UC Berkeley
UC Berkeley Electronic Theses and Dissertations

**Title**
Geospatial strategies to optimize placement of advanced bioenergy crops in marginal landscapes

**Permalink**
https://escholarship.org/uc/item/2328j6qf

**Author**
Lewis, Sarah Marie

**Publication Date**
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

# Geospatial strategies to optimize placement of advanced bioenergy crops in marginal landscapes

By

Sarah Marie Lewis

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy, and Management

In the

Graduate Division

Of the

University of California, Berkeley

Committee in charge:

Professor Maggi Kelly, Chair
Professor Greg Biging
Professor John Radke

Fall 2013

Geospatial strategies to optimize placement of advanced bioenergy crops in marginal landscapes

Copyright © 2013

By

Sarah Marie Lewis

Abstract

Geospatial strategies to optimize placement of advanced bioenergy crops

in marginal landscapes

by

Sarah Marie Lewis

Doctor of Philosophy in Environmental Science, Policy and Management

University of California, Berkeley

Professor Maggi Kelly, Chair

As energy policies around the world progressively mandate increases in bioenergy production, critics point out that the amount of land required to meet these demands might come at the cost of food production or environmental protection. To ensure sustainable practices, planting energy crops on *marginal land* is frequently proposed as a solution to this land use dilemma in order to meet mandates while avoiding negative externalities. To this end, scientists are actively engaged in using Geographic Information Systems (GIS) to spatially quantify marginal lands available for this purpose. Furthermore, both presently and in the future of a changing climate, drought-tolerant species are being targeted for bioenergy production in part because of their ability to grow on arid lands marginal for conventional bioenergy crops. Yet even though the term "marginal land" has become so entwined with discussions surrounding bioenergy, its definition is in fact malleable – easily shaped to fit research needs. This research investigates how marginal lands are commonly viewed in a GIS framework and presents an innovative and flexible model that optimizes the placement of drought-tolerant bioenergy crops in marginal landscapes in the United States.

In Chapter 2, I present the first work to evaluate a suite of studies that use GIS to map marginal lands available for bioenergy production. I conclude that despite similar goals, the current literature exhibits considerable differences in working definitions, including model framework, data inputs, scale, and treatment of uncertainty that limit the potential for comparison across studies. When assessing modeling framework, I find that the majority of studies use simple linear GIS overlays with distinct thresholds. I also find that the majority of studies mention uncertainty only qualitatively, or make no mention of uncertainty or model sensitivity at all. These results aid in the development of an innovative and flexible site suitability model for drought-tolerant bioenergy crops.

Drought-tolerant bioenergy crops are of particular interest in the US because of their ability to grow on land not suitable for conventional agricultural production. In Chapters 3 and 4, I focus on the development of a site suitability model rooted in fuzzy logic for several types of drought-tolerant bioenergy crops – a modified variety of switchgrass as well as two species within the

*Agave* genus. In addition to the primary fuzzy suitability model, each chapter explores the sensitivity of the fuzzy model in different ways, including modifications to overlay functions and input parameters, respectively.

In Chapter 3, I demonstrate that incorporating fuzzy logic into the suitability model provides a more progressive suitability index, and that the Gamma overlay function best synthesizes tradeoffs between multiple criteria. The results highlight a range of land suitable for drought-tolerant switchgrass in the Great Plain States. I also identify the dryness index threshold at which land area is maximized for this crop. These results have very direct practical applications for plant science and will help guide research on drought-tolerant varieties of switchgrass.

Finally, in Chapter 4, I present the first evidence to demonstrate that the area suitable for growing *Agave* as a bioenergy feedstock in the Southwestern US is sufficient to contribute to domestic renewable energy needs. At the state-level, Arizona and Texas show the highest potential energy benefit. The sensitivity analysis of model parameters reveals that suitability for *Agave* is most restricted by climatic variables as opposed to physical soil factors; particularly, minimum temperature in winter months is the most influential criterion restricting land area suitable for *Agave* cultivation in the US. I conclude that *Agave* would not be a major competitor to other energy feedstocks, but rather serve a parallel benefit as a renewable energy crop on lands unsuitable for conventional bioenergy production. *Agave* could thereby serve as a complementary contributor to the country's domestic renewable energy needs.

# Table of Contents

List of Figures ..... iii

List of Tables ..... iv

Acknowledgements ..... v

**Chapter One: Optimizing the placement of bioenergy crops in marginal landscapes** ..... 1
*Renewable energy and bioenergy mandates* ..... *1*
*Marginal lands* ..... *2*
*Site suitability mapping* ..... *3*

**Chapter Two: Mapping the potential for biofuel production on marginal lands: differences in definitions, data and models across scales** ..... 7
Introduction ..... 7
Defining marginal land ..... 9
A comparative case study: mapping marginal lands for bioenergy in China ..... 13
Mapping potential for bioenergy ..... 17
*Model selection* ..... *17*
*Input data* ..... *19*
*Biophysical data inputs* ..... *19*
*Socio-economic data inputs* ..... *23*
*Other data* ..... *23*
*Choice of data threshold* ..... *23*
*Choice of thematic class* ..... *24*
*Issues of data representation* ..... *25*
*Issues of scale* ..... *25*
*Issues of uncertainty* ..... *26*
Discussion ..... 28
*Modeling framework* ..... *28*
*Data availability* ..... *29*
*Scale of analysis* ..... *29*
*Uncertainty* ..... *29*
Conclusions ..... 30

**Chapter Three: A fuzzy logic-based spatial suitability model for drought-tolerant switchgrass in the United States** ..... 31
Introduction ..... 31
Methods ..... 34
*National Suitability Model* ..... *34*
*Primary Suitability Criteria* ..... *34*
*Restrictive Layers* ..... *35*

*Resampling* ... *35*
*Fuzzy Suitability Model* ... *35*
*Standardized suitability criteria* ... *35*
*Model Sensitivity to Fuzzy Overlay* ... *36*
*Focal Area Dryness Index* ... *36*
Results ... 37
*Suitability* ... *37*
*Fuzzy overlay sensitivity* ... *39*
*Results from Focus Area* ... *40*
Discussion ... 41
Conclusion ... 42

**Chapter Four: Fuzzy GIS-based multi-criteria evaluation for U.S. *Agave* production as a bioenergy feedstock** ... 43
Introduction ... 43
Methods ... 43
*Species Selection* ... *45*
*Suitability Model* ... *45*
*Selection Criteria and Data Pre-processing* ... *46*
*Fuzzy Membership Transformations* ... *47*
*Fuzzy Overlay Sub-model* ... *49*
*Fuzzy Overlay Final Model* ... *49*
*Constraints* ... *49*
*Sensitivity Analysis* ... *49*
*Scenarios* ... *51*
*Infrastructure* ... *51*
*Energy Estimates* ... *51*
Results ... 52
*Suitability Model Results* ... *52*
*Sensitivity Analysis Results* ... *53*
*Scenario Results* ... *55*
*Proximity to Infrastructure* ... *55*
*Yield Estimates* ... *58*
*State-level Demands for Renewable Fuels & Electricity* ... *60*
*Other feedstocks* ... *61*
Discussion ... 61
Conclusion ... 61

**Chapter Five: Conclusions and directions for future research** ... 65
*Directions for future research* ... *67*

Cited Literature ... 69

# List of Figures

Figure 3-1. ..... 37
Figure 3-2. ..... 38
Figure 3-3. ..... 40
Figure 3-4. ..... 41
Figure 4-1. ..... 46
Figure 4-2. ..... 48
Figure 4-3. ..... 53
Figure 4-4. ..... 54
Figure 4-5 ..... 56

# List of Tables

Table 2-1. .....11
Table 2-2.. .....16
Table 2-3. .....22
Table 4-1.. .....47
Table 4-2. .....51
Table 4-3.. .....55
Table 4-4 .....57
Table 4-5. .....58
Table 4-6. .....59
Table 4-7. .....59
Table 4-8. .....60

# Acknowledgements

I would like to give my sincere thanks to my advisor, Maggi Kelly, for giving me the opportunity to perform my doctorate work at UC Berkeley. Thank you for your years of generous advice, motivation, and shared passion for all things geospatial. I also extend my deepest gratitude to my dissertation committee, Greg Biging and John Radke, for sharing their sound wisdom and guidance.

I am very grateful for the support of Lawrence Berkeley Labs and for my coworkers there - Larry Dale, Bill Morrow and Gary Fitts, Stephen Gross, and Axel Visel. I would also like to give my sincere gratitude to Peter Berck and the Energy Biosciences Institute for their support for many semesters of this PhD.

I also extend my appreciation to the mentors who have given me the inspiration and strengths that have brought me here today, including Jennifer Walwark Dailey, Kate Dean, Tom Gillespie, Gretchen Coffman, and Larry Smith.

This dissertation would not have been possible without the help of many individuals, including family, friends and fellow students. I would like to thank all members of the Kelly Lab group, both past and present, who have helped me to grow and become a better scholar. Thank you to the staff of the GIF, especially Kevin Koy for your mentorship and support during my time at Berkeley.

I would also like to thank the many members of my team outside of Berkeley who have supported me throughout this process – I could not have done it without you. To my parents, Alice and Si Lewis – thank you Mom for your endless support and countless hours copyediting my papers, reports, and ultimately my dissertation drafts, and thank you Dad for inspiring both intellectual debate and laughter throughout these five years. To my sister Emily, thank you for your hugs and calming spirit, and for Marshal, whose music provided the best inspiration for my writing. I also send my deepest gratitude to the Horton and MacDonald Families who have provided me endless encouragement (and nourishment) along the way. Finally thank you Graeme and Honey Bee. Your love, support, and friendship mean the world to me. I love you with all of my heart.

# Chapter One

## Optimizing the placement of bioenergy crops in marginal landscapes

Land availability is one of the most significant challenges facing renewable bioenergy production today. Planting energy crops on *marginal land* is frequently proposed as a sustainable solution, and to this end scientists are actively engaged in quantifying lands available for such a purpose. This research investigates how marginal lands are viewed in a Geographic Information Systems (GIS) framework and presents an innovative and flexible model that optimizes the placement of drought-tolerant bioenergy crops in marginal landscapes in the United States. Prior to this research, the common approach to mapping marginal lands for bioenergy production was to employ simple GIS overlay models that impose distinct thresholds between lands that are suitable for bioenergy crops (marginal lands) and those that are not. In addition, most studies only qualitatively mention uncertainty in their mapped results, or give no mention of uncertainty or model sensitivity at all. To address this problem, the specific objectives of this dissertation are to 1) understand the current trends in mapping marginal lands; 2) use this knowledge to develop a site suitability model that incorporates uncertainty and sensitivity; and finally 3) determine if drought-tolerant bioenergy crops can be cultivated in the US on lands unsuitable for conventional energy crop production and contribute to domestic renewable energy demand. The motivation for this research originates from a need to determine where to plant bioenergy feedstocks to meet energy mandates while ensuring sustainable food production and environmental protection.

In this introductory chapter I examine the context of this dissertation: I review the state of bioenergy mandates and research, explain the attention given to marginal lands, and finally highlight the potential for site suitability analyses that incorporate fuzzy logic to synthesize tradeoffs between multiple criteria with continuous degrees of suitability. As follows, each chapter of my dissertation focuses on the use of GIS to spatially quantify land area available for bioenergy crops production.

*Renewable energy and bioenergy mandates*

As global energy demand increases, countries around the world are recognizing the potential for biomass to be a significant domestic energy resource with the ability to simultaneously reduce greenhouse gas emissions and increase energy security. Because of their ability to provide both environmental and economic benefits, bioenergy crops promise to become a growing sector of the renewable energy market (Nakicenovic *et al.*, 2000; Somerville *et al.*, 2010; Youngs & Somerville, 2012). To this end, policy makers are progressively mandating increases in energy production. For example, in the United States, the Renewable Fuel Standard (RFS2) mandates

that by 2020 16 billion gallons of cellulosic biofuels contribute to the country's renewable biofuel production (Sissine, 2007).

However, meeting these mandates will require a diversity of bioenergy resources – preferably second generation non-food energy crops that have been shown to be more environmentally sustainable than their first generation food-based counterparts, if they can be designed and managed appropriately (Tilman *et al.*, 2006, 2009; Dale *et al.*, 2010). Second generation energy crops consist of dedicated agricultural crops of herbaceous or arborous species, including annual crops, perennial grasses, and short rotation forestry. The ability of these species to simultaneously sequester carbon and reduce greenhouse gas emissions by replacing conventional fuel and energy production make them especially desirable as a renewable energy source.

In addition to concerns regarding greenhouse gas reduction, when considering a future with a changing climate an extra concern is that the frequency of extreme weather and drought has increased in the U.S. over the last decade (Regonda *et al.*, 2005; Wang *et al.*, 2010; Mo, 2011). While some researchers estimate that droughts are decreasing in severity and duration (Mishra *et al.*, 2010), most climate scientists report that advancing climate change is expected to worsen droughts throughout the U.S., especially in the Midwest, which is the nation's largest producer of food crops (Friedel, 2011; Romm, 2011; Wehner *et al.*, 2011). These changes make it increasingly important to grow drought-tolerant bioenergy crops on dryer land areas not suited for conventional agricultural production. Species like *Agave* and a drought-tolerant variety of switchgrass have recently been attracting attention as a potential complement to other bioenergy feedstocks in areas where those crops do not grow well (Somerville *et al.*, 2010).

However, the sustainability of biofuels depends on where they are produced, processed and used (Tilman *et al.*, 2009), and the demands for production often come ahead of reliable estimates of land area available for such a purpose (Robertson *et al.*, 2008; Solomon, 2010). Consequently, land availability is a major limitation to bioenergy production, and critics point out that these crops are likely to use land needed for other purposes, particularly those used food production or natural habitats (Dale *et al.*, 2010). Therefore some have argued that the sustainable goals of bioenergy production could be counteracted through instigation of both direct and indirect land use change (Kim & Dale, 2011). This *bioenergy land use dilemma* is challenging scientists to determine where exactly to plant biofuels to meet mandates, but also to ensure sustainable food production and environmental protection (Cassman & Liska, 2007; Elobeid & Hart, 2007; Lam *et al.*, 2009).

*Marginal lands*

Many have proposed the use of marginal lands as a sustainable solution to this bioenergy land use dilemma, in part because second-generation biofuels have been documented to grow well on non-prime agricultural land. In fact research has shown that bioenergy has the potential to be carbon neutral, or even carbon negative, if grown on marginal land (Tilman *et al.*, 2006; Gelfand *et al.*, 2013a). For this reason, marginal lands are overwhelmingly referred to as the solution to the negative externalities that may be brought on by bioenergy production.

In fact, the term "marginal land" has become so intertwined with discussions surrounding bioenergy that its definition might be assumed to be specific and certain. However the concept of marginal land is in fact malleable – and has varied across time, space and discipline to meet multiple management goals (Kang *et al.*, 2013a). In the current context of bioenergy production, marginal land is commonly assumed to be low quality land not suitable for prime agricultural production, yet it still has the potential to support bioenergy crops without significantly compromising yield (Gopalakrishnan *et al.*, 2011). While there is also skepticism that large-scale bioenergy production on marginal land is feasible (Bryngelsson & Lindgren, 2013), indeed, "marginal lands" have become commonly used shorthand for more complex solutions to the land use dilemma surrounding bioenergy dialogues. To this end there have been an increasing number of studies attempting to spatially quantify the amount of marginal land available for bioenergy production.

In Chapter 2, I present the first work to explore how the concept of marginal land is carried out in a Geographic Information Systems (GIS) framework. I evaluate the emergent body of published work that uses GIS to map the availability of marginal land as a proxy for bioenergy crop potential. I examine China as a comparative case study, and through this lens develop a framework to evaluate the other work based on several factors. I assess the various generic definitions of marginal land given and then investigate how each working definition is implemented in spatial models through definitions, data and analysis. Despite the similar goals of the studies, I conclude that there is no common working definition of marginal land across all of these works – there are considerable differences in mapped results that are driven by differences in definitions, model framework, data inputs, scale, and treatment of uncertainty. Two important conclusions from this work guide the development of the subsequent chapters of this dissertation. First, I found that most studies use simple linear GIS overlays that assume distinct thresholds between land that is marginal and land that is not. Second, I found that the overwhelming majority of the studies mention uncertainty in their mapped results only qualitatively, or make no mention of uncertainty or model sensitivity at all. Finally I provide a framework to help researchers evaluate existing scholarship in mapping marginal land and suggest possible changes that might be made to modeling protocols so that more realistic comparisons across projects can be made.

*Site suitability mapping*

The first necessary step in quantifying the production potential of bioenergy crops is to accurately measure and map land area available. As evidenced in Chapter 2, Geographical Information Systems (GIS) are a powerful set of tools that provide a framework to spatially model and interpret multiple layers of data (Caver, 1991; Jankowski, 1995). GIS-based site suitability models are increasingly used as a technique to identify locations for renewable energies like wind, solar and geothermal (Angelis-Dimakis *et al.*, 2011). Several GIS-based methods have been developed to assess the potential availability of bioenergy resources,

however more often than not, these studies use discrete thresholds that signify whether or not an area is suitable.

The most commonly used form of GIS-based land suitability analysis, and also the oldest, relies on linear combinations of spatially referenced input criteria, also known as "Boolean" overlays. This methodology has a long history that dates back to the Olmstead era of landscape architecture, when Charles Eliot and Warren Manning (both Olmstead protégés) layered several transparent sheets on a window to view multiple site characteristics simultaneously in order to visualize land cover suitability (McHarg, 1998). In subsequent decades, several planning studies utilized overlay mapping – the predominant method being "sieve mapping," whereby constraints eliminated unsuitable areas and areas passing through constraint layers were deemed suitable (Steinitz *et al.*, 1976). Still, while the use of the method was widespread, a theoretical reasoning for the methodology was lacking until Ian McHarg's 1969 work *Design with Nature* illustrated the use of overlay mapping using vector data, where individual criteria were superimposed to highlight the most suitable areas (McHarg Ian, 1969).

Since the era of transparent overlays, the expanding use of Geographic Information Systems (GIS) has fueled the application of modern suitability analyses (Malczewski, 2004). Today the same basic principles of the overlay methods are still employed, but the transition of suitability models from analog to digital provided greater power to the users and decision makers. The current digital models allow for much more complex mathematical combinations, including the use of weights and hierarchy (Malczewski, 2004). However, Boolean, weighted, and even hierarchical modeling approaches consider each feature on the landscape as having distinct boundaries (Jiang & Eastman, 2000). Yet in reality, natural boundaries are often much more complex and continuous. One way to address these problems concerning class boundaries and to better incorporate expert knowledge concerning them is to consider the natural gradient of the input datasets as well as the uncertainty inherent in model thresholds.

Fuzzy logic, introduced in 1965 by Lofti Zadeh of UC Berkeley, provides one way to express a degree of membership into multiple categories so that when applied to a site suitability model, a specified area can be both suitable and not-suitable at the same time (Zadeh, 1965). Fuzzy set theory transforms raw values into a degree of membership in a set, or a "truthness" value between 0 and 1, with a value of 1 representing the absolute truth that a parcel is suitable and 0 representing the absolute falsehood. A fuzzy value between 1 and 0 thereby expresses exactly how suitable an area is in comparison to its neighbors. Fuzzy logic has been used within site suitability models in varying ways, including industrial allocation and housing development (Jiang & Eastman, 2000; Malczewski, 2006), however the use of fuzzy theory in suitability modeling for renewable energy – particularly bioenergy – is limited (Cai *et al.*, 2011; Charabi & Gastli, 2011).

Both Chapters 3 and 4 employ site suitability models that incorporate fuzzy logic to identify a range of suitable land area for two drought-tolerant crops – a modified variety of switchgrass and

two species within the plant genus *Agave*. In addition to the primary fuzzy suitability model, each chapter explores the sensitivity of the fuzzy model in different ways, including modifications to overlay functions and input parameters, respectively.

Switchgrass (*Panicum virgatum*) was one of the first species to be targeted by the U.S. Department of Agriculture as an exemplary bioenergy crop, and research is underway to develop a modified variety that can withstand periods of drought while not substantially affecting overall yield. In Chapter 3, I develop an innovative suitability model rooted in fuzzy logic that highlights a continuous range of suitable areas to place a drought-tolerant variety of switchgrass into the landscape. I test the model's sensitivity to several fuzzy overlay techniques to determine which best synthesizes tradeoffs between suitability criteria. I also develop a specialized dryness index using high resolution (spatial & temporal) weather and soil data to provide a spatially explicit measure of dry spell severity for switchgrass across a landscape. My results demonstrate that the Gamma fuzzy overlay function best incorporates the tradeoffs between multiple criteria. Increasing the value of gamma shifts the distribution of the suitability index to a more equalized histogram that more clearly represents the suitability contrast between pixels. The results of the suitability model highlight the Great Plains region of the United States as a suitable region, which I then used to guide where to next focus a more detailed analysis in Kansas to quantify land area experiencing varying degrees of aridity. I found that that 96% of the suitable land area in Kansas falls within a dryness index equivalent to one 55-day long dry stretch, after which suitable land area shows diminishing returns. By identifying the dryness threshold where land area is maximized, the results of this analysis may inform the development of drought-tolerant varieties of switchgrass and contribute to plant science research.

In addition to switchgrass, another drought-tolerant crop – *Agave* – has recently garnered attention as a bioenergy crop. In Chapter 4, I explore the potential for *Agave* to be cultivated on lands unsuitable for conventional bioenergy production and to contribute to the country's domestic renewable energy needs. For this research I focus on two specific species – *Agave tequilana* and *Agave deserti* – and for each I develop a spatially explicit site suitability model based on empirical data to optimize ideal locations for each crop. I test the sensitivity of the model to input parameters, examining the influence of moderate changes to each independent suitability criteria on the dependent mapped result. I then adjust all parameters simultaneously to create multiple suitability scenarios. The results of the spatially-explicit suitability model present a range of land area suitable for *Agave* in the Southwestern US. The sensitivity analysis reveals that climatic variables, particularly minimum winter temperature, are the most limiting factors for *Agave* suitability in the US. From these results I estimate the potential contribution of *Agave* to regional energy demands (electricity and fuel) and find that at the state-level, Arizona and Texas show the highest potential energy benefit. I also compare the results from this analysis to estimates of future production of annual energy crops and perennial grasses from the U.S. Billion-Ton Update (Downing *et al.*, 2011). The evidence in this chapter demonstrates that

*Agave* has true potential to become a complementary bioenergy feedstock in areas unsuitable for conventional feedstocks.

Chapter 5 summarizes the key findings of my dissertation as they relate to bioenergy and site suitability research. I offer my perspectives on future directions for the growing body of research mapping marginal land for bioenergy crops. To this end, this dissertation presents an original and timely contribution to the renewable and sustainable energy literature.

# Chapter Two

## Mapping the potential for biofuel production on marginal lands: differences in definitions, data and models across scales

*Energy policies around the world are mandating increases in bioenergy production, yet there are concerns that the resources needed to fulfill these requests might come at the expense of food crops, wildlife habitat, or cultural amenities. For this reason, an emergent body of literature supports growing bioenergy feedstocks on marginal lands. There has been an increase in published work that uses Geographic Information Systems (GIS) to map the availability of marginal land as a proxy for bioenergy crop potential. However, despite the similarity in stated intent among these works there remain a number of inconsistencies across studies that make comparisons and standardization difficult. I reviewed a collection of recent literature that mapped bioenergy potential on marginal lands at global, national, regional and local scales, and found that there is no common working definition of marginal land across all of these works. One country – China – provides many case studies, yet despite parallel goals, these examples illustrate considerable differences in mapped results that are driven by differences in definitions, model framework, data inputs, scale and treatment of uncertainty. These differences were echoed throughout the broader literature, where data choices and modes of analysis were rarely consistent. Most papers reviewed here employed relatively simple GIS overlays using linear combination of input criteria with distinct thresholds identifying land as either marginal or not. In addition, these papers provided few details describing accuracy and uncertainty, even though the results of such analyses are especially important to help curb overly optimistic estimates. These differences are likely to be major impediments to integration of studies mapping marginal lands for bioenergy production. I suggest that there is future need for spatial modeling of bioenergy, yet further scholarship is needed to compare across countries and scales and to understand the cumulative global potential for bioenergy crops.*

### Introduction

As energy demands increase globally, there is a growing pressure for renewable energy sources to help meet requirements and simultaneously mitigate for climate change. While no single type of renewable energy will fulfill all of our needs, bioenergy promises to become a growing portion of the energy market (Nakicenovic *et al.*, 2000; Somerville *et al.*, 2010; Youngs & Somerville, 2012). Energy policies around the world are progressively mandating increases in bioenergy production, and most are targeting second generation non-food biofuels that promise to be more environmentally sustainable than first generation crop-based biofuels (eg. corn and soy) if they can be designed and managed appropriately (Tilman *et al.*, 2006, 2009; Dale *et al.*, 2010; Kim & Dale, 2011). In 2007 the United States Congress passed the Energy Independence and Security Act (EISA), which creates a more aggressive Renewable Fuel Standard (RFS2) – mandating at least 16 billion gallons of cellulosic ethanol by 2022 (Sissine, 2007). The United Kingdom has also passed government incentives, like the 2008 Energy Crop Scheme (ECS), to encourage the establishment of bioenergy crops. And in 2007 China's policies proclaimed a shift to non-food biofuels, which are expected to exceed 12 million tons by 2020 (Tian *et al.*, 2009).

