT) to simulate its water quality effects as a bioenergy crop. *Environmental Science & Technology,* 44(18), 7138-7144.

Georgy, M. E., Chang, L. M., & Zhang, L. (2005). Utility-function model for engineering performance assessment. *Journal of construction engineering and management*, *131*(5), 558-568.

Malcolm, S., M. Aillery, and M. Weinberg. 2009. Ethanol and a changing landscape, ERR-86, U.S. Department of Agriculture, Economic Research Service.

Maringanti, C., Chaubey, I., Arabi, M., & Engel, B. (2008). A multi-objective optimization tool for the selection and placement of BMPs for pesticide control. *Hydrology and Earth System Sciences Discussions*, *5*(4), 1821-1862.

Maringanti, C., Chaubey, I., Arabi, M., & Engel, B. (2011). Application of a multi-objective optimization method to provide least cost alternatives for NPS pollution control. *Environmental management*, *48*(3), 448-461.

Marshall, E., Caswell, M., Malcolm, S., et al. (2011). Measuring the Indirect Land-Use Change Associated With Increased Biofuel Feedstock Production: A Review of Modeling Efforts: Report to Congress. *Administrative Publication No. (AP-054)* 56pp.

McElroy, A. D. (1976). *Loading functions for assessment of water pollution from nonpoint sources* (Vol. 1). US Environmental Protection Agency, Office of Research and Development, [Office of Air, Land, and Water Use].

Menzel, R.G. (1980). Enrichment ratios for water quality modeling. In *W.G. Knisel (ed.) CREAMS, A Field Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems*. U.S. Dept. Agric. Conserv. Res. Rept. No. 26.

Messac, A. (1996). Physical programming-Effective optimization for computational design. *AIAA journal*, *34*(1), 149-158.

Miguez, F., Villamil, M. (2008). Meta-Analysis of the effects of management factors on miscanthus x giganteus growth and biomass production. *Agricultural and Forest Meteorology*, 148, 1280-1292.

Monti, A., Fazio, S., Lychnaras, V., Soldatos, P., & Venturi, G. (2007). A full economic analysis of switchgrass under different scenarios in Italy estimated by BEE model. *Biomass and Bioenergy*, *31*(4), 177-185.

Muleta, M., & Nicklow, J. (2002). Evolutionary algorithms for multiobjective evaluation of watershed management decisions. *Journal of Hydroinformatics*, *4*, 83-97.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., (2011). Soil and Water Assessment Tool Theoretical Documentation, Version 2009. Grassland, Soil and Water Research Laboratory, ARS. Texas Water Resources Institute Technical Report No. 406

Nickerson, C., Ebel, R., Borchers, A., & Carriazo, F. (2011). *Major uses of land in the United States, 2007*. US Department of Agriculture, Economic Research Service.

Nicklow, J. W. (2000). Discrete-time optimal control for water resources engineering and management. *Water international*, *25*(1), 89-95.

Osyczka, A. (1984). Multicriterion optimisation in engineering. Halsted Press, New York, NY.

Pike, Ralph W. (1986). Optimization for engineering systems. Van Nostrand Reinhold Company, New York.

Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. *Swarm intelligence*, *1*(1), 33-57.

Robertson, G. P., & Swinton, S. M. (2005). Reconciling agricultural productivity and environmental integrity: a grand challenge for agriculture. *Frontiers in Ecology and the Environment*, *3*(1), 38-46.

Saga, K., Yokoyama, S., Imou, K., & Kaizu, Y. (2008). A comparative study of the effect of CO2 emission reduction by several bioenergy production systems. *Int Energy J*, *9*, 53-60.

Schwefel, H. P. (2000). Advantages (and disadvantages) of evolutionary computation over other approaches. *Evolutionary computation*, *1*, 20-22.

Sharpley, A. N. & Williams, J. R. (1990). EPIC-Erosion Productivity Impact Calculator, 1. Model Documentation. *U.S. Department of Agriculture, Agricultural Research Service, Tech. Bull. 1768.*

Singh, V. P. (1988). Hydrologic systems. Volume I. Rainfall-Runoff Modeling, Prentice Hall, NJ.