However, as these ambitious policies mandating biofuel production are implemented, they often come ahead of the provision of reliable and accountable information on the extent of lands available for such a purpose (Robertson *et al.*, 2008). For example, in the United States alone, somewhere between 16 million and 21 million hectares (Mha) of non-crop land would be needed to meet the EISA target for cellulosic ethanol by 2022 (Perlack *et al.*, 2005; Schmer *et al.*, 2008; Swinton *et al.*, 2011). There are additional concerns that bioenergy production might come at the cost of food crops if croplands are converted to growing fuel, or at the expense of wildlife habitat or cultural amenities if protected areas are used for bioenergy planting. This *bioenergy land use dilemma* is challenging scientists to determine where exactly to plant bioenergy feedstocks to meet mandates, but also to ensure sustainable food production and environmental protection (Cassman & Liska, 2007; Elobeid & Hart, 2007; Lam *et al.*, 2009).

Many have proposed the use of marginal lands as a possible sustainable solution to the bioenergy land use dilemma, as one of the primary sustainable qualities of second-generation biofuels is their documented ability to grow on non-prime agricultural land (Tilman *et al.*, 2006; Gelfand *et al.*, 2013a). Much of the larger discussion in support of bioenergy as a viable energy resource makes reference to the global abundance of "marginal lands" as a practical solution to growing bioenergy while at the same time avoiding competition with productive agricultural lands and sensitive habitats, and maximizing net carbon benefits (Gelfand *et al.*, 2013a). While there is also skepticism that large-scale bioenergy production on marginal land is feasible (Bryngelsson & Lindgren, 2013), indeed, "marginal lands" have become commonly used shorthand for more complex solutions to the land use dilemma surrounding bioenergy dialogues.

This optimism is widespread globally, as several countries, including the United States, Canada, United Kingdom, Australia, India, Indonesia and China have adopted policies mandating expansion of non-food biofuel crops on degraded or marginal land (OECD/FAO, 2012). China's policies are especially adamant that bioenergy production not interfere with food security. Along with setting minimum limits on arable land, recent Chinese policies quintupled the tax on the use of arable land for other purposes, and in 2007 mandated a shift to exclude the use of grains in bioenergy production (Tian *et al.*, 2009).

The quantification of land available for bioenergy crops initially began with broad-based estimates where studies used reported data from existing land cover inventories, while adding new assumptions or imposing limits in the context of bioenergy (Berndes *et al.*, 2003). For example, Hall et al. (1993) assumed that 10% of the reported global crop, forest and woodland area may be available for bioenergy production. Other studies based their estimates on a compilation of existing literature. Hoogwijk et al. (2003) combined the estimates reported in literature published by Hall et al. (1993) and Houghton et al. (1993) for a total estimate of 430 to 580 Mha of degraded lands available for biomass production globally. More recently, in China two studies provided estimates of available land for biofuels based on the existing literature and government reports (Tian *et al.*, 2009; Tang *et al.*, 2010). What these studies lack, however, is spatially explicit information on exactly where these lands can be found.

To visualize the potential spatial patterns of bioenergy production on marginal land there has been an increase in published work that uses quantitative and spatially explicit models within Geographic Information Systems (GIS) to map the availability of marginal land for bioenergy

crops. However, despite the similarity in stated goals among these works – mapping marginal land for bioenergy production – there remain a number of inconsistencies across studies that make comparisons and standardization difficult, and to date none of the current methods are widely accepted (Calvert *et al.*, 2013; Kang *et al.*, 2013a).

In this paper I shed light on the practical challenge of mapping bioenergy potential across scales by reviewing recent literature that uses geospatial technology to explicitly map marginal, abandoned or degraded lands specifically for the purpose of planting bioenergy crops. First I examined China as a comparative case study. Through this lens I developed a framework to evaluate the other work based on several factors: I first examined the various definitions of marginal land given, then investigated how each working definition is implemented in spatial models through model choice, data selection, scale, and treatment of uncertainty. My goal was to provide a framework to help researchers evaluate existing marginal land scholarship, and to suggest possible changes that might be made to modeling protocols so that more realistic comparisons across projects can be made.

This study reviewed papers that use geospatial techniques to spatially link placement of bioenergy production on marginal land. I used the Google Scholar Database, the Web of Knowledge and citation tracing to identify papers on bioenergy and marginal lands published over a 20-year period from 1993 to 2013. My search words included: bioenergy, biofuel, biomass, biodiesel, marginal land, abandoned land, degraded land, land, GIS, spatial, and scale. Articles were removed if they did not meet all of the following criteria; 1) target second generation bioenergy feedstocks, 2) focus on land that may be categorized as marginal, and 3) use spatially-explicit (GIS) techniques. Papers that used mapped results from other studies were removed from the list (for example, Zhuang et al. (2011) use the work published by Cai et al. (2011) in their analysis of marginal land for algae production). My search criteria narrowed the review to 22 papers from years 2008 to 2013. Of the studies reviewed here, five were conducted at a global scale, seven were national, nine were regional, and one was conducted at the local (city) scale (Table 2-1). The specific bioenergy crops targeted by the studies also varied. Some addressed bioenergy generally, others focused on single crops, and others concentrated on a suite of crops specific to the region.

**Defining marginal land**

The term “marginal land” is currently so intertwined with discussions surrounding bioenergy that its definition might be assumed to be specific and certain. Yet the definition of marginal land has varied across time, space and discipline to meet multiple management goals (Kang *et al.*, 2013a). The concept of marginal land made an appearance in the literature in the 19th century when David Ricardo identified variations in the desirability of land based on its proximity to essential resources, such as water supply or food markets (Ricardo, 1817). Ricardo’s land rent theory posited the logical view that certain land units have a comparably higher value if located in proximity to essential resources. In the early 1900’s an economic concept of marginal land emerged. Peterson and Galbraith (Peterson & Galbraith, 1932) discussed marginal land in a theoretical framework, wherein rational land use decisions stem from and respond to changes on the “extensive margin,” where revenue is equal to the costs of production. They also highlighted

| | Year | Author | Scale | Term Used | General Definition | Crop | Soil[a] | Land Cover[a] | Slope[a] | Climate[a] |
|---|---|---|---|---|---|---|---|---|---|---|
| Global | 2008 | Campbell et al. | Global | Abandoned agriculture lands | "areas that have been abandoned to crop and pasture due to the relocation of agriculture and due to degradation from intensive use" | bioenergy crops | none | 1) 2) | none | none |
| | 2008 | Field et al. | Global | Abandoned cropland, abandoned pastureland | "land that was previously used for agriculture or pasture but that has been abandoned and not converted to forest or urban areas" | biomass energy | none | 1) 2) | none | none |
| | 2011 | Cai et al. | Semi-Global | Marginal agricultural land | "has low inherent productivity for agriculture, is susceptible to degradation, and is high-risk for agricultural production" | 2nd generation biofuel feedstocks and LIHD (low input high diversity) perennial grasses | 1) 2) | 1) | 1) | 1) |
| | 2012 | Nijsen et al. | Global | Degraded land | degraded lands not in use as forest, cropland, pastoral land, or urban | perennial woody and grassy energy crops | 3) 4) | 3) | none | none |
| | 2009 | Milbrandt & Overend | Semi-Global (APEC countries) | Marginal land | "characterized by poor climate, poor physical characteristics, or difficult cultivation" | general lignocellulosic | 5) | 4) | 2) | none |
| National | 2009 | Lovett et al. | National (UK) | Marginal and idle land | class 3 & 4 agricultural lands, minus constraints | Miscanthus | 12) | 10) | 3) | none |
| | 2011 | Odeh et al. | National (Australia) | Marginal agricultural lands | here the definition "is based on precipitation since the most limiting factor for dryland agricultural intensification is low water availability caused by low amount of precipitation. However, the marginal agricultural lands may be characterized by degraded soils, particularly the saline soils" | pongam - perennial tree and Indian mustard - oilseed crop | 14) | none | none | 2) |
| | 2011 | Schweers et al. | National (China) | Degraded and abandoned land | "land degradation is a long-term loss of ecosystem function and services, not least production, caused by disturbances from which the system cannot recover unaided" | generic bioenergy crops | 6) | 6) 5) | 4) | none |
| | 2011 | Swinton et al. | National (USA) | Marginal land | non-crop land at "the extensive margin, where land quality is low enough that the value of biomass produced just covers its cost of production" | cellulosic biofuels | none | 2) | none | none |
| | 2011 | Zhuang et al. | National (China) | Marginal land | "land that has relatively poor natural condition but is able grow energy plants, or land that currently is not used for agricultural production but can grow certain plants" | 5 spp. bioenergy crops | 13) | 13) | 5) | 3) |
| | 2012 | Lu et al. | National (China) | Marginal land | "land unsuitable for crop production, but ideal for the growth of energy plants with high stress resistance. These lands include barren mountains, barren lands and alkaline lands" | *Pistacia chinensis* | 13) | 12) | 6) | 4) |
| | 2012 | Liu et al. (a) | National (Canada) | Marginal land | "lands which are not profitable for food crops due to low productivity" | hybrid poplar | 9) | 9) | none | none |
| Regional | 2011 | Fahd | Regional (Campania, Italy) | Marginal land | "characterized by several negative features (no water easily available, poor organic matter in soil, distance from transportation routes, excess slope and sometimes pollution due to previous use by industrial activities, etc.) that make them unsuitable for food cropping." | *Brassica carinata* | none | 7) | 7) | none |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **2010** | **Fiorese & Guariso** | Regional (N. Italy) | Agricultural marginal land and set-aside land | agricultural marginal land (i.e. abandoned agricultural land) would be targeted for short-term forestry; set-aside land would be targeted for sorghum crop; available land = parcel area * (marginal agricultural land/ total agricultural land) | arborous and herbaceous species | 11) | none | unknown | 6) |
| | **2010** | **Wu et al.** | Regional (SW China) | Marginal land | land with social and economic constraints | *Jatropha curcas* | 10) | 11) | 6) | 7) |
| | **2011** | **Gopalakrishnan et al.** | Regional (NE, USA) | Marginal land | "land that is marginal for conventional crops but not marginal for biofuel crops or other functions, based on economic, soil health, and environmental criteria" | cellulosic biofuels | 16) | 8) | 8) | none |
| | **2012** | **Kukk et al.** | Regional (Tartu Co., Estonia) | Abandoned agricultural land | "soils of abandoned areas are generally of low quality and thereby limited suitability for crop production" | 5 spp. bioenergy crops | 8) | none | none | none |
| | **2012** | **Liu et al. (b)** | Regional (SW China) | Marginal land | "land that has relatively poor natural condition but is able grow energy plants, or land that currently is not used for agricultural production but can grow certain plants" | *Manihot esculenta* (cassava) | 13) | 13) | 5) | 5) |
| | **2012** | **Tenerelli and Carver** | Regional (Apulia Region, S. Italy) | Marginal land or less productive land | "less productive land (grade three, four and five in the Agricultural Land Capability Map)" | perennial energy crops | 7) | 10) | 9) | 8) |
| | **2013** | **Kang et al.** | Regional (MN, USA) | Marginal land | "characterized by low productivity and reduced economic return or by severe constraints for agricultural cultivation. They are generally fragile and their use is environmentally risky. Land can be marginal physically, biologically, environmentally-ecologically, economically" | general | 15) | none | 10) | none |
| | **2013** | **Gelfand et al.** | Regional (USA Midwest) | Marginal land | “those poorly suited for food crops because of low productivity due to inherent edaphic or climatic limitations or because they are located in areas that are vulnerable to erosion or other environmental risks when cultivated” | cellulosic biofuels | 15) | 8) | 11) | none |
| **Local** | **2012** | **Niblick et al.** | Local (Pittsburgh, USA) | Urban marginal lands | "lots with poor agricultural potential and unfit for residential purposes" | sunflower | 15) | 14) | 10) | none |

**Table 2-1.** Summary of papers reviewed.

[a] See Table 2-3 for soil and land cover references.

the importance of socio-economic factors, including cost of living, landholding size, accessibility to credit, and land tenure policies. This concept of the extensive margin has been well studied since (Thünen & Heinrich, 1966; Barlowe, 1978; Hardie *et al.*, 2004; Taheripour *et al.*, 2012) and is still recognized by the USDA as playing a role in land use change today (Lubowski *et al.*, 2006).

More recently the definition of marginal land began to take on an explicit spatial characteristic. In the 1980's several studies began to map physically marginal lands for the purpose of inventorying underproductive agricultural land, often with the goal of taking vulnerable and risky lands out of production (Runge *et al.*, 1986). These studies were largely based on locating soils with physical restrictions and production constraints. In 1988, a regional study in Minnesota was conducted to find "marginal agricultural land" to be targeted for a cropland set-aside program based on erodible soils and poor land productivity (Larson *et al.*, 1988). In 1990, Breuning-Madsen performed an overlay mapping assessment of wet and droughty soils and those on steep slopes marginal for agricultural use in Denmark to isolate areas of the country most economically vulnerable to low yields (Breuning-Madsen *et al.*, 1990). A decade later, researchers in Poland set out to map "marginal agricultural lands" and "less favored farming areas", respectively, based on biophysical characteristics, the later as part of a mandatory inventory requirement before joining the European Union (Ostrowski, 1999; Bielecka, 2002). Such studies relied solely on biophysical definitions of marginal land, including poor soils, poor drainage, and steep slopes.

Currently, the term marginal land is often intimately linked with bioenergy research. This emergent conversation is evidenced by GoogleScholar search statistics, which show that since 1993 there have been an increasing number of papers addressing marginal lands, biofuels, GIS and any combination of these. The initial jump may arguably be credited to Hall et al. (1993), who highlighted the potential benefits of biomass production to restore degraded lands. Since then, interest in the topic further expanded after publication of a paper by Tilman et al. (2006), and especially in 2008 with both the emergence of the food versus fuel debate (Ignaciuk *et al.*, 2006; Johansson & Azar, 2007; Rathmann *et al.*, 2010) and finally with the publication of the first studies to map the global extent of abandoned or degraded agricultural lands for bioenergy production (Campbell *et al.*, 2008a; Field *et al.*, 2008). However, as the subject becomes an increasing focus in the literature, the working definitions of marginal land become increasingly diverse, making comparisons between studies, and standardization of estimates difficult.

Each of the papers reviewed here offer definitions of marginal land as they relate to bioenergy. However, there is a common distinction between the general definition of marginal land referred to in the introduction of each paper, and the working definition of marginal land as implemented in the methods by way of crop choice, input criteria and modeling framework. General definitions of marginal land are fairly generic and are also relatively consistent across studies. The definition offered by Gelfand et al. (2013) is a typical example, where they describe marginal lands as "*those poorly suited for food crops because of low productivity due to inherent edaphic or climatic limitations or because they are located in areas that are vulnerable to erosion or other environmental risks when cultivated*." The general agreement is that although these lands are unsuited for conventional agricultural crops, they are conversely well suited for

bioenergy crops. For example, Gopalakrishnan et al. (2011) define marginal land as "*marginal for conventional crops but not marginal for biofuel crops or other functions, based on economic, soil health, and environmental criteria."* General definitions overwhelmingly describe lands that are not prime for conventional crops because they are high risk for economic payoff owing to low productivity resulting from climate or soil limitations.

In contrast, the working definitions of marginal land implemented in the papers differ considerably between studies. Working definitions of marginal land vary by nation and target crop, which in turn drive input criteria and modeling framework. Thus direct comparisons between published outputs are difficult. Moreover, there are few areas in the world that have been mapped with multiple methods making comparisons practical. One exception is China, which is currently under enormous pressure by energy and food security policies not to grow bioenergy crops on agricultural land. For this reason, bioenergy potential in China has been mapped several times and in different ways. As a precursor to developing a framework for evaluating the larger body of work, it is useful and illustrative to examine how different studies map bioenergy potential on marginal land in China.

**A comparative case study: mapping marginal lands for bioenergy in China**

Of the studies reviewed here, five reported nationwide results for the People's Republic of China (three at the national scale and two at the global scale). Each used different working definitions of marginal land: they focused on different crops, used different input criteria, models and assumptions, and all had different mapped results.

Lu et al. (2012) mapped the national potential for growing Chinese Pistache (*Pistacia chinensis*) on marginal lands in China. Their general definition of marginal land is land "*unsuitable for crop production, but ideal for the growth of energy plants with high stress resistance. These lands include barren mountains, barren lands and alkaline lands."* They used a GIS overlay analysis to first target marginal land uses suitable for planting bioenergy crops, including sparse forest, natural grassland, and unused land (alkaline, bare, shoal/bottom lands). Then they characterized three levels of suitability for Chinese Pistache production based on eco-environmental requirements (temperature, precipitation, soil and slope), and modified their results with social, economic and environmental constraints, including sensitive and protected areas, national reserves, and cultivated lands. Before producing their final mapped results, they employed a minimum mapping unit of 200 ha. All data are referenced in Table 2-2. This method yielded 19.9 Mha of available land (2.08% of China's total land area).

Zhuang et al. (2011) mapped the national potential for five bioenergy species, including Jerusalem artichoke (*Helianthus tuberous L.*), Chinese Pistache, Chinese castor oil (*Jatropha curcas L.*), Cassava (*Manihot esculenta*), and Tung Tree (*Vernicia fordii*). They generally defined marginal land as "*land that has relatively poor natural condition but is able to grow energy plants, or land that currently is not used for agricultural production but can grow certain plants."* They used a GIS overlay analysis with binary thresholds that first identified marginal lands based on land use, terrain (slope >25% and elevation specific to species), climate (specific to species) and soil (they excluded sand, sandy gravely, saline and alkalized soil, and soil depth). Then they modeled the optimum location on the identified marginal lands for each species based

on the eco-environmental requirements. Land uses considered suitable for bioenergy included woodlands (shrub land, sparse forest land), grasslands and barren lands (including shoal/bottomland, saline and alkaline land, and bare land) (Table 2-2). The results suggested that 43.75 Mha were available for these five species (4.57% China's total land area).

Schweers et al. (2011) performed a national study in China examining the potential for general bioenergy production on marginal land defined as degraded and abandoned land. They qualified that, "*land degradation is a long-term loss of ecosystem function and services, not least production, caused by disturbances from which the system cannot recover unaided."* They used a GIS overlay analysis with multiple input criteria that defined degraded land as the loss of net primary productivity (NPP) between 1981 and 2003. They calculated abandoned land by subtracting the maximum value from the percentage grassland and cultivated land between 1700 and 2000 from those values in 2005, where negative change values indicated abandoned land. Lands suitable for bioenergy crops included land cover with mixed vegetation and cropland, land cover with mixed grassland, forest or shrub land, lands with closed to open shrub land, lands with closed to open herbaceous vegetation, and lands with sparse vegetation. Land covers unsuitable for bioenergy crops included croplands, forested areas, wetlands, urban areas, water, snow and ice, bare and un-defined lands. They also excluded conservation areas, including protected areas identified by the World Database on Protected Areas (WDPA), areas of high biodiversity, areas with high percent organic carbon content, and steep slopes (Table 2-2). The authors also conducted ground verification in two regions including GPS, photos, and interviews – with 60% of the verification sites totally agreeing with the remote suitability assessment. This method yielded 39.1 Mha for bioenergy crops when conservation areas were not excluded, or 20.2 Mha when conservation areas were excluded (2.11% or 4.09% China's total land area, respectively).

Milbrandt and Overend (2009) mapped the potential for lignocellulosic biomass plants on marginal land in APEC countries, including China, defined as land "*characterized by poor climate, poor physical characteristics, or difficult cultivation. They include areas with limited rainfall, extreme temperatures, low quality soil, steep terrain, or other problems for agriculture*." They employed a GIS overlay analysis that used linear combination with multiple input criteria using binary thresholds. The analysis focused primarily on land cover, soils and slope (Table 2-2). Bare and herbaceous areas (not in use or with only moderately intensive pastoralism) were targeted, and lands with intensive and extensive pastoralism were excluded. Lands with moderate and steep slopes and lands with soil problems (e.g. course, sandy, acidic) or shallow soils were considered marginal lands suitable for planting bioenergy crops. After marginal land was mapped based on the above criteria, they excluded protected lands as well as deserts, cold regions, ice/glacier areas, water features, forests, agricultural lands, urban areas, as well as herbaceous and bare lands under intensive and extensive pastoralism. The analysis predicted 51 Mha (5.34% of China's total land area) should be available for bioenergy production.

Finally, Cai et al. (2011) modeled the potential for second-generation biofuel feedstocks and low-input high-diversity (LIHD) mixtures of native perennials for most of the globe on marginal agricultural land that has, "*low inherent productivity for agriculture, is susceptible to degradation, and is high risk for agricultural production."* As compared to the four studies

previously mentioned, the authors used a more complex model that employed Fuzzy Logic Modeling (FLM) and land cover exclusion. Their model identified marginally productive lands based on soil productivity, slope, and climate (soil temperature regime and humidity index) (Table 2-2). Input criteria were evaluated by applying a membership function to each criterion based on empirical knowledge, thereby converting quantitative values to qualitative ratings on the level of land productivity (low, marginal, or regular). Criteria were then aggregated into probabilities of land belonging to a category of land productivity, which yielded a final land productivity score. After the marginal lands were mapped, the authors developed four scenarios of marginal lands available for bioenergy crops that progressively included varying combinations of land cover. The first scenario included mixed crop and marginally productive natural vegetation. The second scenario included the land cover categories in Scenario 1 in addition to marginally productive cropland. To this list the third scenario added marginally productive grasslands, savanna, and shrub lands and assumed regularly productive regions of these areas are excluded for pasture or for future crops. Finally, the fourth scenario removed marginally productive pastureland possibly accounted for in the grassland class in scenario 3 and instead included regular land used for mixed crop and vegetation and grassland, savannah, and shrub land with either regular or marginal productivity. For China specifically, mapped results yielded 52-213 Mha of marginal agricultural land, depending on scenario (5.43% to 22.26% of China's total land area).

In summary, despite similarity of intent, these five examples from China illustrate considerable differences in working definition, including input criteria, modeling framework, and mapped results. Mapped results ranged from 2.08 to 22.26% of China's total land area. Of the five projects, all five used some representation of land cover, but each model used a different land cover dataset (Table 2-2). In addition, studies using nearly the same dataset (only with different publication dates, but identical thematic classes) differ in which thematic classes they considered marginal. For example, while both Zhuang et al. (2011) and Lu et al. (2012) initially targeted shrub land as a marginal land class, Lu et al. (2012) later ultimately excluded the shrub land category in their study, citing China's policies on forestry that state shrub land should not be modified for other purposes. Another differentiating factor regarding land cover is the inclusion or exclusion of existing cropland. Both Schweers et al. (2011) and Cai et al. (2011) included cropland in their analysis, while the others did not.

Even among studies mapping within the same country and using the same national data sources, there are both subtle and large differences between input criteria and parameters that impact mapped land area. In the present analysis, both Zhuang et al. (2011) and Lu et al. (2012) used nearly identical datasets (they provide one of the closest comparisons of the papers reviewed here), but have different mapped results: 43.75 and 19.9 Mha, respectively. This is in large part due to the fact that the former mapped marginal land for five bioenergy species and the latter only mapped for one species. However, differences may also be attributed to variances in both input data and parameters applied. For example, Lu et al. (2012) used 2000 land use data from the RESDC, while Zhuang et al. (2011) use land used data from the same source and at the same scale, but for a different year – 2008. In addition, Lu et al. (2012) ultimately *excluded* shrub lands from their analysis, while Zhuang et al. (2011) *included* shrubs in their analysis. The two studies also got their soil and terrain data from the same sources (RESDC for soil and SBSC for terrain), but reported two different mapping scales, 1km and 1:100k, respectively for soil, and

| Author | Scale | Term Used | Target Species | General Definition | Method | Marginal area (Mha) | % China's land area | Soil | Topo-graphy | Current Land Use | Historical Land Use | Climate Data |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cai et al., 2011 | Global | Marginal Agricultural Land | low input high diversity (LIHD) mixtures of native perennials | lands with "low inherent productivity for agriculture, is susceptible to degradation, and is high risk for agricultural production" | fuzzy logic modeling & boolean knockout | 52-213 | 5.43 - 22.26% | HWSD (30"); USDA-NRCS STR (2') | GTS (30") | IGBP (30") | none | Climatic Research Unit (2') |
| Milbrandt & Overend, 2009 | Semi-Global (APEC countries) | Marginal Land | lignocellulosic biomass | lands "characterized by poor climate, poor physical characteristics, or difficult cultivation. They include areas with limited rainfall, extreme temperatures, low quality soil, steep terrain, or other problems for agriculture" | GIS overlay analysis with binary thresholds | 51 | 5.33% | GAEZ (5') | GAEZ (5') | FAO (5') | none | none |
| Schweers et al., 2011 | National (China) | Degrading and Abandoned Land | generic bioenergy spp. | "land degradation is a long-term loss of ecosystem function and services, not least production, caused by disturbances from which the system cannot recover unaided" | GIS overlay analysis with binary thresholds | 20.2 - 39.1 | 2.11 - 4.09% | na | USGS GTOPO30 (30") | GlobCover LCD (300m) | HYDE (5') | none |
| Zhuang et al., 2011 | National (China) | Marginal Land | 5 spp.: 1. *Helianthus tuberoush* L., 2. *Pistania chinensis*, 3. *Jatropha curcas L.*, 4. *Manihot esculenta*, 5. *Vernicia fordii* | "land that has relatively poor natural condition but is able to grow energy plants, or land that currently is not used for agricultural production but can grow certain plants" | GIS overlay analysis with binary thresholds | 43.75 | 4.57% | RESDC (1:100k) | SBSC (1:100k) | RESDC (1:100k) 2008 | none | (CMA) (1:100k) |
| Lu et al., 2012 | National (China) | Marginal Land | specific spp. (*Pistachia chinensis*) | lands "unsuitable for crop production, but ideal for the growth of energy plants with high stress resistance. These lands include barren mountains, barren lands and alkaline lands" | GIS overlay analysis with multiple thresholds | 19.9 | 2.08% | RESDC (1km) | SBSC (1:250k) | RESDC (1:100k) 2000 | none | (CMA) (1km) |

| | | |
|---|---|---|
| CMA | ... | China Meteorological Administration |
| GAEZ | ... | Global Agro-Ecological Zones (GAEZ) system |
| GTS | ... | Global Terrain Slope |
| HWSD | ... | Harmonized World Soil Database |
| HYDE | ... | History Database of the Global Environment |
| RESDC | ... | Data Center for Resources and Environmental Sciences |
| SBSC | ... | State Bureau of Surveying and Cartography |
| STR | ... | Soil Temperature Regime |
| USGS | ... | United States Geological Survey |
| WDPA | ... | World Database of Protected Areas |

na = not applied

**Table 2-2.** Summary of studies calculating nationwide marginal land area in China.

1:250k and 1:100k, respectively for terrain. When targeting marginal lands, Lu et al. (2012) and Zhuang et al. (2011) both *included* alkaline lands as mapped in the RESDC land use database. However, Zhuang et al. (2011) then went on to *exclude* "seriously alkalized soil and saline soil" as reported in the RESDC soils database. Because the two studies mapped for different species, it is hard to say how influential these additional differences in input criteria are, however it may be assumed that the mapped results would still not be identical.

**Mapping potential for bioenergy**

These examples from China illustrate considerable differences in mapped results that were driven by differences in definitions, model framework, and data input. Next I evaluated the broader literature and further evaluated possibilities for commonalities. This review included 22 papers from years 2008 to 2013; five were conducted at a global scale, seven were national, nine were regional, and one was conducted at the local (city) scale (Table 2-1).

*Model selection*

Geographical Information Systems (GIS) are a powerful set of tools for site suitability models that incorporate numerous input datasets (Caver, 1991; Jankowski, 1995). GIS-based site suitability analyses are increasingly used to identify potential locations for integrating renewable energies into the landscape (Voivontas *et al.*, 1998; Angelis-Dimakis *et al.*, 2011). The majority of these analyses are based in Multi-Criteria Evaluation (MCE), the underlying principal of which is to synthesize complex problems by examining the coincidence of factors among multiple spatially co-registered variables (Malczewski, 2004).

GIS suitability analyses can range from very simple GIS linear overlays with binary thresholds, to very complex models that incorporate binned thresholds, weights, standardization, fuzzy logic, or all of the above. The oldest and most commonly used form of GIS-based land suitability analysis relies on linear combinations of spatially referenced input criteria, also known as "Boolean" overlays. This methodology has a long history (McHarg & Mumford, 1969) and has the benefit of being transparent and easy to follow. These models typically result in distinct thresholds of suitability, i.e. a parcel of land is either suitable for planting bioenergy crops or is not. The resulting suitability maps are often further restricted by the application of constraints. For example, marginal lands otherwise potentially suitable for bioenergy crops are modified by the elimination of unsuitable land (e.g. urban areas or water). Another category of geospatial land suitability models are those that incorporate Weighted Linear Combination (WLC) (Malczewski, 2004). Models that use WLC bin each criterion into categories, which then can be weighted based on their importance as decided by experts in the field. Therefore when combined, one criterion with relative low suitability can be recompensed by the high score of another. In contrast to Boolean and WLC models, which employ distinct thresholds of suitability, models that incorporate fuzzy set theory assign continuous grades of suitability (Zadeh, 1965; Jiang & Eastman, 2000). In a fuzzy set, the concept of suitability, or 'membership,' is not definitive because all objects belong to the suitability set in varying degrees. Using membership transformation functions, input criteria are given standardized "fuzzy" membership values, which vary continuously between 1 and 0; values approaching 1 are considered more suitable

and values approaching 0 are considered less suitable. This approach incorporates more realistically the continuous nature of biophysical and economic variables (Malczewski, 2004).

The majority of the studies reviewed here used relatively simple GIS suitability models that relied on linear combinations of spatially referenced input criteria with distinct thresholds, i.e. "Boolean" overlays. The one primary difference between the studies that used the linear combination overlay method is simply the succession in which input criteria were added to the model. For example, five of the twenty-two studies reviewed here first mapped for species suitability based on biophysical crop requirements, then mapped for land availability based on marginality (Lovett *et al.*, 2009; Fiorese & Guariso, 2010; Wu *et al.*, 2010; Lu *et al.*, 2012). Four other studies first mapped marginal land, and then mapped the potential suitability of that land for specific bioenergy species (Odeh *et al.*, 2011; Zhuang *et al.*, 2011; Fahd *et al.*, 2012; Liu *et al.*, 2012b). The eight studies that mapped for generic bioenergy species first mapped marginal lands, then employed land use constraints, or masks, where bioenergy crops should not be planted (Campbell *et al.*, 2008a; Field *et al.*, 2008; Milbrandt & Overend, 2009; Cai *et al.*, 2011; Schweers *et al.*, 2011; Liu *et al.*, 2012a; Niblick *et al.*, 2012; Nijsen *et al.*, 2012). Finally, three studies first targeted marginal lands for generic bioenergy species based on multiple input criteria, but employed no land cover constraints (Gopalakrishnan *et al.*, 2011; Swinton *et al.*, 2011; Kang *et al.*, 2013b). For example, Kang et al. (2013a) performed a "hierarchical" GIS overlay in succession mapping lands that were considered *physically* marginal (using slope, rock fragment, bedrock depth, flooding, and ponding), *biologically* marginal (using temperature, moisture, soil erosion, soil depth, sand content, production, CEC, EC, sodicity, pH, drainage, water table, and soil restriction), *environmentally-ecologically* marginal (using soil organic carbon trend, slope, erosion, wetland), and finally lands that were *economically* marginal (from breakeven price or yield). In this way they targeted marginal lands by defining what they are considered to be, but no specific land cover constraints were employed.

The remaining three studies reviewed here used more complex methods for mapping marginal land. Tenerelli and Carver (2012) combined a suite of GIS suitability methods. After identifying bioenergy crop types (perennial grasses, short rotation coppice and short rotation forestry) and ecological requirements based on crop typology, they both standardized and weighted input criteria. At the same time they excluded constraints (e.g. built-up areas, natural habitats, land with high ecological value, and highly productive agricultural land) based on a binary evaluation. These variables (the standardized and weighted criteria, the binary constraints, and the uncertainty and sensitivity analyses) were then used to generate the land capability index specifically for the bioenergy crops being studied. This tailored land capability index was then combined with an existing land capability classification to derive the final land allocation for bioenergy crops.

Only two methods reviewed here incorporated fuzzy set theory into their suitability models. Wu et al. (2010) used the FAO's Agro-Ecological Zone (AEZ) method to map marginal land for Jatropha, and for this the authors employed a fuzzy membership approach. First they sorted three soil quality indicators into five sequential bins. Then based on the crop's soil quality requirements, they identified the best suitable range of each of soil quality factors with a fuzzy membership equaling 1.0 and transition zone with membership values between 0.5 and 1.0. They then aggregated the degree of membership over three soil quality factors into one index with

three categories of land (suitable, moderately suitable, and unsuitable for Jatropha plantation). Cai et al. (2011) delved much deeper into fuzzy set theory for their analysis. As described above, the authors first identified marginally productive land using a fuzzy logic model based on the soil rating for plant growth (SRPG) index. The model applied a fuzzy membership to input criteria based on empirical or expert knowledge, thereby converting quantitative land productivity values into qualitative ratings. Criteria were then aggregated by fuzzy rule inference to determine the probabilities of land belonging to a category of land productivity, and a final land productivity index was generated. Steps which were based on empirical or expert knowledge were then iteratively calibrated through a learning process that incorporated existing land use. Finally they overlaid mapped marginally productive land with existing land cover and designed four land availability scenarios for bioenergy crops.

One modeling step commonly employed among the studies reviewed here, and in GIS-based land suitability analyses in general, was the implementation of land cover constraints, or restrictions – however, not all restriction criteria are alike. Restrictions that eliminate an area from the analysis can be classified as either "hard" or "soft" (Angelis-Dimakis *et al.*, 2011; Calvert *et al.*, 2013). Hard restrictions are land covers like urban areas and ice/snow – lands where it can generally be agreed that large bioenergy plantations *cannot* be planted. For this reason, hard restrictions were generally consistent among the studies reviewed here. This is in contrast to soft restrictions, which may include croplands, pasturelands, and shrub lands – lands where there may be legitimate reasons that bioenergy crops *should not* be planted, but it is feasible that they may be. For example, Lovett et al. (2009) first mapped nine absolute factors that precluded any opportunity for energy crop planting, then they mapped two secondary factors, where planting perennial biomass crops would not be *encouraged*, but also not necessarily *excluded*. Soft restrictions may be modified over time and can also vary regionally (for example shrub lands may be considered a hard restraint in China, but a soft restraint elsewhere), which makes them harder to standardize, especially at international and global scales (Angelis-Dimakis *et al.*, 2011).

*Input data*

*Biophysical data inputs*

Of the studies reviewed here, soil, land cover and topography (i.e. slope) were the most common input criteria used. Of these, land cover was the most frequently used input criteria (used by 19 out of 22 studies), and was most often employed as a masking or elimination factor in the analysis. For example, it was generally assumed that bioenergy crops will not be grown on lands classified as water, ice or snow. Soil variables were also a common input factor, with 18 out of 22 of the studies reviewed here using at least one soil variable as an input criterion. For instance, Milbrandt and Overend (2009) used 10 soil variables to target marginal lands potentially suitable for bioenergy crops, including soil texture, fertility, pH, etc. Topography (particularly slope) was the third most considered factor, with more than half (15 of 22), of the studies using slope as an input criterion defining land marginality. Primarily slopes too steep for planting bioenergy crops were excluded, for example three studies mapping marginal land in China excluded slopes >25% (Schweers *et al.*, 2011; Zhuang *et al.*, 2011; Lu *et al.*, 2012).

| ***Soil Datasets*** | | **Source** | **Geography** | **Date** | **Resolution** | **Used By** |
|---|---|---|---|---|---|---|
| 1) | *HWSD* | FAO/ IIASA | Global | 2009 | 30 arc sec | Cai et al. |
| 2) | *STR* | USDA-NRCS | Global | 1997 | 2' (1:5x$10^6$) | Cai et al. |
| 3) | *GLASOD* | ISRIC | Global | 1945-1990 | mmu 5,652$km^2$ (1:$10^6$) | Nijsen et al. |
| 4) | *FAO World soil map* | FAO-UN | Global | 1991 | 32$km^2$ (1:5x$10^6$) | Nijsen et al. |
| 5) | *GAEZ* | FAO/IIASA | Global | ? | 5' (~100$km^2$) | Milbrandt and Overend |
| 6) | *WDPA* | UNEP-WCMC | Global | 2008? | ? | Schweers et al. |
| 7) | *Soil Associations in England* | NSRI | National (England) | 2008 | 1:250k | Tenerelli and Carver |
| 8) | *Estonian Board digital soil map* | ELB | National (Estonia) | ? | 1:10k | Kukk et al. |
| 9) | *Land Suitability Rating System* | SDSC | National (Canada) | 1995 | ? | Liu et al.(a) |
| 10) | *Multi-source* | ? | National (China) | ? | ~1:250k | Wu et al. |
| 11) | *Unknown source* | ? | National (Italy) | ? | 1:250k | Fiorese & Guariso |
| 12) | *NatMap 1000 database* | NSRI LandIS | National (Whales & England | 2001 | 1$km^2$ | Lovett et al. |
| 13) | *Soil Map of China* | RESDC | National (China) | 2005 | ~1$km^2$ (1:100k) | Lu et al.; Liu et al.(b); Zhuang et al. |
| 14) | *CISIRO* | AGO | National (Australia) | 2000 | ? | Odeh et al. |
| 15) | *SSURGO* | USDA-NRCS | Regional (USA) | '08; '10 | 1:24k | Kang et al.; Niblick et al.; Gelfand et al. |
| 16) | *STATSGO* | USDA-NRCS | State (USA) | 1991 | mmu 617ha; 1:250k | Gopalakrishnan et al. |

| ***Land Cover Datasets*** | | **Source** | **Geography** | **Date** | **Resolution** | **Used By** | **Accuracy** |
|---|---|---|---|---|---|---|---|
| 1) | *IGBP DISCover* | Biradar et al. | Global | 1997 | 30" (~1$km^2$) | Cai et al.; Campbell et al.; Nijsen et al. | 59.4-78.7% |
| 2) | *MODIS* | NASA | Global | '04, '06, '08 | 500m | Campbell et al.; Field et al.; Swinton et al. | ? |
| 3) | *GLC2000* | GVM, JRC | Global | 2000 | 1$km^2$ | Nijsen et al. | 68.60% |
| 4) | *Global Agro-Ecological Zones (GAEZ)* | *FAO & IIASA* | Global | 2000 | 5' | Milbrandt & Overend | ? Can't find |
| 5) | *GlobCover* | ESA | Global | 2005 | 300m | Schweers et al. | 67.10% |
| 6) | *HYDE 3.0* | MNP | Global | 1700-2000 | 5' | Field et al; Campbell et al; Schweers et al. | uncertainty 5% (2000) to 25% (1800) & > |
| 7) | *CORINE* | EEA | Europe | 2006 | 100m | Fahd et al. | 87% +/- 0.8 |
| 8) | *Cropland Data Layer (CDL)* | USDA | National (USA) | 2007 | 56m | Gopalakrishnan et al; Gelfand et al. | mid 80% to mid 90% |
| 9) | *Canadian Land Cover Map* | AGR | National (Canada) | 2000 | 30m | Liu et al. (a) | overall class accuracy ~82% |
| 10) | *Land Cover Map 2000* | CEH UK | National (United Kingdom) | 2000 | mmu 0.5ha; 25$m^2$ | Lovett et al.; Tenerelli and Carver | 80–85% |
| 11) | *China land use* | CAS | National (China) | | 1:100k | Wu et al. | ? |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 12) | *China land use* | RSEDC | National (China) | 2000 | 1:100k | Lu et al. | ? |
| 13) | *China land use* | RSEDC | National (China) | 2008 | 1:100k | Zhuang et al.; Liu et al. (b) | ? |
| 14) | *Allegheny County greenways* | PASDA | Local | 2000 | variable, ~ 50ft | Niblick et al. | ? |

| ***Topography Datasets*** | | **Source** | **Geography** | **Date** | **Resolution** | **Used By** | **Threshold used** |
|---|---|---|---|---|---|---|---|
| 1) | *Global Terrain Slope (GTS)* | FAO & IIASA | Global | | 30" | Cai et al. | eight slope classes: 0-0.5%, 0.5-2%, 2-5%, 5-10%, 10-15%, 15-30%, 30-45%, and >45%. |
| 2) | *GAEZ* | FAO & IIASA | Global | | 5' | Milbrandt & Overend | includes slopes <8% and >30% |
| 3) | *Digimap* | EDINA | National (United Kingdom) | | 50m | Lovett et al. | excludes >15% |
| 4) | *USGS GTopo30* | USGS | Global | | 30" | Schweers et al. | excludes >25% |
| 5) | *SBSC* | SBSC | National (China) | 2008 | 1:100k | Liu et al. (b); Zhuang et al. | excludes >25% |
| 6) | *SBSC* | SBSC | National (China) | | 1:250k | Lu et al.; Wu et al. | Lu: excludes >25%; Wu: slopes >15% suitable, slopes 15% to 25% moderately suitable, exlcudes slopes >25% |
| 7) | *Unknown source* | ? | ? | | 1:250k | Fiorese & Guariso | excludes >20% |
| 8) | *STATSGO* | USDA-NRCS | National (USA) | | 1:250k | Gopalakrishnan et al. | includes >15% |
| 9) | *Digital Terrain Model (DTM)* | EDINA | National (United Kingdom) | | 1:50k | Tenerelli & Carver | unclear |
| 10) | *SSURGO* | USDA-NRCS | National (USA) | | 1:24k | Kang et al.; Niblick et al. | Kang: <8% envi-eco marginal, >30% phys. narginal; Niblick: excludes slopes >15% |
| 11) | *SRTM* | NASA | National (USA) | | 30m | Gelfand et al. | excludes >20% |

| ***Climate Datasets*** | | **Source** | **Geography** | **Date** | **Resolution** | **Used By** |
|---|---|---|---|---|---|---|
| 1) | *Temperature and precipitation* | Climatic Research Unit | Global | 1961-1990 | 2' | Cai et al. |
| 2) | *Temperature and precipitation* | Australian Bureau of Meteorology | National (Australia) | ? | ? | Odeh et al. |
| 3) | *Temperature and precipitation* | CMA | National (China) | ? | ? | Zhuang et al. |
| 4) | *Temperature and precipitation* | CMA | National (China) | 1971 -2000 | 1km | Lu et al. |
| 5) | *Temperature and precipitation* | CMA | Regional (SW China) | 1981-2010 | ? | Liu et al. (b) |
| 6) | *Temperature and precipitation* | ? | Regional (N. Italy) | ? | 1:500k | Fiorese & Guariso |
| 7) | *Temperature and humidity index* | CMA | National (China) | 1950 - 2000 | ? | Wu et al. |
| 8) | *UKCIP02 National Climate Scenarios: Growing degree days and precipitation* | UKCIP | National (United Kingdom) | 1961 -2000 | 5km | Tenerelli et al. |

| ***Land Capability Datasets*** | **Source** | **Geography** | **Date** | **Resolution** | **Used By** | **levels employed** |
|---|---|---|---|---|---|---|
| *Agricultural Land Classification* | DEFRA | National (United Kingdom) | ? | ? | Lovett et al. | 3 and 4 |
| *Land Suitability Rating System* | Agronomic Interpretations Working Group | National (Canada) | 1995 | ? | Liu et al. | 4 through 6 |
| *Agricultural Land Capability* | Natural England | National (United Kingdom) | ? | ? | Tenerelli and Carver | 3 through 5 |
| *SSURGO Land Capability Class* | USDA-NRCS | National (USA) | ? | 1:24k | Gelfand et al.; Kang et al. | 5 through 7; >4 |

| ***Protected Areas Datasets*** | **Source** | **Geography** | **Date** | **Resolution** | **Used By** |
|---|---|---|---|---|---|
| *Nature reserve* | RESDC | National (China) | 2000 | 1:100k | Lu et al. |
| *World Database of Protected Areas* | UNEP-WCMC | Global | 2009 | ? | Schweers et al.; Milbrandt & Overend |
| *Protected areas* | MAGIC | National (United Kingdom) | 2008 | ? | Lovett et al.; Tenerelli and Carver |

| ***Other Datasets*** | **Source** | **Geography** | **Date** | **Resolution** | **Used By** |
|---|---|---|---|---|---|
| *Polluted areas* | ARPAC | Italy | ? | ? | Fahd et al. |
| *Riparian and roadway buffers* | ESRI | National (USA) | ? | 10 to 50m | Gopalakrishnan et al. |
| *Brownfield sites* | NDEQ | Regional (Nebraska) | 2007 | ? | Gopalakrishnan et al. |
| *Contaminated water resources* | USEPA, USGS | National (USA) | 2002 | ? | Gopalakrishnan et al. |
| *Significantly irrigated lands* | USDA | National (USA) | 2002 | county-level | Gopalakrishnan et al. |
| *Climate change data* | OzCLim, CSIRO | National (Australia) | 2020, '30, '40, '50, '60, '70 | 25km | Odeh et al. |

**Table 2-3.** Summary of primary datasets used.

Among the 19 studies that used land cover as an input criterion, 14 different land cover datasets were used, and among the 18 studies to use soil, 16 different soil datasets were used (Table 2-3). Moreover, for "current" land cover datasets, publication dates span over a decade, ranging from the "end of last millennium" (released 1997) (IGBP DisCover) to 2008 (CDL).

*Socio-economic data inputs*

Several studies include what they term "socio-economic" factors – usually used as constraints in the model – however these data were often inconsistently utilized. In general "socio-economic" criteria consisted of varying aspects of land use / land cover, as decided by the scope and region of the study. For example, Lovett et al. (2009) incorporated the following "socio-economic" factors into their multi-criteria suitability analysis: urban boundaries and cultural heritage sites comprising doorstep greens, millennium greens, historic parks & gardens, monuments, registered battlefields, and world heritage sites. Wu et al. (2010) also incorporated "social-economic constraints" by excluding all but 2% of land used for food production and all but 50% of barren, grass and open forestlands from their mapped results. The only socio-economic constraint used by Lu et al. (2012) was shrub land, citing that this land use cannot be used for other purposes based on Chinese policy. In short, the socio-economic criteria employed were subjective and non-standardized. Despite the common use of the term "socio-economic", these factors may be better categorized as land cover/ land use criteria.

*Other data*

A small number of papers used other input criteria less common between studies. For example, polluted areas were considered in two of the papers reviewed here. Fahd et al. (2012) point out that some lands become marginal when excess pollution is generated by human-dominated processes (including illegal disposal of liquid and solid waste), therefore they used polluted areas as a primary input into their overlay model. In the United States Gopalakrishnan et al. (2011) included in their map of marginal land what they termed "environmentally degraded land," including land with brownfield sites, areas with water contamination, and areas with excessive irrigation.

Other less common input layers included highly localized datasets, like roadways and riparian corridors, which were mapped by Gopalakrishnan et al. (2011), but only qualitatively mentioned elsewhere (Kang *et al.*, 2013b). Projected changes in climate were infrequently used. Odeh et al. (2011) mapped marginal land in Australia at first by isolating annual precipitation to 300 – 600 mm/yr., and then made adjustments for climate change based on six emissions scenarios.

*Choice of data threshold*

Even amongst studies that used the same datasets for input criteria and exact same thematic classes, differences can be made when determining exact thresholds of delineation determining marginality. Topography (slope) is a prime example of how a threshold decided on by the authors or a "panel of experts" can vary widely between studies and influence mapped results. In the studies reviewed here, primarily slopes too steep for planting were excluded from the

analysis; however, exact thresholds can be subjective and therefore results differ considerably. For example, in their working definition of marginal lands, Lovett et al. (2009) *excluded* slopes >15%, while Gopalakrishnan et al. (2011) *included* slopes >15%. Milbrandt & Overend (2009) included slopes between 8 and 30%. Fiorese & Guariso (2010) and Gelfand et al. (2013) excluded slopes >20%, and four studies in China, including Schweers et al. (2011), Zhuang et al. (2011), Lu et al. (2012), and Liu et al. (2012b) excluded slopes >25%. Studies with numerous levels of marginality had multiple thresholds for slopes, including Kang et al. (2013a) who used a physical definition of marginal lands that excluded slopes >30%, and an environmental-ecological definition that excluded slopes <8%. Cai et al. (2011) identified eight slope classes for their fuzzy analysis.