Singh, V. P., & Frevert, D. K. (Eds.). (2005). *Watershed models*. CRC Press.

Singh, V. P., & Woolhiser, D. A. (2002). Mathematical modeling of watershed hydrology. *Journal of hydrologic engineering*, *7*(4), 270-292.

Soil Conservation Science (1972). Section 4: Hydrology. In National Engineering Handbook. Soil Conservation Service, Washington, D.C.

Srinivas, N., & Deb, K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation*, *2*(3), 221-248.

Taboada, H., & Coit, D. (2008). MOEA-DAP: a new multiple objective evolutionary algorithm for solving design allocation problems. Department of Industrial and Systems Engineering, Rutgers University. Piscataway, NJ.

Tallis, H., & Polasky, S. (2009). Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. *Annals of the New York Academy of Sciences*, *1162*(1), 265-283.

Tewfik, S.R. 2004. Biomass utilization facilities and biomass processing technologies. *Energy Edu Sci Technology*., *14*: 1-19.

U.S. Congress. (2007). Energy Independence and Security Act of 2007. HR 6. 110th Congress. 1st session. December 2007.

United States Department of Agriculture, (2010). A USDA Regional Roadmap to Meeting the Biofuels Goals of the Renewable Fuels Standard by 2022. Biofuels Strategic Production Report. June 23, 2010.

Vitousek, P.M., Mooney, H.A., Lubchenko, J., and Melillo, J.M. (1997). Human domination of earth's ecosystems. *Science, 277*: 494-499.

Williams, J.R. (1975). "Sediment routing for agricultural watersheds." Water Resour. Bull., 11(5): 965-974.

Williams, J.R. (1995). The EPIC model. In V.P. Singh (ed). *Computer Models of Watershed Hydrology*, Water Resources Publications, Highlands Ranch, CO.

Williams, J.R. and R.W. Hann (1978). Optimal operation of large agricultural watersheds with water quality constraints. Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Wischmeier, W.H. and Smith, D.D. (1965). Predicting rainfall erosion losses: A guide to conservation planning. Agriculture Handbook 282, USDA-ARS.

Wong, T.S., Koh, X.C., (2008). Which model type is best for deterministic rainfall-runoff modeling?. Water Resources Research Progress, 241-260.

Zhang, X., Srinivasan, R., & Van Liew, M. (2008). Multi-site calibration of the SWAT model for hydrologic modeling. *Transactions of the ASABE*, *51*(6), 2039-2049.

Zhang, X., Srinivasan, R., & Liew, M. V. (2010). On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. *Hydrological Processes*, *24*(8), 955-969.

Zhou, A., Qu, B. Y., Li, H., Zhao, S. Z., Suganthan, P. N., & Zhang, Q. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, *1*(1), 32-49.

Zitzler, E., & Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *Evolutionary Computation, IEEE Transactions on*, *3*(4), 257-271.

Zitzler, E., Laumanns, M. and Thiele, L. (2001). SPEA2: Improving the strength pareto evolutionary algorithm. TIK-Report 103, Computer Engineering and Network Laboratory (TIK), Swiss Federal Institute of Tech., Zurich, Switzerland.

Vita

Rodney Vance was born in El Paso, Texas and is the third child of Jerry Vance and Juanita Garcia Vance. He graduated from Andress High School in El Paso, Texas in 1999 and continued his education at the University of Texas at El Paso. He is the first and only of his family to attend college. While pursuing a Bachelor's of Science in Biology, he worked for the Department of Defense, Department of the Army at Ft. Bliss, Texas. After receiving his degree he conducted research for Texas A&M's Agricultural Extension and taught high school science at Cesar Chavez Academy. While taking courses towards his Master's in Education he decided to transfer to the PhD program in Environmental Science and Engineering. Rodney has had the opportunity to intern in Washington D.C. at USDA's National Institute of Food and Agriculture and with the Office of the Chief Scientist. He currently resides in Arlington, Virginia and works at NIFA's Institute of Bioenergy, Climate, and Environment.

Permanent address: 2701 S. Adams St., #405
Arlington, VA, 22206
rvance@miners.utep.edu

This dissertation was typed by Rodney W. Vance.