*Choice of thematic class*

Considering data layers of the same subject and scale, there were often differences across studies in the thematic classes used in the analysis. The majority of studies used soil (81%) and land cover (86%) as input criteria to determine marginal land, but in non-standardized ways. For instance, certain land cover datasets sometimes included other less-common classifications specific only to those datasets, which make comparisons problematic. For instance, both Tenerelli and Carver (2012) and Lovett et al. (2009) used the Land Cover Map (LCM) 2000 (Table 2-3), but Lovett et al. (2009) only used the LCM2000 grassland classification yet Tenerelli and Carver (2012) used unique land cover classes including arable cereals, horticulture, improved grassland, and set-aside grassland that were exclusive to the LCM2000 (Fuller *et al.*, 2002). An Italian study by Fahd et al. (2012) used the CORINE (Table 2-3) land cover dataset to isolate non-irrigated arable lands. In the United States, Gopalakrishnan et al. (2011) used the 56m resolution 2007 USDA Cropland Data Layer (CDL) (Table 2-3) and targeted the thematic category "idle and fallow cropland" for their model, which they assumed includes Cropland Reserve Program (CRP) lands (i.e. set-aside lands). In sum, the categories listed above were not commonly available in all land cover datasets, therefore making cross comparison between studies difficult.

One of the most common soil-related input criteria in the studies reviewed here was land capability class (LCC). In the U.S., the Natural Resources Conservation Service (NRCS) describes LCC as, *"a system of grouping soils primarily on the basis of their capability to produce common cultivated crops and pasture plants without deteriorating over a long period of time"* (USDA, 2008). Several adaptations of this system have been used worldwide (Bibby & Mackney, 1969; Munn, 1986). Across datasets, LCC is derived from a combination of criteria, including erosion risk, soil depth, wetness, slope and climate (Klingebiel & Montgomery, 1961). Low LCC values designate prime agricultural land with few restrictions and higher values represent lands with increasing restrictions for agricultural production. In the studies reviewed here, LCC was most often used as a restrictive layer to exclude prime agricultural land from the analysis; however varying numeric classes were implemented. For example, in the U.K. Lovett et al. (2009) restricted marginal land to LCC levels 3 and 4 from the Agricultural Land Classification (ALC), in Canada Liu et al. (2012) restricted their analysis to Land Suitability Rating System (LSRS) classes 4 through 6, and in the U.S. Gelfand et al. (2013) restricted marginal lands to NRCS LCC levels 5 through 7, and in Italy, Tenerelli and Carver (2012) restricted their final mapped result to predetermined Agricultural Land Capability Map classes 3

through 5. In sum, none of the studies that implemented LCC into their model used the same thematic classes.

*Issues of data representation*

Finally, many of the GIS suitability models reviewed here assumed that input criteria exhibit crisp boundaries between what is marginal land what is not. However, as in the case of physical variables, like slope, the reality is often more continuous. Sometimes uncertainty in thresholds delineating marginal lands can be addressed in methods of analysis as a way to map input criteria in a continuous manner. For example, Tenerelli and Carver (2012) standardized their input criteria according to their compatibility with the ecological requirement of the crops. Based on expert knowledge, numeric input criteria (e.g. degree days, pH, rainfall and slope) were standardized using continuous benefit or cost functions, and nominal input criteria (e.g. qualitative values of soil depth and soil texture) were standardized using a ranking approach. Instead of binary thresholds, these methods employed more continuous grades of marginality that more accurately resemble the continuous nature of the input data.

*Issues of scale*

The papers reviewed here present work from global, national, regional and local scales. It is clear that the scale of analysis impacts the way in which the working definition of marginal land is structured, as it impacts model choice, data availability and selection, and resolution. For example, available datasets for *national* studies ranged in spatial resolution from 56m to 1km, while *global* datasets varied in resolution from 300m to 5 arc-minutes. In addition, different modes of scale were reported, including ratio (1:100k), pixel size (1km), and spatial resolution (30 arc-seconds), making direct comparisons between studies even more difficult.

In general, global datasets are available at a relatively coarse resolution and are also static in time – sometimes produced over a decade before the study using the data was conducted. The largest differences in land cover datasets were apparent in the global studies examined, where four global land cover datasets were used among the five global studies reviewed here. The analysis by Cai et al. (2011) used the International Geosphere-Biosphere Programme (IGBP) land cover dataset released in 1997 (Loveland *et al.*, 2000; Biradar *et al.*, 2009), which is available at 30 arc-seconds. Both Field et al. and Campbell et al. (2008; 2008) used both past land use and current land use (1700 to 2000) from the HYDE database at 5 arc-minutes (Goldewijk, 2001) to map abandoned pasture and cropland, along with a 2004 MODIS land cover product to mask current unsuitable land covers. Nijsen et al. (2012) used the Global Land Cover 2000 dataset (GLC2000) with a 1-km resolution (Bartholomé & Belward, 2005). Milbrandt and Overend (2009) used 2000 GAEZ land cover data available at 5 arc-minutes. In sum, of the "current" global land cover datasets used in the studies reviewed here, one was released in 1997, two have a publication date of 2000, and studies that used MODIS datasets had varied dates including 2004, 2006 and 2008. These examples illustrate the challenge global modeling has in acquiring timely data.

The other drawback of global studies is the coarse resolution of the datasets. Nijsen (2012) attempted to address this issue by downscaling the coarse GLASOD soils database (a database of

human-induced soil degradation from 1945 to 1990) from a minimum mapping unit of 5652 $km^2$ to 5 arc-minute scale, using several spatially explicit databases with finer resolution (Nijsen *et al.*, 2012). Other studies relied on currently available soil datasets, yet no two global studies reviewed here used datasets from the same source. Cai et al. (2011) used two datasets, the Harmonized World Soil Database (HWSD) available at 30 arc-seconds to get a soil productivity rating for each pixel and also 16 indices on soil temperature regime (STR) available from USDA-NRCS. Milbrandt and Overend (2009) used the GAEZ soils dataset with a resolution of 5 arc-minutes.

Global studies were not alone in these differences. There were also international differences between national datasets, as methodologies to generate national datasets vary by country. For instance, of the papers discussed here, many national studies used national-level soils datasets. Lovett et al. (2009) used the NatMap 1,000 database for their study in the United Kingdom; Odeh et al. (2011) used the CISIRZ dataset available for Australia, and both Lu et al. (2012) and Zhuang et al. (2011) used the Soil Map of China (RESDC) (Table 2-2).

Only one study reviewed here was conducted at a local scale. For their city-level analysis in Pittsburgh, Niblick et al. (2012) used very localized datasets, like commercial zoning at the parcel-level. Their analysis also employed a greenways feature class, which included agricultural easements, forested floodplains, designated greenways, land trust properties, rivers buffered by 100 ft., conservation streams buffered by 50 ft., sensitive slopes, wetlands 1 acre or more buffered by 50 ft., golf courses, parks and trails. This type of dataset better represents real world limitations to bioenergy plantations, however such data are not available for global-level analyses.

*Issues of uncertainty*

A complete accounting of error in any GIS analysis is important, yet not often completed. Issues of accuracy, error propagation and sensitivity are especially important because proponents of bioenergy production often cite these studies, more often quoting the upper bound rather than the lower bound of the mapped results (Fritz *et al.*, 2013). To best incorporate uncertainty and error in land suitability analysis, four components should be examined: 1) accuracy of input criteria, 2) validation of final results, 3) error propagation, or the compounding of errors in input datasets through the model, and 4) sensitivity of the model outputs to inputs. I reviewed the target papers for their attention to these issues. At least nine studies reviewed here failed to mention accuracy or uncertainty at all (Fiorese & Guariso, 2010; Kukk *et al.*, 2010; Wu *et al.*, 2010; Swinton *et al.*, 2011; Zhuang *et al.*, 2011; Liu *et al.*, 2012a, 2012b; Lu *et al.*, 2012; Gelfand *et al.*, 2013a). As these studies mapped crops that have yet to be planted, this is not surprising. However, there are extensive protocols developed in the GIScience discipline that outline methods of evaluating the quality of input criteria and understanding possible error propagation and model sensitivity (Arbia *et al.*, 1998; Crosetto *et al.*, 2000; Congalton, 2009).

Though most studies reviewed here do not cite the accuracy of their input datasets, metadata reveals reported accuracy of land cover datasets ranged from 67.1% (GlobCover) to 87% (CORINE) (Table 2-2). The accuracy of land cover datasets can have profound impacts on mapped results. Fritz and See (2008; 2005) highlighted thematic uncertainty as well as spatial

uncertainty in global land cover maps. When comparing GlobCover to MODIS land cover data, both input criteria used in the studies reviewed here, they found the combined forest and cropland disagreement to be 893 Mha (Fritz et al, 2011). As Field et al. (2008) pointed out, the MODIS land cover dataset does not distinguish between grassland and pasture—a potentially important distinction for planting bioenergy crops.

Regarding final mapped results, uncertainties were often addressed only qualitatively in the paper's discussion or conclusion. For example, Field et al. (2008) qualified the uncertainty in their results, saying that while "*the regional distribution of agriculture and pastures is relatively certain, the uncertainty for this abandoned area estimate is substantial (probably ± 50% or more).*" Likewise, Gopalakrishnan et al. (2011) qualitatively acknowledged that uncertainty arises from classification of marginal land and from using data layers at varying scales (scales of input data used in their analysis range from 10 m road and riparian buffers to soils data with a minimum mapping unit of 617 ha). Owing to the coarse resolution of inputs, some global-level analyses, including Field et al. (2008) and Campbell et al. (2008), added disclaimers to their mapped results, suggesting that general estimates of spatial distribution should not be prescribed at the local level.

Cai et al. (2011) assert that concerns of uncertainty are inherently addressed in their methods, saying *"FLM [Fuzzy Logic Modeling] is used to treat the uncertainty of the global data sets and the fuzzy nature inherent in land classification according to multiple criteria. FLM has been proven to be a powerful tool to address data variability, imprecision, and uncertainty and to treat the ambiguity and uncertainty involved in generating realistic continuous classifications."* However, it is still difficult to fully account for land cover uncertainty and global-level analyses such as this one still face a difficult obstacle when their results are compared to local-level analyses. In a recent study, mapped results from Cai et al. (2011) were calibrated by Fritz et al. (2013), who downgraded the 2011 estimates based on statistical adjustments derived from crowdsourcing Google Earth images. Their estimates reduced available land area between 264 and 376 Mha, depending on scenario, which might suggest that global studies overestimate land available for bioenergy production.

One study reviewed here verified their results with existing datasets that may be proxies for marginal lands. Kang et al. (2013a) compared their results to existing Conservation Reserve Program (CRP) lands, as well as Land Capability Classes 5 through 8, arguing that these designations have in the past been used to quantify marginal land. They argued that their hierarchical analysis was more comprehensive than just using LCC or CRP land alone, in part because their results showed more area being mapped. Another study by Schweers et al. (2011) conducted ground surveys to verify their mapped results in two regions, using GPS data, photos, interviews, and government data. They found that only 60% of the ground verification sites (out of 19 locations overall) totally agreed with the remotely sensed assessment. They also reported that using 2005 land cover data did not capture more recent land use changes as observed in their 2009 field survey, and that the resolution of the DEM (30 arc-seconds) used to derive slope was insufficient to reveal subtle nuances of the landscape.

Only one study quantitatively addressed both uncertainty and sensitivity. Tenerelli and Carver (2012) conducted an extensive sensitivity and uncertainty analysis, acknowledging that results

can be affected by both input and model errors, and that uncertainty propagates from the input criteria and parameters to the final output. To assess error propagation, their uncertainty analysis employed Monte Carlo simulations based on data accuracy. Their sensitivity analysis, which assessed how each input affects the model, involved two methods. The first removed each of the input criteria one-by-one through a jack-knife approach. The other was a sensitivity simulation on the criteria weights based on a Monte Carlo approach. The results showed which criteria are more or less influential to the mapped results, and how sensitive each criterion is to the input parameters, including applied weights.

**Discussion**

I examined a collection of recent literature that describes the mapping of bioenergy potential across space and scale. Projects that targeted second generation bioenergy feedstocks, focused on land that may be categorized as marginal, and used spatially-explicit GIS models were evaluated for commonalities and differences in methodology. While most papers provided similar general definitions for marginal land, their working definitions – those which are implemented in a spatial modeling framework – differed greatly. The concept of "marginal land" is often assumed to be static, yet this review suggests that the concept is better understood as relative: considered in proportion or in relation to something else. For example, many papers reviewed examined lands that are marginal when compared to agricultural lands. These lands might not be marginal when compared to their ability to provide wildlife habitat or other ecosystem services. Alternatively, many papers reviewed here were defined in response to what data a researcher can acquire, suggesting that difficult-to-map, but nonetheless important landscapes might be absent from the discussion.

I found no common working definition of marginal land across all of these studies, including considerable differences across models, input data, scales and validation methods. One country – China – provided a case study to examine potential comparisons. Despite the potential similarity of intent, the examples from China illustrated considerable differences in mapped results that were driven by differences in crop choice, model framework, data inputs, scale and treatment of uncertainty. These differences were echoed throughout the broader literature.

*Modeling framework*

The Geographic Information System framework is a useful and flexible one in which to perform suitability modeling. Most papers reviewed here employed relatively straightforward GIS overlays using linear combination of input criteria with distinct thresholds identifying land as either marginal or not. Models specifying varying degrees of marginality were less common (Wu *et al.*, 2010; Cai *et al.*, 2011; Lu *et al.*, 2012; Tenerelli & Carver, 2012). However, as this study shows, the thresholds determining marginality vary greatly between studies and are far from exact. Although simple GIS overlays have the benefit of being transparent and easy to replicate, they may not be the best measure of the dynamic concept of marginal land. And so there is a need to incorporate the natural continuity of datasets – either by standardization of input criteria or incorporation of Fuzzy Set Theory – in ways that more effectively represent this fluid subject.

*Data availability*

I found a large range in data choice for input criteria, therefore it was not surprising that data choices were rarely consistent across studies. For example, most studies across scales used some kind of land cover and soil dataset, yet I found 13 different land cover datasets and 16 different soil datasets were used. Additionally, differences in data thresholds, thematic class selection, and ways in which data is represented mean that analysis with identical data can yield different results. I also found that lack of appropriate data was an additional driver of differences between studies. For example, Kang et al. (2013a) initially identified 30 key variables for their analysis, however only 21 were applied because data for the other nine desired variables (eg. nutrient loss, biodiversity, resilience, resistance, buffer-zones or corridors, and placeholders for "other restrictions") were not readily available or easily quantified. Socio-economic factors were the least consistent category of input criteria, highlighting that standardization was especially lacking in this category.

*Scale of analysis*

I reviewed projects that focused on global, national, regional and local scales. The scale of analysis clearly impacted the way in which a working definition of marginal land is structured, as it impacts model choice, data availability and selection, and resolution. In some projects data matched the scale of analysis, in the manner that global studies used global datasets, but in some projects, there was a miss-match. Schweers et al. (2011) performed their analysis at a national scale, but because they were mapping *abandoned* lands they used global-level datasets. This makes results more comparable to global estimates, but less so to national ones.

National studies excel at addressing national policies and bioenergy targets. They also have the flexibility of either mapping for specific bioenergy species, like *Miscanthus*, or for a set of generic bioenergy crops, like perennial grasses or short rotation forestry. Input criteria for national studies are necessarily available at a national or global scale; therefore the data is generally more coarse than that available for regional studies, but can be of higher resolution than global studies. Datasets are often country-specific and are not internationally standardized.

Studies conducted at a regional scale are best suited for incorporating detailed datasets, including fine scale soils data (eg. the Soil Survey Geographic Database (SSURGO), 1:24,000) as well as road and riparian boundaries (Gopalakrishnan *et al.*, 2011). Regional studies also have the advantage of being able to incorporate socio-economic inputs at a more realistic level of analysis (Fiorese & Guariso, 2010). This allows regional studies to better address specific management goals and mapping for a specific bioenergy species. Methodologies for studies at regional scales vary from standard binary GIS overlays (Fahd *et al.*, 2012) to very complex analyses (Tenerelli & Carver, 2012). However, regional studies can be limited in that the results are often crop-specific and it is therefore difficult to extrapolate models to larger areas.

*Uncertainty*

Most of the papers reviewed here provided few details describing accuracy and uncertainty, suggesting great opportunity in the future for this kind of work. A review of the metadata

revealed that the accuracies of input land cover data ranged from 67.1% to 87%, yet these numbers were rarely reported in the works themselves and the impacts of propagation of error were seldom addressed. The results of accuracy and uncertainty analyses are especially important since the high end of the range of mapped estimates is often the most widely cited. Because these estimates have the potential to influence policy and real-world investments, it is especially important to ensure they are truly representative of the land resource availability. At the most basic level, studies can express these uncertainties by offering a range of mapped marginal land areas (as adjusted by soft constraints or accuracy analyses) that may be suitable for bioenergy crops instead of one exact number. However there is a still danger that only the highest value of the range estimate gets cited in future literature promoting bioenergy.

**Conclusions**

The challenges in planning for bioenergy mandates globally are large and numerous. Foremost among them is the need to determine where to plant bioenergy feedstocks to meet energy mandates while ensuring sustainable food production and environmental protection. The breadth of the projects reviewed here underpin the benefits of spatial modeling and GIS – even when simply implemented – in projecting where bioenergy might be planted. However, the considerable differences in definitions, models and data revealed in this review allow limited potential for comparison across studies, as well as for synthesis work that quantifies global biofuel potential. Some of the differences highlighted in this work might be minimized with standardization, through the use of similar datasets and similar analyses, and some might be minimized with better application of existing protocols, such as common accuracy assessment practices. Thus, the mapping of bioenergy potential is ready for a meta-analysis or shared examinations that use common data and protocols.

Challenges also remain in effectively addressing the bioenergy land use dilemma. The studies reviewed here individually and broadly reiterate the importance of refining theoretical estimates of bioenergy suitability with real world conditions that reflect explicit understanding of the balance between fuel, food and conservation. Lands that may be identified as marginal and mapped as such based on solely physical conditions may not necessarily be available for bioenergy plantation when considering economic, social, or environmental factors. However, incorporating these variables in a GIS environment is not always straightforward. Hard-to-depict land uses such as the presence of pastoralists, the use of land for cultural purposes, or for biodiversity protection are difficult to capture via remote sensing techniques (which excel in mapping land cover) and thus are not often included in broad-scale geospatial datasets. These mapping limitations make effective understanding of the tradeoffs between food, fuel and land elusive, but nonetheless tremendously important, and worthy of much more applied research.

This study is the first to systematically review projects that map bioenergy potential on marginal lands. My goal was to provide a framework to help researchers evaluate existing scholarship in mapping marginal land. I have identified areas where understandable differences in definitions, models, data and applications result in differences in mapped results. I suggest that there is tremendous future need for spatial modeling of bioenergy, yet further work should be done to allow for comparative work across countries and scales and understanding of cumulative global potential for bioenergy crops.

# Chapter Three

## A fuzzy logic-based spatial suitability model for drought-tolerant switchgrass in the United States

*Switchgrass (Panicum virgatum) has been targeted by the U.S. Department of Agriculture as an exemplary bioenergy crop, however it requires a significant amount of water and experiences reduced yields in water-stressed conditions. To avoid competition for prime agricultural areas, lands that receive adequate rainfall but are marginal due to highly variable timing of rain are potentially ideal locations to grow drought-tolerant biofuels. As scientists develop a modified variety of switchgrass that can withstand periods of drought while not substantially affecting overall yield, it is important to identify the potential geographical niche for this xerophytic crop to maximize its environmental and economic sustainability. This project uses a spatial suitability modeling approach that incorporates fuzzy logic and utilizes both physical and economic variables. I assess several fuzzy overlay techniques to identify and synthesize tradeoffs between suitability criteria. Our results highlight the Great Plains region of the United States as a suitable region, and within this area I focus on Kansas for a more detailed analysis to calculate land areas within varying dryness index thresholds. For this I develop a specialized dryness index using high resolution (spatial & temporal) weather and soil data to provide a spatially explicit measure of dry spell severity for switchgrass across a landscape. I estimate that 80% of the suitable land area in Kansas falls within a dryness index equivalent to about four 22-day long dry stretches, or one 45-day long dry stretch. By identifying the dryness threshold where land area is maximized, the results of this analysis inform the development of drought-tolerant varieties of switchgrass and identify marginal areas where efforts to plant such a species may prosper.*

### Introduction

As world-wide energy demand increases, there is growing pressure for renewable energy sources to help meet requirements and at the same time mitigate for climate change. While no single type of renewable energy will fulfill all of our energy needs, biofuels promise to become a growing sector of the global renewable energy market. In the 1980's the U.S. Department of Agriculture (USDA) recognized switchgrass *(Panicum virgatum*) as an exemplary mass bioenergy crop for several reasons, including its ability to grow on non-prime agricultural land. Since then many studies have illustrated the potential benefit of non-food, second generation biofuel crops, also known as advanced biofuels (Tilman *et al.*, 2006; Youngs & Somerville, 2012; Gelfand *et al.*, 2013b). As a result, a growing number of countries are incorporating mandates for second generation biofuels into policy. In 2007 the United States Congress passed the Energy Independence and Security Act (EISA) which creates a more aggressive Renewable Fuel Standard (RFS2), including a call to increase domestic production of 16 billion gallons of cellulosic biofuels per year by 2022 (Sissine, 2007). Out of a growing concern that such a large mandated sum of biofuels will compete for prime agricultural land, there is increasing pressure

to plant biofuels on marginal, abandoned, or otherwise non-agricultural land (Searchinger *et al.*, 2008). However, it is not straightforward how to locate such lands.

Spatial suitability modeling with Geographical Information Systems (GIS) is increasingly used as a technique to identify potential locations for renewable energy generation (Voivontas *et al.*, 1998; Ramachandra & Shruthi, 2007; Angelis-Dimakis *et al.*, 2011), including the placement of geothermal (Yousefi *et al.*, 2007; Noorollahi *et al.*, 2008), wind (Hansen, 2005; Rodman & Meentemeyer, 2006), solar (Janke, 2010) and bioenergy (Lovett *et al.*, 2009; Odeh *et al.*, 2011). The most common spatial models used in GIS-based suitability studies fall into two fundamental classes of multicriteria evaluation (MCE): Boolean overlay or Weighted Linear Combination (WLC) (Malczewski, 2004). In Boolean and WLC models, a common type of operation employs a distinct threshold of suitability. In Boolean overlay models, each criterion is classified into two subsets delineating whether or not a particular area is suitable. Criterion maps are then layered using logical connectives (i.e. AND, OR). One primary shortcoming of Boolean analysis is that criteria can only be TRUE or FALSE, which creates discrete boundaries between variables. This imposes artificial precision on mapped results and fails to model more nuanced degrees of suitability. In contrast, models that use WLC bin each criterion into categories. These categories are then weighted based on their importance as decided by professionals in the field, so when combined, one criterion with relative low suitability can be recompensed by the high score of another. Because binning still requires discrete thresholds, such models are also inherently rigid and do not best represent the real world (Jiang & Eastman, 2000).

Numerous studies identifying marginal lands for specific biofuel crops have used Boolean approaches using linear combination GIS overlays (Campbell *et al.*, 2008b; Gopalakrishnan *et al.*, 2011; Zhuang *et al.*, 2011). One Australian analysis by Odeh and colleagues (Odeh *et al.*, 2011) identified marginal lands for two biofuel crops based on an overlay of climatic and physical variables. Similarly, in the United Kingdom, Lovett and colleagues (Lovett *et al.*, 2009) identified marginal lands for Miscanthus with a GIS overlay of 11 physical and socio-economic factors along with sub-prime agricultural land classification classes. Although less common, at least one study mapping marginal land incorporated WLC into their model to quantify marginal land for agro-fuel crops in Italy by applying weights to the input criteria depending on level of importance as defined by the user (Tenerelli & Carver, 2012).

This paper adopts a different approach and employs a spatial suitability model based in fuzzy set theory (Zadeh, 1965). As compared to traditional overlay models, fuzzy logic better addresses data variability, imprecision, and ambiguity (Hall *et al.*, 1992). Unlike Boolean logic or WLC, fuzzy logic assigns a continuous membership value between 1 and 0 for each criterion. Typically, fuzzy logic values approaching 1 are considered more suitable and values approaching 0 are considered less suitable. In a fuzzy set, the concept of suitability, or 'membership,' is not definitive because all objects belong to the set in varying degrees. This approach better represents the continuous nature of biophysical and economic variables (Malczewski, 2004). The fuzzy spatial suitability model provides a formal framework to represent the uncertainty of where a threshold of suitability falls within a continuous landscape. In this way, fuzzy logic incorporates variation in expert judgment and provides more flexible classifications of suitability (Jiang & Eastman, 2000).

Recent spatial suitability studies that incorporate fuzzy set theory and GIS-based suitability modeling frameworks to identify ideal sites for renewable energy generation include Charabi and Gastli (Charabi & Gastli, 2011), who assessed land suitability for large Photovoltaic farms in the country of Oman with GIS-based spatial multi-criteria evaluation approach using Fuzzy Logic Ordered Weighted Averaging (FLOWA). Cai and colleagues (Cai *et al.*, 2011) conducted a global study targeted towards mapping marginal land for generic cellulosic biofuel species using MCE and fuzzy logic modeling. They first identified land with marginal productivity based on soil productivity, slope, soil temperature regime, and humidity index. Datasets were transformed into standardized layers using fuzzy membership functions based on empirical knowledge or expert opinion, and were then disaggregated to produce a final productivity index over which the team layered results with land cover to create four scenarios of land availability restriction. To the author's knowledge, the paper presented here is the first to use a fuzzy overlay approach specifically for mapping suitable areas for a specific biofuel species – switchgrass *(Panicum virgatum*).

In addition to locating enough land to meet bioenergy mandates, an added concern is that the frequency of extreme weather and drought has increased over the last century (Dai *et al.*, 2004; Regonda *et al.*, 2005; Wang *et al.*, 2010) and is expected to intensify with climate change (Karl & Trenberth, 2003; Mishra *et al.*, 2010; Wehner *et al.*, 2011). These changes make it increasingly important to grow drought-tolerant biofuels on dryer land areas not suited for crop production (Youngs & Somerville, 2012). Scientists at the University of California at Davis are currently developing a variety of switchgrass that is more drought-tolerant than current varieties (Ronald et al. *in progress*). At present, the precise response of switchgrass to water scarcity at its roots has been understudied. However, it has been shown that exposure to water stressed conditions reduces yields (Yi *et al.*, 2012) and that seven weeks of drought conditions significantly diminishes harvestable yields (Barney *et al.*, 2009). The theoretical benefit of the modified plant currently being developed for drought tolerance is that it can maintain the same yield while receiving the same amount of precipitation overall, but has the added advantage that it can withstand intermittent periods of drought where other varieties cannot. This paper investigates the potential land areas in the United States that might be able to sustain this crop.

For this reason, a new drought index was developed for the present analysis because existing indices deal with time periods too long for switchgrass growth or focus only on informing irrigation management decisions. For example, the most widely used drought indices, including the Palmer Drought Index (Alley, 1984) and the Koppen climate classifications (Peel *et al.*, 2007), are most effective in quantifying the severity of long-term droughts on the scale of months or years. Another technique capable of shorter time-scales, the Crop Water Stress Indicator, measures the amount of transpiration based on plant leaf temperatures to inform farmers when irrigation is required to maintain crop yields, but cannot be used to derive the length of dry spells in the past (Jackson *et al.*, 1981). The Water Deficit Index and other remote sensing-based techniques estimate crop water deficits based on inferred sensing and measurements of surface air temperatures (Moran *et al.*, 1994; Gu *et al.*, 2007). However, these methods are meant for irrigation intervention and do not enable quantification of the severity of past droughts. For these reasons, I developed a specialized dryness index using high resolution (spatial & temporal) weather and soil data to provide a measure of dry spell severity for

switchgrass, which is then combined with the results of the fuzzy spatial suitability model to quantify available land area within varying dryness thresholds.

The objectives of this paper are to: 1) develop a suitability map for switchgrass in the conterminous United States based on modeled potential yields, land value and land availability; 2) evaluate model sensitivity to fuzzy overlay techniques to identify target area for dryness study; and 3) develop a spatially explicit dryness index for the study area and calculate maximum land area within dryness thresholds.

## Methods

### *National Suitability Model*

#### *Primary Suitability Criteria*

The following variables were identified as primary suitability criteria for the explicit purpose of targeting the most fitting areas to plant drought-tolerant switchgrass.

First, modeled switchgrass yield data were available from Wullschleger and colleagues for a combination of both lowland and upland varieties (Wullschleger *et al.*, 2010). Their empirical analysis, based on over 1,190 field observations, is a function of precipitation, temperature, latitude, and fertilizer application. While only existing varieties of switchgrass were considered, where standard varieties do well, drought-tolerant varieties will also do well.

Second, I incorporate county-level market land value per acre from the 2007 USDA-NASS Census of Agriculture (USDA-NASS, 2007), for the reason that owners of these lands may be more likely to plant biofuels (Jensen *et al.*, 2007). In addition, lower land values on non-irrigated lands are significantly correlated with lower average annual precipitation (Mendelsohn & Dinar, 2003), therefore by identifying low-value lands I was able to target drier land areas preferred by the objectives of the study.

Finally, I include land area expiring from the USDA Cropland Reserve Program (CRP) by 2017 (USDA-FSA, 2011). The CRP is a volunteer program whereby farmers owning eligible croplands that are vulnerable to erosion or having other certain conservation benefits can option to take their land out of production in exchange for a rental payment for a contract period of 10-15 years. CRP land has been targeted as an ideal place to plant dedicated perennial biofuel crops, as growing perennial grass crops reduces erosion and has been shown to be carbon neutral, or even carbon negative (Perlack *et al.*, 2005; Gelfand *et al.*, 2011). Gelfand and colleagues conclude, “the direct use of existing CRP grasslands for cellulosic feedstock production would avoid carbon debt entirely and provide modest climate change mitigation immediately” (Gelfand *et al.*, 2011). For both environmental and economic reasons, it has been shown that farmers may choose to plant biofuels instead of re-enlisting in the CRP program (Jensen *et al.*, 2007; Gill-Austern, 2011). To prioritize land becoming available, I summed all acres scheduled to expire from CRP contract within each county between 2011 and 2017.

*Restrictive Layers*

Land uses considered inappropriate for planting biofuels for physical, economic, or environmental reasons were also identified and removed from further analysis. The 2011 30-meter Cropland Data Layer (CDL) was used to identify areas of exiting cropland, forest land, urban areas, water, snow/ice, and wetlands (Boryan *et al.*, 2011; Han *et al.*, 2012; USDA-NASS, 2012). Accuracy of cropland layers within the CDL varies by state, and accuracy of non-agricultural land areas in the CDL are based on the 2006 USGS National Land Cover Dataset (NLCD) (Wickham *et al.*, 2013). In addition to the CDL layers listed above, the national Protected Areas Database (PAD-US v1.2) identified areas of protected lands based on management intent to conserve biodiversity, as described by GAP Status Codes (USGS, 2011). GAP lands of Status Code 1 or 2 are permanently protected lands and were considered restrictive layers in this analysis. Lands in GAP Status 3 & 4 are less protected and potentially subject to biofuel production, therefore these lands were not included in the restrictive layers of this study. All restrictive layers were combined via linear combination into one "knock-out" layer where lands were not considered suitable for planting dedicated perennial grass biofuel crops.

*Resampling*

All primary suitability criteria and restrictive layers were resampled to the same resolution as the switchgrass yield dataset (1km). This scale was selected as a suitable compromise for generalization given the characteristics of the data sources. Raster and vector layers were converted to this resolution on a majority area basis using the ArcGIS Version 10.1 software (ESRI, 2012).

*Fuzzy Suitability Model*

*Standardized suitability criteria*

To address the continuous nature of the input criteria and their degree of suitability for switchgrass, raw data of the three primary suitability criteria were transformed using fuzzy membership transformation functions to create standardized suitability criteria. Fuzzy values were assigned in ArcGIS 10.1 (ESRI, 2012) by converting raw input values to a membership scale from 0 to 1 based on a transformation function defined by expert opinion. Values approaching 1 are more suitable and those approaching 0 are less suitable. For this preliminary analysis, linear relationships were employed. A positive relationship was applied to modeled switchgrass yields, as lands modeled with potential for higher yields are more suitable for planting switchgrass to be used for biofuels. Lands with modeled yields less than 3 tons per acre (6.57 Mg/ha) were considered non-suitable. A positive relationship was also applied to acreage of expiring CRP land, as more CRP land coming out of contract within a county suggests more area may be available for planting dedicated biofuel crops. A negative linear relationship was applied to market land value, as lower land values are more likely to be converted to biofuel crops. Land values greater than $2,500 per acre were assigned minimum suitability. This resulted in three final standardized suitability criteria representing modeled yield, expiring CRP land, and market land value.

*Model Sensitivity to Fuzzy Overlay*

To investigate the relationships between all factors, all standardized suitability criteria were combined into a final suitability map using a fuzzy overlay approach and multiple overlay techniques were tested. The overlay function "AND" selects the lowest value of all inputs (Eq. 1) whereas the function "OR" selects the highest value of all inputs (Eq. 2). The function "PRODUCT" multiplies all fuzzy inputs, so that the combined evidence is lower in value than any single input (Eq. 3). The function "SUM" linearly combines the inputs, so the combined evidence is more important than any single input (Eq. 4). Finally, the overlay function "GAMMA" multiplies the fuzzy sum by the fuzzy product to the power of gamma (Eq. 5). A gamma of 1 is closer to the fuzzy sum whereas a gamma value approaching 0 is closer to the fuzzy product (Bonham-Carter, 1994). Outputs were generated and tested for values of GAMMA in the range [0.00, 1.00] in 0.1 increments. See equations below:

Eq. 1 Fuzzy AND Value = $\min(arg_1, ..., arg_n)$
Eq. 2 Fuzzy OR Value = $\max(arg_1, ..., arg_n)$
Eq. 3 Fuzzy PRODUCT Value = $\text{product}(arg_1, ..., arg_n)$
Eq. 4 Fuzzy SUM Value = $1 - \text{product}(1 - arg_1, ..., 1 - arg_n)$
Eq. 5 Fuzzy GAMMA Value = $(\text{FuzzySum})^{\gamma} * (\text{FuzzyProduct})^{1-\gamma}$

*Focal Area Dryness Index*

Finally, the results of the nation-wide suitability model were used to guide where to next focus a more detailed analysis to quantify land area experiencing varying degrees of aridity. For this a dryness index model was developed.

In the processed-based dryness index model, soil moisture was calculated as a function of daily precipitation that increases the soil water level, which is capped by local field capacity, and evapotranspiration depletes the stored water in the soil. The index was in part based on a 4 km grid of daily precipitation and maximum and minimum temperature data from 1996 to 2005, produced from an integration of NOAA weather stations and PRISM data, provided by Schlenker (Schlenker & Roberts, 2006). The U.S. General Soil Map (STATGSO2) data on field capacity and water table depth available from USDA/NRCS were used to calculate the local soil storage capacity (NRCS-USDA). For each 4 km weather grid cell, irregular soil polygon data were averaged by spatial weighted averaging. Evapotranspiration was calculated by the Hargreaves equation. This method uses the maximum and minimum temperature, and extra-atmospheric solar radiation, which is calculated based on the latitude of each grid cell (Hargreaves, 1994). Based on current understanding, an evapotranspiration crop coefficient of 1, which is an average for corn and switchgrass, was assumed (Kang *et al.*, 2003) (personal communication, Dan Putnam, University of California, Davis).

Next the model identified extended periods of drought or "dry spells," defined as periods when the soil moisture was below the maximum allowable depletion for a given crop and root depth. At this point, a given crop's roots have difficulty extracting moisture from the soil thereby causing water deficit stress. Based on existing research, a maximum allowable depletion of 50% for switchgrass was assumed (personal communication, Dan Putnam, University of California,

Davis). Days in which soil moisture fell below this threshold were considered dry days, and periods of consecutive dry days were considered dry spells during which the soil moisture remains below this level. The dryness index for each 4 km grid cell is the sum of the squares of the 4 longest dry spells during the growing season (April - September) measured in days. Based on this index, longer spells were given a higher weight and multiple dry spells are included. The switchgrass yield losses reported in Barney et.al (2009) correspond to an index value of 1,225 using this drought index. Finally, the results from the fuzzy analysis were incorporated to calculate suitable and available land area in varying dryness thresholds within the focus region targeted by the suitability study described above.

## Results

*Suitability*

The standardized suitability criteria represent continuous degrees of suitability for each input criteria (Figure 3-1). Modeled switchgrass yields show the highest degrees of membership in the southeastern United States. Land market values (transformed to have lower values be more suitable) show higher membership west of the Mississippi owing to higher land values in the

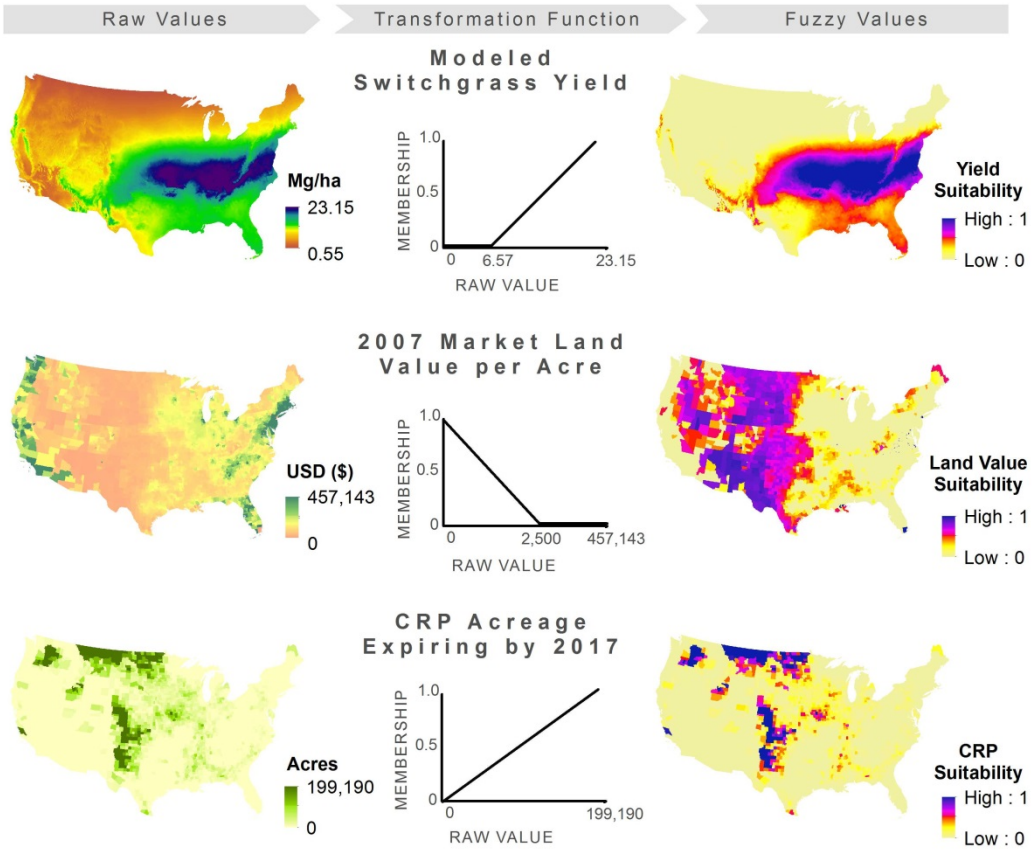



**Figure 3-1.** The standardized suitability criteria represent continuous degrees of suitability for each input criteria.

productive croplands and urban areas of the eastern United States. Finally, CRP acreage expiring by 2017 shows the highest membership in the Great Plain states, along the northern borders of Montana and North Dakota, and in eastern Washington.

Different fuzzy overlay functions portray the interaction of the membership of all criteria in significantly diverging ways (Figure 3-2). The "AND" overlay function selects the minimum value of the standardized suitability criteria, which gives emphasis to the most limiting suitability criterion. This highlights Kansas and north Texas as regions with comparatively minimum restrictions on suitability. The "PRODUCT" overlay function combines input criteria by cross multiplying the values of the standardized suitability criteria. This overly function is more restrictive than the "AND" function because the product of multiple standardized criteria (with values less than one) is lower than the value of any one standardized criterion. In this way, the "PRODUCT" overlay results in an output where the combination of multiple evidence has a lower suitability value than any of the inputs alone, but unlike the "AND" function is influenced by the values of the other criteria. The result of the "OR" function outputs the maximum value of the standardized suitability criteria. The "OR" function is comparatively liberal because it only considers the highest membership of a single criterion while ignoring any criteria with low membership value. The "SUM" function is also liberal because "SUM" is an increasing linear combination function, where the combination of multiple input criteria is larger than any of the inputs alone. In the case of the "SUM" overlay, unlike the "OR" function, the output is influenced by the combination of criteria, however only one criteria must have a high value for

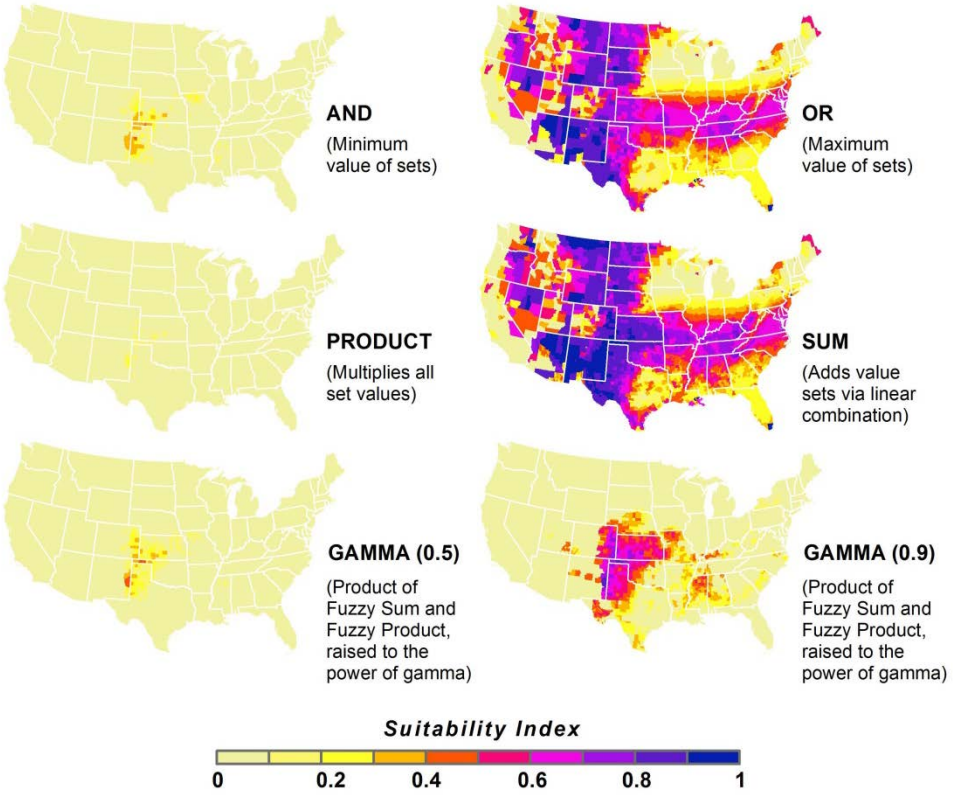



**Figure 3-2.** Different fuzzy overlay functions portray the interaction of the membership of all criteria in significantly diverging ways.

the overlay to display a high value. Both the "OR" and "SUM" overlay functions give substantial weight to criteria with high suitability without recognizing the limitation of criteria with low suitability in those areas. In these scenarios a single criteria with high membership drives the results despite the occurrence of other criteria with low memberships in the same area.

The "GAMMA" function, which combines both the fuzzy "PRODUCT" and the fuzzy "SUM", offers a way to balance multiple input criteria as needed to best represent suitability. This approach gives weight to limiting factors, but also allows more positive values of the other inputs to have more weight. In this way, the "GAMMA" function identifies gradient of suitability rather than a surface with discrete and unrealistic cutoffs.

*Fuzzy overlay sensitivity*

As mentioned above, the "GAMMA" function is a combination of two functions: the fuzzy-algebraic "PRODUCT" and fuzzy-algebraic "SUM". At low values of gamma, the "GAMMA" function is closer to the fuzzy-algebraic "PRODUCT", resulting in an output surface with values lower than any input in a given location. This is an 'all or nothing' scenario, where all suitability criteria must have high values for a pixel to be suitable and if any one criterion has low suitability, the entire pixel is classified as having low suitability. At high values of gamma, the "GAMMA" function is dominated by the fuzzy-algebraic "SUM" and is cumulative in character; that is, favorable input values result in an output that is larger than any of the inputs. This is an 'either or' scenario, where a pixel only needs one mapped criteria to be suitable for a pixel to be classified as suitable. Adjusting the range of "GAMMA" to be either closer to the fuzzy-algebraic "PRODUCT" or closer to the fuzzy-algebraic "SUM" enables the three suitability criteria to interact in converging or diverging ways.

For this analysis different values for "GAMMA" were tested on the standardized suitability criteria to determine the most suitable area. At gamma values below 0.35, the resulting suitability index can have values lower than any one of the input criteria, and for values between 0.8 and 1, it is possible to have a suitability index with values larger than any one input criteria. The histograms in Figure 3-3 show the frequency (y-axis) of pixels in a given suitability index (x-axis). Increasing the value of gamma shifts the distribution of the suitability index. As the gamma value approaches 0.9, the frequency of higher suitability values increases, the frequency of lower suitability values decreases, and the standard deviation increases. The result is a more equalized histogram that more clearly represents the suitability contrast between pixels. In contrast, the lowest values of gamma help to identify more precisely where all three criteria have reasonable memberships of suitability – low land values, expiring CRP land and reasonable switchgrass yields.

Figure 3-3 illustrates that the fuzzy overlay using the "GAMMA" function highlights a highly suitable area in the Great Plains States. Within this region, and in all scenarios, Kansas contains the highest percentage of available suitable land as compared to other adjacent states in the suitable area. For these reasons, Kansas was chosen for a more detailed state-level analysis to identify area of suitable land for biofuels under different drought index levels.

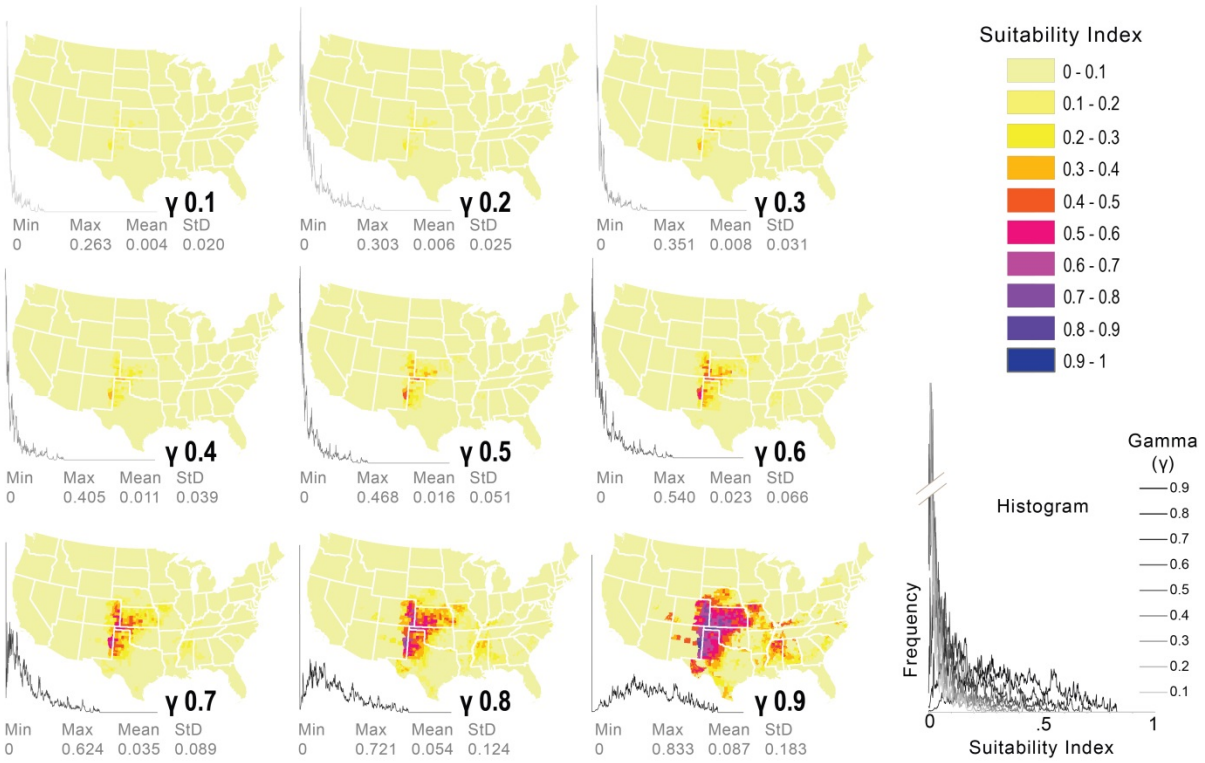



**Figure 3-3.** Increasing the value of Gamma (γ) shifts the distribution of the suitability index; as the gamma value approaches 0.9, the frequency of higher suitability values increases, the frequency of lower suitability values decreases, and the standard deviation increases

*Results from Focus Area*

Of the suitable, non-restricted land areas in Kansas that have a dryness index greater than 0, approximately 56% fall within a dryness index of 1,000 or below (equivalent to about four 16-day long dry stretches or one 32-day long dry stretch). That is to say more than half of the available land area in Kansas can support a switchgrass crop that will withstand a month-long drought. A dryness index value of 1,225 corresponds to yield losses of non-drought-tolerant switchgrass crops, as reported in Barney et.al (2009). Of the available lands in Kansas, 63% occur within this threshold. Roughly 80% fall within a dryness index of 2,000 or below (equivalent to about four 22-day long dry stretches, or one 45-day long dry stretch). And approximately 96% fall within a dryness index of 3,000 or below (equivalent to about four 27-day long dry stretches, or one 55-day long dry stretch). Beyond a dryness threshold of 2000, land area shows increasingly diminishing returns (Figure 3-4).

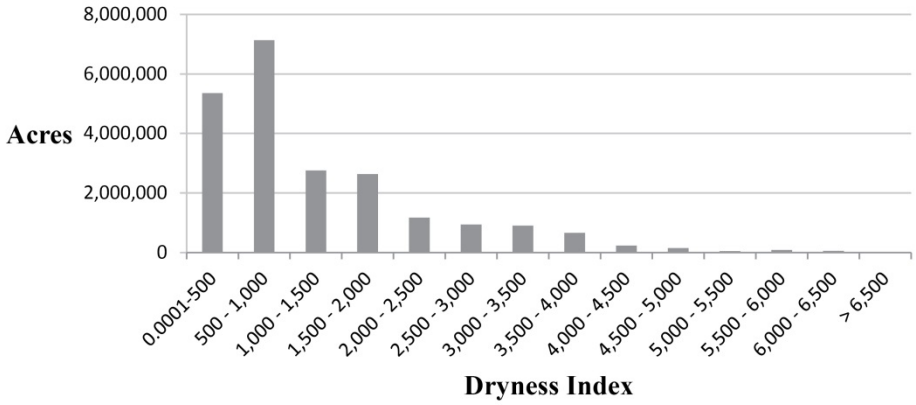



**Figure 3-4.** Roughly 80% of the available land area in Kansas falls within a dryness index of 2,000 or below (equivalent to about four 22-day long dry stretches, or one 45-day long dry stretch) under current climate conditions.

**Discussion**

New renewable energy mandates for second generation biofuels are challenging land change scientists to consider their location within current land use patterns. This study highlights the potential of fuzzy spatial suitability modeling to better represent the natural continuum of suitability across a landscape to guide placement of a specific biofuel species – switchgrass. The advantage of fuzzy spatial suitability modeling is that it provides more flexible weighing of evidence and combinations of evidence than the other more commonly used suitability models. In fuzzy theory, a parcel of land can have a partial membership of suitability that provides a more realistic measure of continuous factors as opposed to distinct thresholds. This has the advantage over other modeling techniques in studies where the boundaries of suitability are not precisely known by empirical analysis, as is often the case in new and expanding markets, such as biofuels. This preliminary broad-scale analysis allows us to identify highly suitable regions, like Kansas, where more detailed analyses can be focused.

Concerning fuzzy overlay methods, the "GAMMA" function provided the best combination of evidence. Other overlay methods gave too much weight to single variables at a given location while downplaying others. Comparatively, the "GAMMA" function best integrates the low and high memberships of multiple suitability criteria. Different values for gamma between 0 and 1 were tested on the standardized suitability criteria to generate a range of suitability maps, illustrating the effect of the gamma value on mapped results. Higher values of gamma resulted in more land mapped at higher values of suitability. Since the input criteria hadn't changed, those lands are not necessarily more suitable than they are at low values of gamma, however higher values of gamma do show more differentiation between levels of suitability, providing a superior comparative illustration of suitability across space. As switchgrass becomes a common biofuel crop in the United States, forthcoming research should incorporate empirical data to optimize gamma value for future suitability analyses.

Still, as with many site suitability studies, this interpretation should be taken with restraint. Results obtained by this model were dependent on data quality and the fuzzy membership values assigned to the input criteria. While based on expert opinion, assumptions on the range and slope of the fuzzy transformation functions affect model outputs, making the model more or less restrictive. The final output map may be used for planning purposes, however site-specific evaluation will be needed before the implementation of any biofuel crop plantation.

This study is especially applicable to plant science, as it presents a dryness threshold at which potential land area for production of a drought-tolerant crop is maximized. As follows, the results of this study can help guide research of drought-tolerant switchgrass by illustrating that over half (56%) of available land within Kansas falls within a dryness index threshold under 2,000 (equivalent to about four 22-day long dry stretches, or one 45-day long dry stretch). And approximately 96% fall within a dryness index of 3,000 or below (equivalent to about four 27-day long dry stretches, or one 55-day long dry stretch). Therefore the great majority of land in Kansas may be available to a switchgrass plant able to withstand a nearly two month long dry spell. This work will aid researchers in developing drought-tolerant biofuels in order to best maximize available planting area.

Future biofuels research will need to take variable rainfall into consideration, considering that scientists predict that climate change will significantly impact natural water cycles leading to an increased frequency of extreme weather events throughout the Midwest, including floods and droughts. When considering climate change and risk factors associated with crop production of marginal lands, it may be better to aim for higher drought-tolerance, but only to a point. Beyond a dryness index threshold equivalent to a 55-day long dry period, this study found diminishing returns on land available.

**Conclusion**

This paper demonstrates the ability of fuzzy spatial suitability modeling to highlight a continuous range of suitable areas to place a specific biofuel species – switchgrass – into the landscape. The results of the sensitivity analysis indicate that the GAMMA fuzzy overlay function best recognizes tradeoffs between combinations of multiple criteria. Within the suitable area Kansas was targeted for a spatially explicit dryness analysis. The majority of land within the state falls within a dryness index equivalent to an approximately 45 day long dry stretch. This information has implications for plant science, as it will help guide research on drought-tolerant varieties of switchgrass.

# Chapter Four

## Fuzzy GIS-based multi-criteria evaluation for U.S. *Agave* production as a bioenergy feedstock

*In the United States, renewable energy mandates calling for increased production of cellulosic biofuels will require a diversity of bioenergy feedstocks to meet growing demands. Within the suite of potential energy crops, plants within the genus Agave promise to be a productive feedstock in hot and arid regions. The potential distributions of Agave tequilana and Agave deserti in the United States were evaluated based on plant growth parameters identified in an extensive literature review. A geospatial suitability model rooted in fuzzy logic was developed that utilized a suite of biophysical criteria to optimize ideal geographic locations for this new crop, and several suitability scenarios were tested for each species. The results of this spatially-explicit suitability model suggest that there is potential for Agave to be grown as an energy feedstock in the southwestern region of the U.S. – particularly in Arizona, California and Texas – and a significant portion of these areas are proximate to existing transportation infrastructure. Both Agave species showed the highest state-level renewable energy benefit in Arizona, where agave plants have the potential to contribute 4.8 to 9.6% of the states' ethanol consumption, and 2.5 to 4.9% of its electricity consumption, for A. deserti and A. tequilana, respectively. This analysis supports the feasibility of Agave as a complementary bioenergy feedstock that can be grown in areas too harsh for conventional energy feedstocks.*

### Introduction

Increasing global energy demand, coupled with a changing global climate, necessitates a search for alternatives to energy production that promote environmental and economic sustainability. To this end, there has been substantial interest in the development of renewable bioenergy feedstocks. However, one significant obstacle facing these demands is land availability. In order to avoid conflicts with existing food production and to minimize use of water and other resources, effective biofuel production will rely on a diversity of bioenergy species, including feedstocks that may be both productive and sustainable on semiarid lands, thereby providing both economic and environmental benefits without impacting food production (Perlack & Stokes, 2011).

Species within the plant genus *Agave* have recently attracted attention as a potential complement to other bioenergy feedstocks (Somerville *et al.*, 2010; Davis *et al.*, 2011; Holtum *et al.*, 2011). Native to the hot, semi-arid regions of Mexico and the Southwestern United States, *Agave* spp. are naturally adapted to regions where conventional agricultural production is challenging, if not impossible. As perennial evergreen xerophytes, *Agave* spp. are capable of surviving in areas that experience long dry spells between rain events. The ability of *Agave* spp. to withstand such conditions is in large part due to their use of crassulacean acid metabolism (CAM) – a specialized form of photosynthesis, which allows the plant to keep leaf stomata closed during the hot day, thereby minimizing water loss through evapotranspiration and allowing them to achieve remarkable heat tolerance and water use efficiency (WUE) (Nobel *et al.*, 1992). This is in

contrast to $C_4$ feedstocks, like switchgrass and *Miscanthus*, which have a WUE that limits them to areas with relatively high annual rainfall. Even in these extreme environments, and with few nitrogen (fertilizer) inputs, agave plants produce yields comparable to other second-generation bioenergy feedstocks (Somerville *et al.*, 2010; Davis *et al.*, 2011).

Around the globe, commercial production of agave plants for the manufacturing of both fiber and tequila has been occurring for decades, and agronomic practices are well established (Somerville *et al.*, 2010). This existing knowledge of the production chain makes commercial production of agave plants for bioenergy even more feasible from an economic perspective. Yan et al. (Yan *et al.*, 2011) conducted a life cycle analysis and showed that *Agave* is likely to outperform corn and switchgrass in terms of ethanol output and net greenhouse gas emissions per unit land area. As such, *Agave* spp. represent a new opportunity for arid lands to be utilized for biofuel production. Although *Agave* spp. also grow well in areas with abundant rainfall and good soil water retention, in considering where to plant these species as a biofuel feedstock it is important to prioritize areas where limited rainfall prevents conventional agriculture (Davis *et al.*, 2011). To quantitatively scale the potential opportunity of *Agave* spp. as a feedstock and to minimize environmental and economic impacts, there is a need to determine the geographic potential for agave plantations with spatially explicit methods.

Geographical Information Systems (GIS) is a powerful tool for spatial suitability analyses involving numerous input datasets (Caver, 1991; Jankowski, 1995). GIS-based site suitability analyses are increasingly used to identify potential locations for integrating renewable energies into the landscape (Voivontas *et al.*, 1998; Angelis-Dimakis *et al.*, 2011). The majority of these analyses are based in Multi-Criteria Evaluation (MCE), the underlying principal of which is to synthesize complex problems with multiple variables. GIS-based MCE analyses traditionally rely on two main approaches: Boolean overlay or Weighted Linear Combination (WLC) (Malczewski, 2004). Boolean overlays consist simply of true and false designations given a specified threshold; an area is either suitable or it isn't. In WLC, weights are attributed to criteria according to the importance of each variable in defining the optimal solution. Both methods employ discrete thresholds to delineate suitability and are usually defined by the opinion of a panel of anonymous experts (Jiang & Eastman, 2000). There are inherently many assumptions in these models, and the association between empirical data and land suitability is often lacking. In addition, sensitivity analyses are not often performed nor are uncertainties about the results even addressed.

In contrast, fuzzy logic, as introduced by Zadeh (1965), allows for more flexibility in analyses where the boundaries between suitable and non-suitable are uncertain. The concept of fuzzy set theory involves classes with continuous grades of membership ranging from 0 to 1, with 0 representing the absolute falsehood and 1 being the absolute truth. In the case of suitability analyses, a value of 1 translates to an absolute truth that an area is suitable, values between 0 and 1 have a partial membership of suitability, and a value of 0 translates to a complete falsehood that an area is suitable.

A common argument within the literature is that fuzzy models allow for the inclusion of uncertainty in expert opinion and address continuous environmental factors that are not easily represented by discrete thresholds (Jiang & Eastman, 2000). However, even in fuzzy models,

expert decision makers still determine minimum and maximum input parameters, i.e., the 1's and 0's, which may impact the final mapped results. Because overlay functions within fuzzy models are often multiplicative, any input criterion with a 0 value immediately casts that location as unsuitable – despite the presence of several other suitable factors at that same location. The sensitivity of these models in regards to spatial suitability mapping has not been explored.

In addition to finding a suitable location, one of the biggest monetary expenses in the bioenergy production chain is transportation costs. Therefore proximity to infrastructure is essential to make biofuel projects economically sustainable. The costs and environmental impacts are substantially higher for road transport than for rail or ship, thus railroad networks are the most economically efficient mode of land transportation for bulk cargo [4]. However rail network density in the west is much sparser than in the east where biofuel feedstocks like *Miscanthus* are being targeted. In the case of agave plants, which can grow in marginal areas not currently being utilized by crop production, there may be gaps in the network where production potential is limited by distance to infrastructure.

To date the geographic potential for *Agave* as a biofuel feedstock in the United States is unknown. To explore this concept, I 1) develop a multi-criteria land suitability model based on fuzzy logic to site potential areas for two *Agave* species, *Agave deserti* and *Agave tequilana*, to be grown as biofuel feedstocks; 2) test the sensitivity of this model to small changes in input parameters; 3) construct multiple suitability scenarios adjusting input parameters to be more strict or optimistic based on extremes available in the empirical data; 4) integrate model results with existing transportation infrastructure to see where agave may be optimized in the landscape both physically and economically.

**Methods**

*Species Selection*

Two target species were identified for this work: *Agave tequilana and Agave deserti*. *A. tequilana* is a species with known commercial-scale operations and is well documented to produce high yields (24 Mg $ha^{-1}$ $yr^{-1}$) with limited inputs (Nobel & Valenzuela, 1987). The species has considerable photosynthetic plasticity in response to changes in temperature, light and water in addition to an ability to sequester a great amount of carbon (Pimienta-Barrios *et al.*, 2001). The second species reviewed here, *A. deserti,* is native to California and southwestern Arizona. It produces comparatively limited yields (7 Mg $ha^{-1}$ $yr^{-1}$) but is native to very arid climates and represents one extreme of the thermotolerance and drought tolerance spectrum of the *Agave* genus (Nobel & Hartsock, 1986). In addition to having a high tolerance to extreme heat, *A. deserti* also has a relative tolerance to cold temperatures – especially as compared to other *Agave* species.

Based on potential yield *A. tequilana* is the more likely crop of the two to be used for bioenergy purposes and may also be a better proxy for other high-yielding *Agave* species. In contrast *A. deserti* is less productive but shows some desirable characteristics for growing the crop in the U.S. (e.g. a wide temperature tolerance) and is therefore a useful comparative species to use in the model. These two species are among the most extensively studied in the *Agave* genus,

therefore the broad amount of literature makes them ideal test cases to use in this model. While other Agave species currently grown for fiber (eg. *A. sisalana* and *A. fourcryodes*) may also be potential candidates for bioenergy production, at the time of this publication there is not enough data on physiological responses to environmental stimuli for analysis of additional species in the present model.

*Suitability Model*

For this study, a geospatial suitability model was developed that incorporates fuzzy logic and utilizes a suite of biophysical variables to identify ideal geographic locations for this new crop. The following steps are outlined in Figure 4-1.

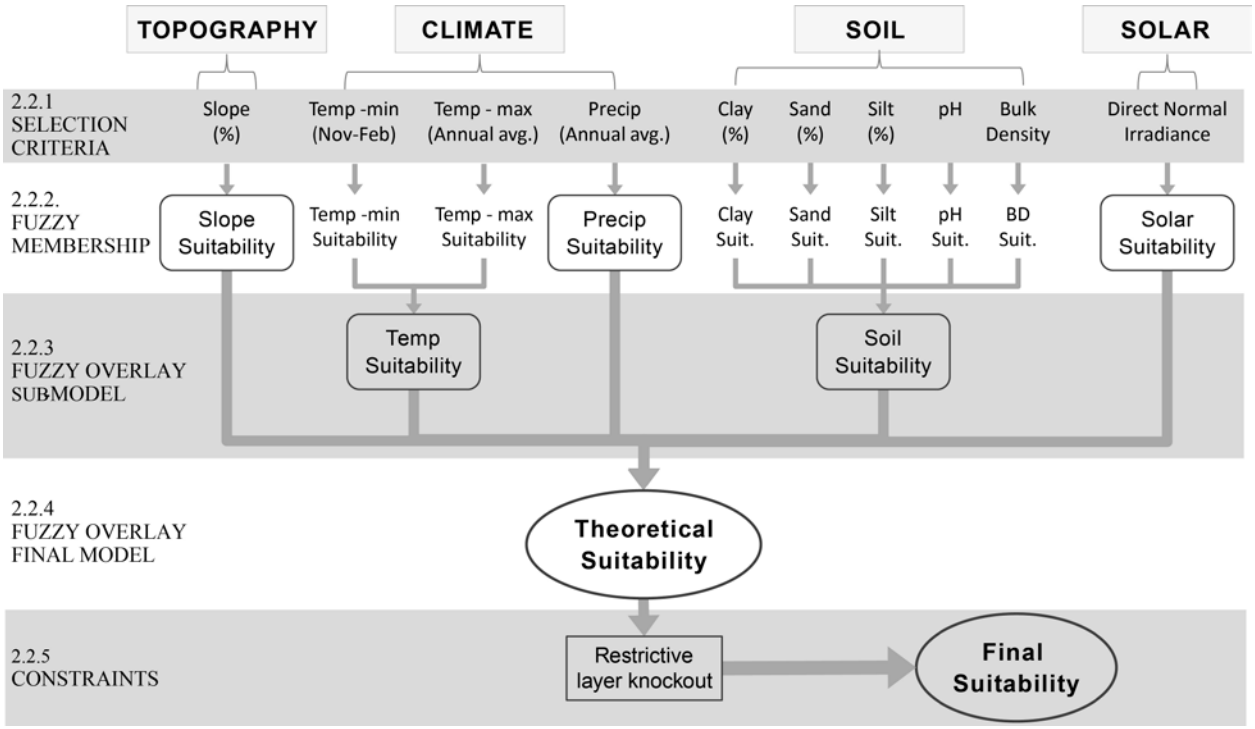



**Figure 4-5.** The famework for the suitability model incorporates fuzzy logic modeling in both the membership and overlay processes.

*Selection Criteria and Data Pre-processing*

Criteria determining *Agave* growth parameters were identified based on a review of the existing literature, and national-level geospatial datasets best emulating data from empirical studies were collected (Table 4-1). Data on climate conditions (years 1971-2000), including temperature (min, max) and precipitation were acquired from PRISM 30-arcsec (800m) (PRISM Climate Group, 2004). STATSGO soils data were acquired from (NRCS-USDA), including percent sand, percent

clay, percent silt, pH, and bulk density (1:250,000) (Miller & White, 1998). Solar data on Direct Normal Irradiance (DNI) was acquired from SUNY Albany and NREL (0.1dd X 0.1dd) (Perez *et al.*, 2002). Topography data (percent slope) from the HYDRO1k dataset was acquired from USGS/EROS (1km) (USGS, 1996).

| | *Data* | Source | Resolution | Description | Citation |
|---|---|---|---|---|---|
| ***Climate*** | *Tmin* | PRISM | 30 arcsec (~800m) | United States Average for Annual Minimum Temperature, 1971 - 2000 | (PRISM, 2004) |
| | *Tmax* | | | United States Average for Annual Maximum Temperature, 1971 - 2000 | |
| | *Precip* | | | United States Average for Annual Precipitation, 1971 - 2000 | |
| ***Soil*** | *Bulk Density* | CONUS - STATSGO - USGS NRCS | 1:250,000 (mmu ~2.3mi$^2$) | Bulk density is the ratio of the mass of soil to its total volume (solids and pores together) | (Miller & White, 1998) |
| | *Clay %* | | | Proportion of clay in the soil | |
| | *Silt %* | | | Proportion of silt in the soil | |
| | *Sand %* | | | Proportion of sand in the soil | |
| | *pH* | | | pH is a measure of the acidity or alkalinity of the soil | |
| ***Topography*** | *Slope* | HYDRO1k - USGS/EROS | 1km = 1000m | Maximum change in the elevations between each cell and its eight neighbors. The slope is expressed in integer degrees of slope between 0 and 90 | (USGS, 1996) |
| ***Solar*** | *Direct Normal Irradiation (DNI)* | SUNY Albany and NREL | 0.1 dd x 0.1 dd (~10km) | Annual average solar resource potential; data is shown in watt hours per meter squared per day | (Perez *et al.*, 2002) |

**Table 4-4.** General description of the research data.

Prior to model implementation, some additional pre-processing of the data was necessary to determine the best fit of the empirical data used to parameterize the model. First, soils data were estimated as the depth-weighted average in the top 20cm of soil – Layers ($L$) 1-3 derived from the STATSGO dataset (Eq. 1). This depth was chosen based on available field samples [*Gobeille et al.*, 2006; *Nobel and Valenzuela*, 1987].

Eq. 1 depth-weighted average $= \frac{L_1 * Depth_1 + L_2 * Depth_2 + L_3 * Depth_3}{Depth_1 + Depth_2 + Depth_3}$; where $Depth_1 = 5\ (0 - 5\ cm)$; $Depth = 5\ (5 - 10\ cm)$; $Depth_3 = 10(10 - 20\ cm)$

Second, monthly minimum temperature data for the months of November through February were combined to estimate the absolute minimum temperature during those months. Finally, all data were re-projected to NAD83 with appropriate datum transformations and then resized to match the extent & cell size of the PRISM grids (30 arcsec).

*Fuzzy Membership Transformations*

Raw data of the input suitability criteria were transformed into standardized suitability criteria by process of fuzzy transformation functions, which convert raw values (x-axis) into "fuzzy" values (0 to 1) (y-axis). A membership function expresses the degree of membership between 0 and 1.

Figure 4-2 illustrates an example of a trapezoidal shaped fuzzy membership transformation function used for the majority of the suitability criteria in this study (Pedrycz, 1994). This function assigns linear grades of membership (Eq. 1):

Eq. 1 $\mu(x) = 0 \text{ if } x < min, \mu(x) = 1 \text{ if } x > max,$

$$\text{Otherwise, } \mu(x) = \frac{(x - min)}{(max - min)}$$

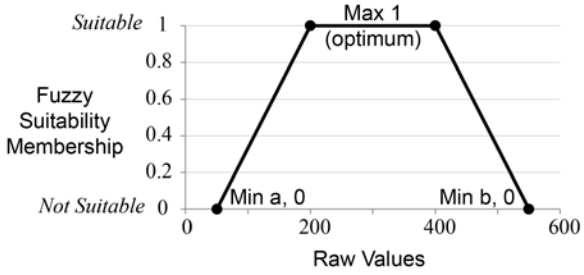



**Figure 4-6.** Fuzzy membership transformation functions used for the majority of the suitability criteria in this study transforms raw values of the input criteria to degrees of membership between 0 and 1.

Minimum and maximum input parameters were based on empirical information in the published literature. For *A. tequilana* suitable soil physical properties determined from the work of Nobel and Valenzuela (Nobel & Valenzuela, 1987) as well as Gobeille et al. (Gobeille *et al.*, 2006). Optimum temperatures for *A. tequilana* were determined from the work of Ruiz Corral et al. (Ruiz-Corral *et al.*, 2002), Cedeño (Cedeño, 1995), and Pimienta (Pimienta-Barrios *et al.*, 2001). The parameters for precipitation were determined from the work of Cedeño. For *A. deserti*, suitable soil physical properties were acquired from the work of Young and Nobel (Young & Nobel, 1986). Remaining criteria for *A. deserti* were acquired from the Agave Hill research center (Riverside County, CA; Lat/Long: 33.6386, -116.3983, elevation = 833m) for weather data collected between March 1973 and November 2011.

Non-linear membership functions were applied to two suitability criteria: slope and solar radiation. The Large function, which creates a sigmold shape where large values have high membership (Eq. 2), was used for the solar radiation criterion, as higher solar energy for *Agave* has a positive influence on yield (Ruiz-Corral *et al.*, 2002). For this study we applied a midpoint of 3.5 and a spread of 10.

Eq. 2 $\mu(x) = \frac{1}{1+\frac{x-f_1}{f_2}}$ where $f_1$ is the spread and $f_2$ is the midpoint.

The Small function, which creates a sigmoid shape where small values have high membership (Eq. 3), was used for the slope criterion, as I assumed commercial feedstock production becomes less and less suitable as the percent slope increases. For this study I applied a midpoint of 5 and a spread of 3.

Eq.3 $\mu(x) = \frac{1}{1+\frac{xf_1}{f_2}}$ where $f_1$ is the spread and $f_2$ is the midpoint.

*Fuzzy Overlay Sub-model*

Prior to the final overlay, it was necessary to first combine standardized suitability criteria within like-classes (soil and temperature criteria) to create categorical standardized suitability criteria. This was done using a fuzzy overlay function. The overlay function "Gamma" multiplies the fuzzy sum by the fuzzy product to the power of gamma (Eq. 6). At high values of gamma, the Gamma function is dominated by the fuzzy algebraic sum and is additive in character; that is, favorable input values result in an output that is larger than any of the inputs. A gamma of 1 is closer to the fuzzy sum (Eq. 4), whereas a gamma value approaching 0 is closer to the fuzzy product (Eq. 5). See equations below.

Eq. 4 $Fuzzy\ Sum\ Value\ =\ 1\ -\ product(1\ -\ arg_1, \ldots, 1\ -\ arg_n)$

Eq. 5 $Fuzzy\ Product\ Value\ =\ product(arg_1, \ldots, arg_n)$

Eq. 6 $Fuzzy\ Gamma\ Value\ =\ (FuzzySum)\gamma\ *\ (FuzzyProduct)1-\gamma$

The five standardized soils criteria (pH, bulk density and percent clay, sand and silt) were combined into a single standardized soil suitability criterion using a Gamma fuzzy overlay function with a gamma value of 0.9 (Lewis et al. in review). The standardized suitability criteria for temperature minimum & and maximum were also combined using the Gamma overlay function with a gamma value 0.9 into one final standardized temperature suitability criterion.

*Fuzzy Overlay Final Model*

All five final standardized suitability criteria, including slope, temperature, precipitation, soil and solar, were combined into a final suitability map using a fuzzy overlay approach. A Gamma fuzzy overlay function was used with a gamma value of 0.9. This resulted in values ranging from 1 to 0, with 1 being most suitable, and 0 being least suitable.

*Constraints*

Finally, land uses considered inappropriate for planting biofuels for physical, economic, or environmental reasons were also identified and removed from further analysis. The 2012 Cropland Data Layer (CDL) was used to identify areas of existing cropland, forest land, urban

areas, water, snow/ice, and wetlands (Boryan *et al.*, 2011). Urban areas and military installations available from the 2012 TIGER/Line Census were also excluded. These include boundaries of military installations from the U.S. Department of Defense (DOD) for Air Force, Army, Marine, and Navy installations and from the U.S. Department of Homeland Security (DHS) for Coast Guard installation. In addition, the national Protected Areas Database (PAD-US v1.2) was used to identify areas of protected lands based on management intent to conserve biodiversity, as described by GAP Status Codes (USGS, 2011). GAP lands of Status Code 1 or 2 are permanently protected lands and were considered restrictive layers in this analysis. Lands in GAP Status 3 & 4 are less protected and potentially subject to biofuel production, therefore these lands were not included in the restrictive layers of this study. All restrictive layers were assigned a suitability value of 0 (i.e. non-suitable) and combined via linear combination into one masking layer where lands were not considered suitable for planting dedicated perennial grass biofuel crops.

*Sensitivity Analysis*

To examine the sensitivity of the dependent outcome to moderate changes in the input parameters of each independent suitability criteria, a sensitivity analysis was conducted. These analyses were performed by varying the input parameters for the fuzzy membership transformation function of a single criterion while holding the parameters of other criteria fixed. One at a time, the parameters for a single suitability criterion were adjusted to be either 10% higher or 10% lower than the total range of the variable parameters identified in the literature review (Table 4-2 – original/moderate scenario).

Finally, a tornado diagram was constructed to summarize the individual impact of many independent criteria on the mapped results (Eschenbach, 1992). The suitability criteria tested include temperature minimum (tmin), temperature maximum (tmax), precipitation (precip), slope, solar, pH, bulk density, sand, silt, clay, and all soils combined. Each diagram demonstrates the impact of independently varying each of these fuzzy membership parameters. The base-case suitable land area of the moderate (i.e. original) suitability model (100%) defines the vertical axis of the diagram. From here, changes in percent mapped suitable land area owing to any one independent suitability criteria deviate from the base-case.

| | ***A. deserti*** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Moderate (Original)** | | | **Optimistic** | | | **Strict** | | |
| | Min a (0) | Max (1) | Min b (0) | Min a (0) | Max (1) | Min b (0) | Min a (0) | Max (1) | Min b (0) |
| **pH** | 3 | 6.3-8.1 | 9.1 | 3 | 6.3-8.1 | 9.1 | 5 | 7-8 | 9 |
| **BD** | 0 | 1.451-1.657 | 2.65 | 0 | 1.078-1.657 | 2.65 | 0 | 1.078-1.657 | 2.65 |
| **Clay (%)** | 0 | 5.75-6.21 | 70.16 | 0 | 5.75-6.21 | 70.16 | - | ≤6.21 | 40 |
| **Sand (%)** | 0 | 69.56-81.28 | 92 | 0 | ≥40 | - | 40 | 69.56-81.28 | 92 |
| **Silt (%)** | 0 | 16.29-18.37 | 82 | 0 | 16.29-18.37 | 82 | 0 | 16.29-18.37 | 82 |
| **tMin °C** | 0 | ≥5 | - | 0.1 | ≥4.4 | - | ≤4 | ≥8 | - |
| **tMax °C** | 18.5 | 22.5-27.4 | 29.9 | 9 | 13.8-35.1 | 39.7 | 18.5 | 22.5-27.4 | 29.9 |
| **Precip (mm $yr^{-1}$)** | 37.8 | 112.65-357.8 | 528.1 | 37.8 | 112.65-528.1 | 800 | 70 | 192.4-357.8 | 528.1 |

| | *A. tequilana* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Moderate (Original)** | | | **Optimistic** | | | **Strict** | | |
| | Min a (0) | Max (1) | Min b (0) | Min a (0) | Max (1) | Min b (0) | Min a (0) | Max (1) | Min b (0) |
| **pH** | 3 | 4.5-6.43 | 9.1 | 3 | 4.5-6.43 | 9.1 | 5;9 | 7-8 | 9 |
| **BD** | 0 | 0.68-1.54 | 2.65 | 0 | 0.68-1.54 | 2.65 | 0 | 0.68-1.54 | 2.65 |
| **Clay (%)** | 0 | ≥37.94 | - | 0 | ≥37.94 | - | 0 | 40-70 | 100 |
| **Sand (%)** | 0 | 12.72-40.53 | 92 | 0 | ≥40 | - | 0 | 12.72-40.53 | 92 |
| **Silt (%)** | - | - | - | - | - | - | - | - | - |
| **tMin °C** | ≤0.1 | ≥5 | - | ≤0.1 | ≥4.4 | - | ≤4 | ≥8 | - |
| **tMax °C** | 10 | 15-25 | 33 | 10 | 15-35 | 40 | 10 | 15-25 | 33 |
| **Precip (mm $yr^{-1}$)** | 250 | 500-1000 | 1200 | 250 | 500-1000 | 1200 | 600 | 800-1000 | 1200 |

**Table 4-5.** Model input parameters to fuzzy transformation functions.

*Scenarios*

In addition to the initial suitability model run, additional suitability scenarios were created for the *A. deserti* and *A. tequilana* suitability models, in which input fuzzy membership parameters were adjusted to be more or less restrictive. These parameters were based on physical extremes as observed in the available empirical data (Table 4-2). In one scenario, input parameters were altered to be more restraining than those in the original (termed the "strict" model). In another scenario, input parameters were altered to be more liberal than those in the original model (termed the "optimistic" model). As such, the original model represents a "moderate" scenario and will be referred to as such throughout the remainder of the paper. Not all parameters were adjusted for each scenario; solar radiation and slope remained unchanged for all scenarios. Table 4-2 displays the final parameters used.

Steps 2.2.2 through 2.2.5 in Figure 4-1 were repeated for each of the six scenarios in the sensitivity analysis, resulting in six mapped suitability results with values ranging from 1 to 0. To aid comparison between scenarios, based on the distribution of continuous suitability measure, suitable values were aggregated into four suitability classes: Very Suitable (VS) [1, 0.95], Suitable (S) [0.95, 0.85], Moderately Suitable (MS) [0.85-0.01], and Non-Suitable (NS) [0].

*Infrastructure*

Finally, the North American CTA Railroad Network (Peterson, 2012) was used to assess the proximity of the potential agave resources to existing infrastructure. This operational freight network only contains current operators and has a geographic accuracy of 100m. A buffering methodology that screens for regions surrounding active railroad nodes was utilized at distances of 25km and 50km. Areas of suitable land falling both within and outside of the designed buffers were calculated.

*Energy Estimates*

To estimate potential agave-derived ethanol production and energy requirements, including potential excess electricity that could be exported to the electric grid, first generation biofuels conversion assumptions were used. Recent U.S. corn ethanol conversion efficiencies were used to estimate the energy requirements for agave-based biorefineries (Perrin *et al.*, 2009). Davis et al. (Davis *et al.*, 2011) observed that approximately 1 liter of 40% ethanol is produced from 5.5 kg of dry agave biomass in modern large-scale tequila production facilities in Mexico. For simplification, here it was assumed that the sugar from the piña is the feedstock for fermentation into ethanol with the conversion efficiency observed by Davis et.al. (Davis *et al.*, 2011) and that both the bagasse (remaining solids from the piña after sugar extraction) as well as the harvested agave leaves are utilized in a Combined Heat and Power (CHP) unit in the biorefinery. Although bagasse and harvested agave leaves could be used in other power generation settings, for the present analysis it was assumed that an onsite CHP unit would provide the biorefinery process heat and electricity and that excess energy would be converted and supplied to the electric grid. It could be argued that the leaves of harvested agave plants should be returned to the soils for nutrients replenishment, however here I assumed that all of the harvested agave is utilized for energy production (as ethanol, heat, and power). Ethanol potential, energy consumption at biorefineries, and excess electricity supplied to the electric grid were calculated for *A. deserti* and *A. tequilana,* respectively.

**Results**

*Suitability Model Results*

The *A. deserti* suitability model results show California, Arizona, and Texas as potentially suitable areas for planting *Agave* as a biofuel feedstock (Figure 4-3). Results for the *A. tequilana* suitability model highlight Texas as a very suitable region in part because it has minimum land cover constraints. California and Arizona also show a substantial amount of suitable area, however they have significantly less area than Texas. Employing model constraints reduced and fragmented many of the potentially suitable land areas in California, particularly the Central Valley, which is dominated by cropland, along with the urbanized areas of Los Angeles and San Diego Counties. Large land tracts in Arizona remained less fragmented. Still, overall mapped suitable land area remained higher in California.

Looking at the current land cover in the three states with the most mapped suitable land area, in Arizona over 95% of the potentially suitable area is shrub land. Shrub land is also a significant land cover in the suitable areas of California (~40%) and Texas (~50%) but in these states grassland is also a substantial land cover within potentially suitable areas (~45% in California and ~13% in Texas). The present analysis does not, however, exclude pasture and grazing land. Texas was the only state to show pastureland in the suitable area, where ~20% of the land suitable area for *A. tequilana* is classified as pasture/hay.

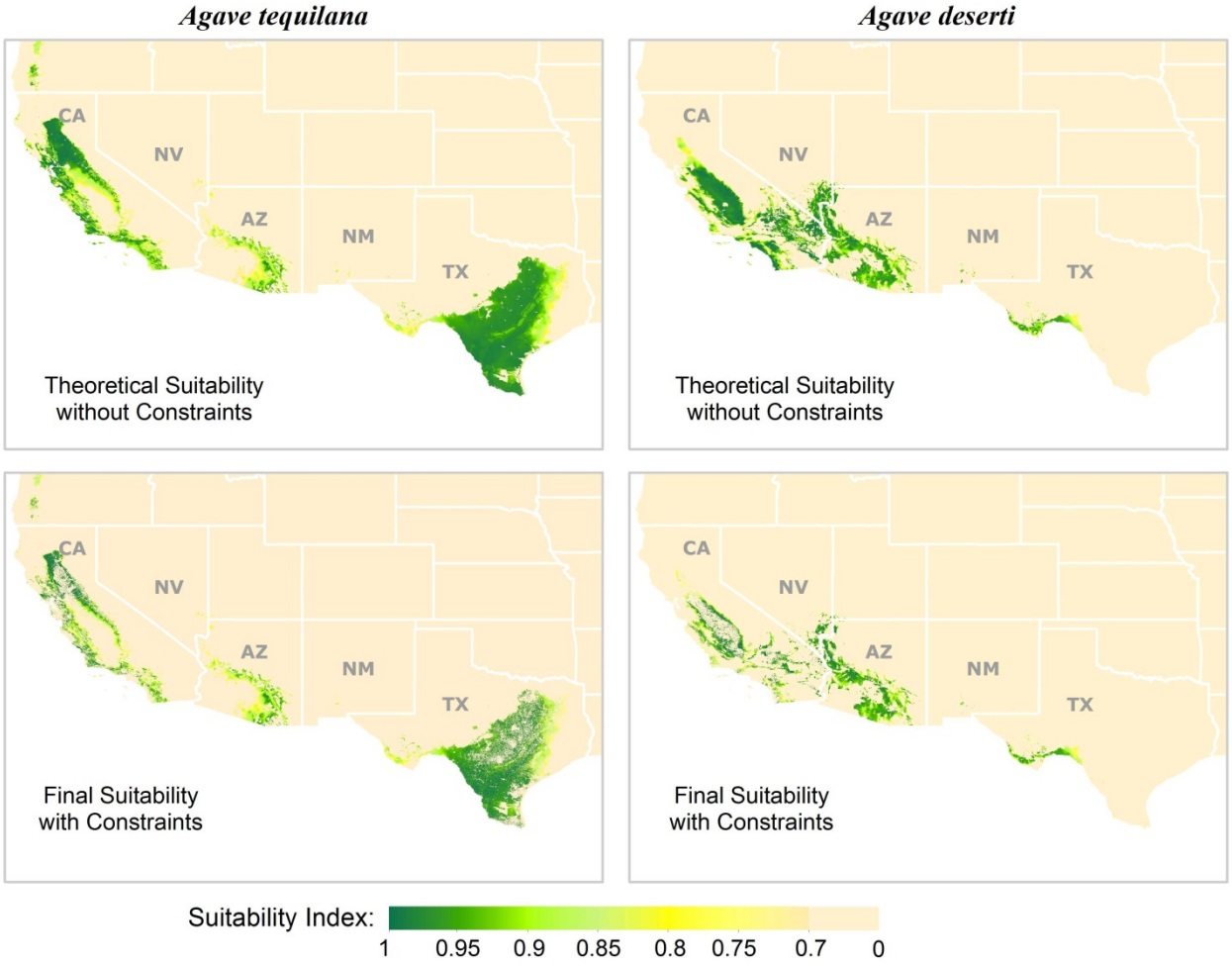



**Figure 4-7.** Results for the *A. deserti* suitability model highlight California, Arizona, and Texas as potentially suitable areas for planting agave plants as a biofuel feedstock.

*Sensitivity Analysis Results*

For both *A. tequilana* and *A. deserti*, adjusting the minimum temperature range had the greatest effect on suitable land area mapped as suitable (Figure 4-4). Decreasing the minimum optimum temperature increased the mapped suitable area by 76% in the case of *A. tequilana* and 40% in the case of *A. deserti*. Changing temperature maximum has little effect on mapped suitable land area for *A. tequilana*; however for *A. deserti*, increasing the optimum temperature maximum 10% of the range (1 °C) increases the mapped suitable land area by 12.7%. The effect of adjusting optimal precipitation has opposite effects on *A. tequilana* and *A. deserti*. For *A. deserti,* increasing optimal precipitation increased the mapped suitable land area by 19%, whereas for *A. tequilana* – a species which has higher water requirements – decreasing the amount of optimal precipitation increased suitable land area 6%. For *A. deserti* it is clear that climatic criteria, including temperature and precipitation have the biggest impact on mapped results.

As compared to temperature and precipitation, adjusting the fuzzy membership parameters for soil criteria had relatively little influence on area mapped as suitable. However, of the soil criteria, percent clay and percent sand had the most influence. For both *A. tequilana* and *A. deserti*, adjusting parameters for slope and solar variables had little influence on mapped results. Increasing the amount of solar radiation required for *A. deserti* reduced the mapped suitable land area by 7%, while the effect of changing the solar radiation input parameter had nearly no detectable effect on *A. tequilana*. Changing the parameters for the slope input criterion had hardly any detectable effect on either species (<1%). Figure 4-4 displays the combined results for all suitability classes: Moderately Suitable, Suitable and Very Suitable.

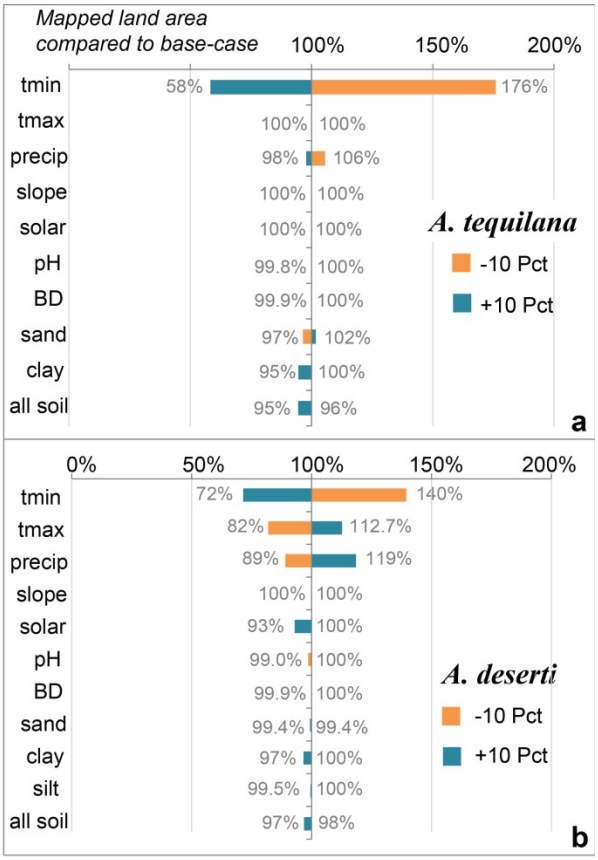



**Figure 4-8.** These tornado diagrams illustrate the impact of varying the parameters of input criteria in the model. The base-case suitable land area of the moderate (original) model (100%) defines the vertical axis; from here, changes in percent mapped suitable land area owing to changes in any one input criterion deviate from the base-case.

*Scenario Results*

As compared to the moderate (original) model parameters, both the optimistic and strict suitability scenarios generated increased and decreased mapped suitable land area estimates, respectively (Table 4-3, Figure 4-5). The optimistic parameters for *A. tequilana* increased the suitable area by half (150%) and the strict parameters decreased suitable area to 20% of the moderate scenario. For *A. deserti*, the optimistic scenario nearly doubled the mapped suitable area (197%) and the strict parameters decreased suitable area to 10% of the moderate scenario. In both the optimistic and moderate scenarios, the modeled suitability results for the two species - *A.tequilana* and *A. deserti* showed the most overlap primarily in southern Arizona as well as certain areas in southern California and Texas.

Climatic variables, including temperature (particularly average minimum winter temperature), and precipitation were the most influential suitability criteria in the model. Owing primarily to high precipitation and moderate minimum temperatures in the Pacific Northwest, the optimistic suitability model for *A. tequilana* showed suitability as far north as Oregon and Washington states. Limiting the minimum precipitation parameter from those in the moderate scenario from 250 mm $yr^{-1}$ to 600 mm $yr^{-1}$ in the strict scenario significantly reduced the amount of suitable area to primarily the southwestern region of Texas.

Although mapped suitable area differs considerably among suitability models, under all scenarios the model accurately predicts suitability for *A. deserti* in all counties identified by the Biota of North America Program where *A. deserti* is native and not rare. As these data were not used to drive model inputs, this comparison serves as an independent assessment (Kartesz, 2010). That being said, the mapped suitable areas in the strict scenario are not abundant, which might imply that this scenario is in fact more restrictive than in reality, and that the moderate mapped scenario more realistically depicts the potential distribution of *A. deserti*.

**Totals (Mha)**

| ***A. tequilana*** | **VS** | **S** | **MS** | **TOTAL** | **% Orig.** |
|---|---|---|---|---|---|
| Optimistic | 17.82 | 19.91 | 4.76 | 42.49 | 150.7% |
| Moderate | 13.56 | 8.97 | 5.67 | 28.20 | 100.0% |
| Strict | 0.00 | 1.46 | 4.23 | 5.69 | 20.2% |

**Totals (Mha)**

| ***A. deserti*** | **VS** | **S** | **MS** | **TOTAL** | **% Orig.** |
|---|---|---|---|---|---|
| Optimistic | 15.82 | 7.23 | 2.88 | 25.92 | 197.5% |
| Moderate | 5.11 | 5.80 | 2.22 | 13.13 | 100.0% |
| Strict | 0.02 | 0.47 | 0.84 | 1.33 | 10.1% |

**Table 4-6.** Estimated land available for *A. tequilana* and *A. deserti*, depending on model scenario. Very Suitable (VS), Suitable (S), Moderately Suitable (MS) land areas sum to the total estimated suitable land area.

*Proximity to Infrastructure*

Restricting suitable areas to 50- and 25-kilometer buffers around railroad nodes decreased the amount of total suitable area mapped by 92% to <1%, depending on species and scenario (Table 4-4). For example, as compared to the unrestricted moderate model for *A. tequilana* (28.2 Mha), suitable land constrained to the 50km buffer was reduced to 25.2 Mha – a difference in mapped land area of 11%. When restricted to a 25km buffer, the original land area mapped as suitable for

*A. tequilana* in the moderate scenario was reduced to 17.2 Mha – a 39% decrease. In the optimistic scenario, the suitable land area mapped decreased by 7% for the 50km buffer and to 46% for the 25km buffer. In the strict scenario, suitable areas only decreased by <1% for the 50km buffer, and <7% for the 25km buffer. Higher reductions were seen for the optimistic scenario because more land area was mapped as suitable.

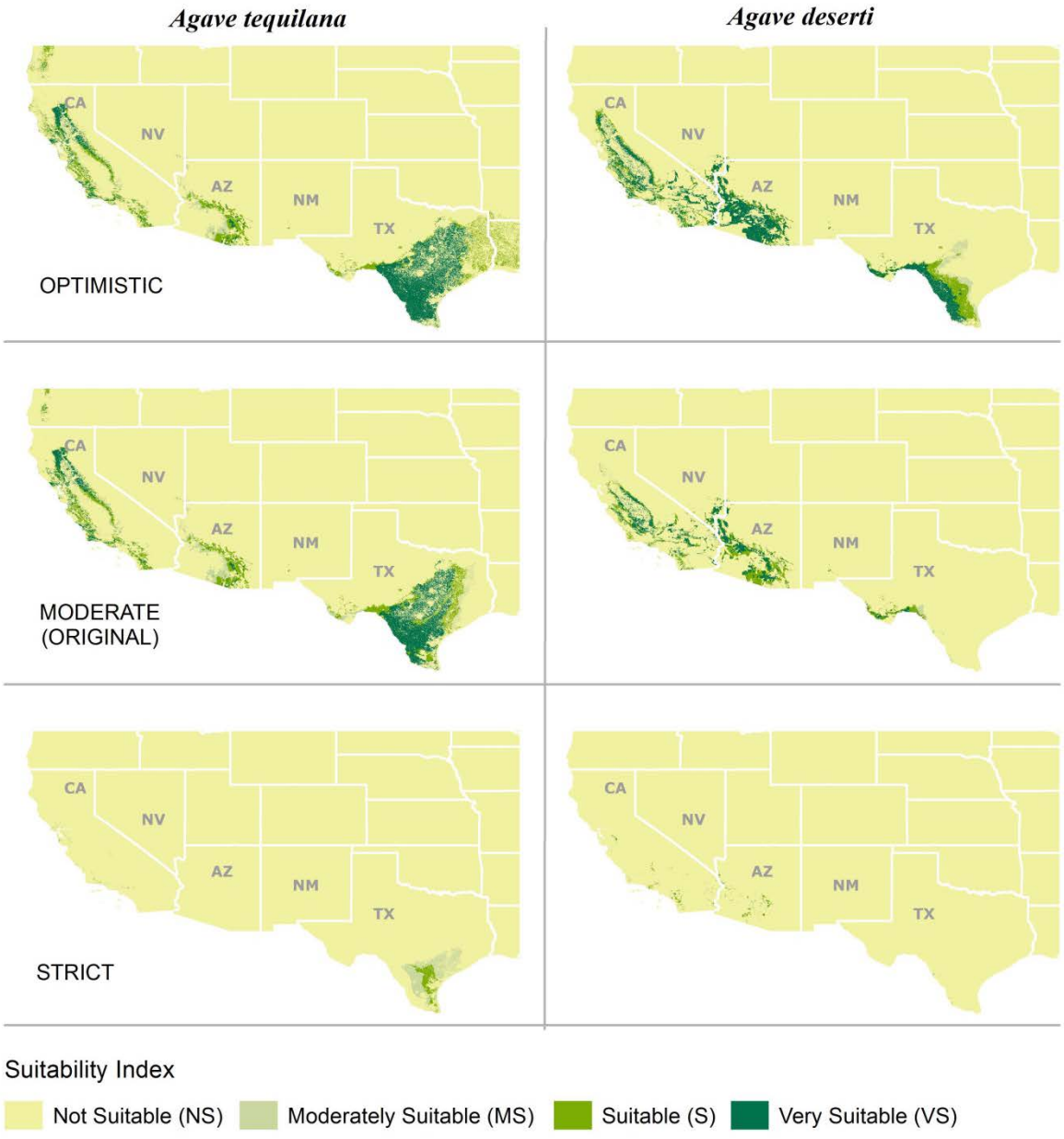



**Figure 4-9.** For both *A. tequilana* and *A. deserti*, California, Texas and Arizona are highly suitable. Both the optimistic and strict suitability scenarios generate increased and decreased mapped suitable land area estimates, respectively.

For *A. deserti*, the original land area mapped as suitable in the moderate scenario (13.13 Mha) decreased 14% (11.26 Mha) when restricted to a 50km buffer, and 48% (6.86 Mha) when restricted to the 25km buffer. For the optimistic scenario, the suitable land area mapped decreased by 26% for the 50km buffer and to 92% for the 25km buffer. In the strict scenario, suitable areas only decreased by 1% for the 50km buffer, and 4% for the 25km buffer.

**Totals, Mha (50km buffer)**

| *A. tequilana* | VS | S | MS | TOTAL | % ORIG |
|---|---|---|---|---|---|
| Optimistic | 17.82 | 19.05 | 3.68 | 40.55 | 143.78% |
| Moderate | 12.39 | 8.12 | 4.68 | 25.19 | 89.31% |
| Strict | 0.00 | 1.43 | 4.08 | 5.52 | 19.56% |

**Totals, Mha (50km buffer)**

| *A. deserti* | VS | S | MS | TOTAL | % ORIG |
|---|---|---|---|---|---|
| Optimistic | 13.82 | 6.24 | 2.46 | 22.52 | 171.59% |
| Moderate | 4.62 | 4.68 | 1.96 | 11.26 | 85.76% |
| Strict | 0.02 | 0.41 | 0.76 | 1.19 | 9.06% |

**Totals, Mha (25km buffer)**

| *A. tequilana* | VS | S | MS | TOTAL | % ORIG |
|---|---|---|---|---|---|
| Optimistic | 11.45 | 15.75 | 2.28 | 29.48 | 104.52% |
| Moderate | 8.42 | 5.60 | 3.18 | 17.20 | 61.00% |
| Strict | 0.00 | 0.93 | 2.77 | 3.70 | 13.13% |

**Totals, Mha (25km buffer)**

| *A. deserti* | VS | S | MS | TOTAL | % ORIG |
|---|---|---|---|---|---|
| Optimistic | 8.70 | 3.60 | 1.53 | 13.82 | 105.30% |
| Moderate | 2.89 | 2.74 | 1.23 | 6.86 | 52.25% |
| Strict | 0.02 | 0.27 | 0.51 | 0.80 | 6.11% |

**Table 4-7.** Estimated land available within buffers around railroad nodes, depending on species and scenario. Very Suitable (VS), Suitable (S), Moderately Suitable (MS) land areas sum to the total estimated suitable land area.

At the state level, when the mapped suitable land area for *A. tequilana* was restricted to land areas within the 25km infrastructure buffer, the majority of the originally mapped suitable land areas in the moderate scenario (15.58 Mha out of 17.20 Mha) occurred within Texas (9.99 Mha), California (3.65 Mha), and Arizona (1.94) (Table 4-5). However, while in both the moderate and optimistic scenarios, California and Arizona have a substantial amount of land potentially suitable for *A. tequilana* (4.61 Mha and 1.94 Mha in the optimistic scenario, respectively), in the strictest scenario, nearly all suitable land area for *A. tequilana* is limited to Texas (3.47 Mha within the 25km buffer and 5.23Mha within the 50km buffer).

The results for *A. deserti* at the state-level show that the majority of the originally mapped suitable land areas in the moderate scenario was restricted to the 25km infrastructure buffer area (6.44 Mha out of 6.86 Mha) within California (3.56 Mha), Arizona (2.27 Mha) and Texas (0.61 Mha) (Table 4-5). In the optimistic scenario, Texas shows more mapped suitable land area than Arizona (4.35 Mha compared to 3.37 Mha, respectively). However with more strict parameters employed, only California and Arizona remain as suitable areas (0.49 and 0.30 Mha, respectively) and the suitable land in Texas drops to 0.01 Mha.

| | 25km buffer | | | 50km buffer | | |
|---|---|---|---|---|---|---|
| ***A. tequilana*** | **Optimistic** | **Moderate** | **Strict** | **Optimistic** | **Moderate** | **Strict** |
| **California** | 4.61 | 3.65 | 0.24 | 6.62 | 6.28 | 0.29 |
| **Arizona** | 1.94 | 1.94 | 0.00 | 3.47 | 3.47 | 0.00 |
| **Texas** | 11.63 | 9.99 | 3.47 | 16.54 | 14.50 | 5.23 |

| | 25km buffer (Mha) | | | 50km buffer (Mha) | | |
|---|---|---|---|---|---|---|
| ***A. deserti*** | **Optimistic** | **Moderate** | **Strict** | **Optimistic** | **Moderate** | **Strict** |
| **California** | 5.66 | 3.56 | 0.49 | 8.05 | 5.09 | 0.61 |
| **Arizona** | 3.37 | 2.27 | 0.30 | 5.92 | 4.38 | 0.53 |
| **Texas** | 4.35 | 0.61 | 0.01 | 7.96 | 1.22 | 0.04 |

**Table 4-8.** State-level results for land area mapped as suitable within the 25- and 50km buffers.

*Yield Estimates*

Assuming yields of 24 Mg $ha^{-1}$ $yr^{-1}$ for *A tequilana* (Nobel & Valenzuela, 1987) and 7 Mg $ha^{-1}$ $yr^{-1}$ for *A. deserti* (Nobel & Hartsock, 1986) and additional conversion and production assumptions outlined above, Tables 4-6 and 4-7 estimate the ethanol potential, energy consumption at biorefineries, and excess electricity supplied to the electric grid for *A. tequilana* and *A. deserti* respectively. The share of ethanol production, and consequently electricity and fuel consumption and exported electricity between states are also presented in Tables 4-6 and 4-7.

Based on the original mapped suitable land area for *A. tequilana* in the moderate scenario within the 50km transportation buffer, there is a production potential of 8,793 Ml of ethanol, an electricity consumption of 1,324 GWh, a fuel consumption of 64,454 TJ, and an exported electricity potential of 26,370 GWh (Table 4-6). The results for *A. deserti*, in the moderate scenario within the 50km buffer show a production potential 1,146 Ml of ethanol, an electricity consumption of 173 GWh, a fuel consumption of 8,401 TJ, and an exported electricity potential of 3,437 GWh (Table 4-7). Considering that in 2012, the U.S. produced 13,300 million gallons of ethanol (50,346 Ml) (Renewable Fuels Association, 2013), these numbers *for A. tequilana* calculate out to 28.1% of total U.S. production in the optimistic scenario within 50km of rail infrastructure, 17.5% in the moderate scenario, and 3.8% in the strict scenario.

The state-level results follow the trend of mapped suitable land area. For *A. tequilana*, in the strictest scenario, potential suitability is limited only to Texas, while California and Arizona are also significant contributors in moderate and optimistic scenarios. The reverse is seen for *A. deserti.* In the optimistic scenario, Texas joins California and Arizona as a potentially significant contributor of ethanol and energy from *A. deserti,* but in the moderate and strict models, production is limited to California and Arizona.

| *A. tequilana* | | Transportation Buffer 50 km (25 km) | | | |
|---|---|---|---|---|---|
| **Region** | **Scenario** | **Ethanol Ml (a)** | **Electricity Consumption (GWh) (b)** | **NG Consumption (JT) ( c)** | **Exported Electricity (GWh) (d)** |
| Total | Optimistic | 14,156 (10,291) | 2,132 (1,550) | 103,767 (75,432) | 42,454 (30,861) |
| | Moderate | 8,793 (6,005) | 1,324 (904) | 64,454 (44,020) | 26,370 (18,010) |
| | Strict | 1,926 (1,293) | 290 (195) | 14,118 (9,475) | 5,776 (3,876) |
| California | Optimistic | 16% (16%) | | | |
| | Moderate | 25% (21%) | | | |
| | Strict | 5% (6%) | | | |
| Arizona | Optimistic | 9% (7%) | | | |
| | Moderate | 14% (11%) | | | |
| | Strict | <1% (<1%) | | | |
| Texas | Optimistic | 41% (39%) | | | |
| | Moderate | 58% (58%) | | | |
| | Strict | 95% (94%) | | | |

**Table 4-9.** State-level potential energy impacts for *A. tequilana.*

| *A. deserti* | | Transportation Buffer - 50 km (25 km) | | | |
|---|---|---|---|---|---|
| **Region** | **Scenario** | **Ethanol Ml (a)** | **Electricity Consumption (GWh) (b)** | **Fuel Consumption (TJ) ( c)** | **Exported Electricity (GWh) (d)** |
| Total | Optimistic | 2,293 (1,407) | 345 (212) | 16,809 (10,315) | 6,877 (4,220) |
| | Moderate | 1,146 (698) | 173 (105) | 8,401 (5,118) | 3,437 (2,094) |
| | Strict | 121 (82) | 18 (12) | 887 (598) | 363 (245) |
| California | Optimistic | 36% (41%) | | | |
| | Moderate | 45% (52%) | | | |
| | Strict | 51% (61%) | | | |
| Arizona | Optimistic | 26% (24%) | | | |
| | Moderate | 39% (33%) | | | |
| | Strict | 44% (37%) | | | |
| Texas | Optimistic | 35% (31%) | | | |
| | Moderate | 11% (9%) | | | |
| | Strict | 3% (1%) | | | |

**Table 4-10.** State-level potential energy impacts for *A. deserti.*

(a) Assumes 1 litter of 40% ethanol per 5.5 kg of dry agave biomass
(b) Assumes 0.151 kWh/l
(c) Assumes 7.33 MJ/l

(d) Assumes 4.324 j/g of total harvested agave biomass is available to the CHP with a 60% boiler efficiency and 40% Rankine cycle efficiency for conversion to electricity. Thermal energy from the CHP first supplies necessary thermal energy to the biorefinery and excess energy is converted to electricity. Biorefinery electricity demand is satisfied and the remainder is exported to the electric grid.

*State-level Demands for Renewable Fuels & Electricity*

In addition to United States federal mandates such as the Renewable Fuels Standard, state-specific transportation fuels programs and renewable portfolio standards for electricity promote the development and use of renewable energy. For this reason, the yield results estimated here were compared with specific state transportation fuel and electricity demands. Table 4-8 shows the relative contribution that agave-based biofuels and resulting electricity exports could contribute to current light duty fleet transportation fuels (gasoline) and electricity consumption in each of the three states. For example, in the optimistic scenario, the amount of *A. tequilana* processed in the 50km transportation buffer would provide 2.68% of California's 2010 state electricity consumption, and 2.8% of its gasoline consumption.

*A. deserti* has the highest state-level impact in Arizona, where it has the potential to produce as much as 4.8% of the state's gasoline consumption, and 2.5% of its electricity consumption. *A. tequilana* may also significantly contribute to Arizona's gasoline and electricity consumption – potentially contributing 9.6% and 4.9%, respectively at the highest levels. *A. tequilana* may also significantly contribute to Texas' gasoline and energy consumption – potentially contributing 7.8% and 4.8%, respectively.

| | | ***A. tequilana*** Transportation Buffer - 50 km (25 km) | | ***A. deserti*** Transportation Buffer - 50 km (25 km) | |
|---|---|---|---|---|---|
| Region | Scenario | Ethanol Percent of State 2010 Gasoline Consumption (a) | Exported Electricity Percent of State 2010 Electricity Consumption (b) | Ethanol Percent of State 2010 Gasoline Consumption (a) | Exported Electricity Percent of State 2010 Electricity Consumption (b) |
| California | Optimistic | 2.8% (2%) | 1% (0.7%) | 1% (0.7%) | 2.68% (1.87%) |
| | Moderate | 2.7% (1.6%) | 0.6% (0.4%) | 0.6% (0.4%) | 2.54% (1.48%) |
| | Strict | 0.1% (0.1%) | 0.1% (0.1%) | 0.1% (0.1%) | 0.12% (0.1%) |
| Arizona | Optimistic | 9.6% (5.4%) | 4.8% (2.7%) | 4.8% (2.7%) | 4.99% (2.79%) |
| | Moderate | 9.6% (5.4%) | 3.5% (1.8%) | 3.5% (1.8%) | 4.99% (2.79%) |
| | Strict | 0% (0%) | 0.4% (0.2%) | 0.4% (0.2%) | 0% (0%) |
| Texas | Optimistic | 7.8% (5.5%) | 1.1% (0.6%) | 1.1% (0.6%) | 4.83% (3.4%) |
| | Moderate | 6.8% (4.7%) | 0.2% (0.1%) | 0.2% (0.1%) | 4.24% (2.92%) |
| | Strict | 2.5% (1.6%) | 0% (0%) | 0% (0%) | 1.53% (1.01%) |

**Table 4-11.** Potential state-level energy benefit from producing agave plants as a bioenergy feedstock.

(a) 2010 State gasoline consumption for California, Arizona, and Texas was 56.6 Gl, 8.7 Gl, and 51.2 Gl respectively (Energy Information Administration, U.S. Department of Energy, Washington D.C., 2013)
(b) 2010 State electricity retail sales for California, Arizona, and Texas was 258.5 TWh, 72.8 TWh, and 358.5 TWh respectively (Energy Information, U.S. Department of Energy, Washington D.C. Administration, 2013)

*Other feedstocks*

Comparing these results to the mapped potential for other feedstocks illustrates the estimated potential for *Agave* to be a complementary bioenergy feedstock because it is suitable in areas where other conventional feedstocks are not. The results from this analysis were compared to county-level estimates of future production of annual energy crops and perennial grasses from the U.S. Billion-Ton Update (Perlack & Stokes, 2011). Areas of suitable land area for *Agave* occurring within counties with estimated future annual energy crop production greater than zero showed the most overlap in the southern tip of Texas. Of the two species studied here, *A. tequilana* showed the most overlap with other potential energy crop production, however even in the optimistic scenario where the most land area was mapped as suitable for *Agave*, the estimated number of hectares to be planted for other annual energy crops (like switchgrass and *Miscanthus*) is only 0.8 Mha in 2020 and only 0.12 Mha in 2030 (assuming a $80 target price). When looking at modeled switchgrass yields by published Wullschleger et al. (Jager *et al.*, 2010; Wullschleger *et al.*, 2010) in the overlapping areas of Texas (in the 2030/$80 model), the average modeled switchgrass yield in the areas also mapped as suitable for the *A. tequilana* was 8.7 Mg $ha^{-1}$ $yr^{-1}$ (*A. tequilana* optimistic scenario) [1] – a yield well below what's documented for *A. tequilana* ( 24 Mg $ha^{-1}$ $yr^{-1}$). In other words, even where *Agave* and switchgrass production could potentially overlap at a macro scale, for example in southwestern Texas, future switchgrass production would only use a small part of that land, and it would be a relatively low yielding crop. Therefore it may be concluded that *Agave* will not be a major competitor to other energy crops, but rather a complementary renewable energy species that can be productive in regions where others cannot.

**Discussion**

This paper is the first to explore the quantitative potential for agave production as a bioenergy feedstock in the U.S. Based on a spatially-explicit site suitability model, the results presented here suggest that there is potential for *Agave* to be grown as an energy feedstock in the southwestern region of the U.S. – particularly in Arizona, California and Texas – and a significant portion of these areas are proximate to existing rail infrastructure. In the moderate scenario, 25.19 Mha of land were modeled as suitable for *A. tequilana* within the 50km rail buffer, which equates to a total production potential of 8,793 Ml of ethanol and an exported electricity potential of 26,370 GWh. Even in the strictest scenario and when limited to a 25km rail buffer, the potential suitable land area for *A. tequilana* is still 3.7 Mha – a land area comparable to the current planted area for sorghum (3.1 Mha planted in 2013) – another drought-resistant crop drawing recent attention as a potential bioenergy feedstock in the U.S. (Almodares & Hadi, 2009). This amount of land under *Agave* production could still produce 1,293 Ml of ethanol and an exported electricity potential of 3,876 GWh. Consequently, the lowest risk

strategy would be to begin production on the most suitable land areas and eventually fan out to the moderately suitable regions when the most suitable regions have been saturated.

Both *Agave* species showed the highest state-level impact in Arizona, where agave has the potential to contribute 4.8 to 9.6% of the states' ethanol consumption, and 2.5 to 4.9% of its electricity consumption, for *A. deserti* and *A. tequilana*, respectively. If additional species are incorporated to the production system, which could potentially diversify and increase the maximum suitable area for agave plantations, these numbers may be increased further. Texas may also significantly benefit from producing agave as a bioenergy feedstock – potentially seeing as much as 7.8% of its ethanol consumption and 4.8% of its electricity consumption produced by *A. tequilana* production. Considering that the total U.S. ethanol production in 2012 was 13,300 million gallons (Renewable Fuels Association, 2013), cumulative estimates for *A. tequilana* calculate a potential production of 28.1% of total U.S. ethanol production in the optimistic scenario within 50km of rail infrastructure, 17.5% in the moderate scenario, and 3.8% in the strict scenario.

Feedstocks that grow where other food crops cannot have the potential to invigorate local economies however the actual maximum energy potential from agave plants will rely on the synergy of multiple technological factors. When converting agave plants to sugar or ethanol, there are both solid and liquid byproducts, which can further add to the energy value potential of *Agave*, if downstream systems are in place to process these byproducts. For example, the solid byproducts can be used to generate electricity or be returned to soils as nutrients. Current generation biofuel conversion technologies (fermentation of sugars to ethanol), as well as potential next-generation technologies, (that could convert the solid material to ethanol or other hydrocarbon liquid fuels) consume heat and electricity in the plant operations.

Although agave plantations do not currently exist in these U.S. states, the agronomic practices are well-established because of the long-standing practices of commercial agave production in Mexico and elsewhere for both tequila and fiber. As a bioenergy feedstock, *Agave* spp. grown on plantations can be managed on a short rotation forestry-type production cycle, with approximately five years interval between planting and harvest. In this way, growing agave as a bioenergy feedstock is more akin to short-rotation forestry production (eg. poplar production) rather than perennial grass species, like switchgrass and *Miscanthus*. However an additional consideration is that current agave plantations typically leave agave leaves on the soil to maintain soil nutrients. The absence of returning solid from biorefineries to soils might result in the plantation's use of organic fertilizers which would reduce the overall life-cycle and sustainability of agave-based biofuels and or electricity (Davis *et al.*, 2011). Thus utilizing agave solid byproducts as a heat and or power feedstock must be weighed against the costs and benefits of returning the plant material to the plantation soils.

Looking forward to a future with a changing climate, CAM species like *Agave*, which are adapted to extreme high temperatures and drought conditions, are more likely to withstand variable temperatures and rainfall predicted in future climate over the next century. In fact, *Agave* and plantation yields are likely to benefit from increased temperatures and $CO_2$ levels that accompany climate change because net $CO_2$ uptake in *Agave* increases as $CO_2$ levels increase (Drennan & Nobel, 2001; Garcia-Moya *et al.*, 2011). In addition, increasing temperatures,

particularly increasing minimum nighttime temperatures and decreasing freezing days (both of which are currently limiting factors for *Agave*), may expand the land potential for the species.

It should also be noted that potential agave production in the United States is not limited to the two species evaluated in this study; in fact 25 *Agave* species are native to the southwestern United States (Garcia-Moya *et al.*, 2011). Of the two species reviewed here, high-yielding *A. tequila*na is more likely to be actually used for bioenergy purposes. *A. deserti* is used here as more of a comparator species that can be used to evaluate the extremes for the genus *Agave*, but is not a likely candidate for immediate deployment in biofuel applications. Moreover, some *Agave* species not native to the U.S., such as *A. salmiana* (42 Mg $ha^{-1}$ $yr^{-1}$) and *A. mapsiaga* (38 Mg $ha^{-1}yr^{-1}$), under ideal cultivated conditions are known to produce even higher yields than the species studied here (Escamilla-Treviño, 2012). However empirical data for these species that is necessary to drive the model used in the present study are not currently available. For this reason, there is a need for more research involving rigorous replicated field trials to better understand tolerance ranges of additional *Agave* species, and how varied climates affect agave production and yield patterns.

As additional *Agave* species are targeted for future research and as more empirical data becomes available, this working model can easily be modified to map their potential in the landscape. The model used here incorporates fuzzy logic, which provides a more differentiated result by showing degrees of membership to the desired criteria. The sensitivity analysis revealed that the suitability for the *Agave* species studied here is most restricted by climatic variables – as opposed to physical soil factors – which prove to be a major factor in determining mapped suitable area. Particularly, minimum temperature in winter months is the most influential criterion restricting land area suitable for agave cultivation in these regions. Therefore *Agave* with a relative tolerance to cold temperatures, while also producing high yields, will have a distinct advantage as a bioenergy feedstock in the U.S., and selective breeding or engineering of *Agave* with these traits should be a priority for the development of bioenergy-optimized agave cultivars.

## Conclusion

Meeting the demand for safe, reliable energy will take a diversity of renewable energy sources. Biofuels have been identified as a critical component to achieving a more sustainable future by supplying a low-carbon option for difficult to decarbonize portions of the transportation sector (Williams *et al.*, 2012; Wei *et al.*, 2013). In addition to electricity generated from a low-carbon electricity generation sector, and in combination with economy-wide energy efficiency improvements, bioenergy will help displace current fossil fuel combustion in building and transportation sectors. Although the arid and semiarid southwestern U.S. regions are endowed with large solar electricity potential, they are limited in their ability to produce conventional low-carbon bioenergy feedstocks and have few options available for supplying residual liquid fuel demand other than importing biofuels from other regions. This not only adds costs to bioenergy use, but also reduces the carbon benefits (through long-distance transportation), and denies these regions the economic prosperity associated with producing their own fuels. This paper illustrates that given the right physical conditions, agave plants have the potential to complement conventional bioenergy crops given its ability to grow on land not currently utilized by agriculture or other potential bioenergy crops. Policymakers seeking to advance low-carbon

intensive energy supplies and regional job development in the dry southwestern U.S. should consider agave-derived bioenergy development.

# Chapter Five

## Conclusions and directions for future research

The challenges in planning for global bioenergy mandates are large and numerous. Foremost among them is the need to determine where to locate bioenergy plantations and infrastructure to meet energy mandates while ensuring sustainable food production and environmental protection. Marginal land has been proposed as an ideal place to grow bioenergy crops to ensure sustainability. Furthermore, drought-tolerant bioenergy crops have also been targeted for future bioenergy production in part because of their ability to grow on lands marginal for conventional bioenergy crops. However, while the term "marginal land" has become so intertwined with discussions surrounding bioenergy, its definition is not always clear. In fact, the concept of marginal land has varied across time, space and discipline to meet multiple management goals. Even in the modern context of bioenergy research, a consistent and common working definition of marginal land is elusive.

Still, as energy policies around the world continue to mandate increases in bioenergy production, the need for quantitative and spatially-explicit estimates of available marginal land is driving a new field of research. Despite the similarity in stated goals among these works – mapping marginal land for bioenergy production – a number of inconsistencies across studies remain that make comparisons and standardization difficult, and to date none of the current methods are widely accepted. However, to allow for reliable estimates of bioenergy production across scales, some measure of consistency is critical. The research presented here is the first to explore how the concept of marginal land is used in a GIS modeling framework. The results of this review help guide the development of an innovative and flexible model to map land available for drought-tolerant bioenergy crops in the United States.

This dissertation presents a review of literature on the topic of identifying marginal lands and then develops a state-of-the-art GIS modeling framework for spatially identifying where to plant bioenergy crops on suitable marginal lands. In Chapter 2, I reviewed a collection of studies that targeted second generation bioenergy crops, focused on land that may be categorized as marginal, and used spatially-explicit GIS techniques. My results revealed that the concept of "marginal land" is relative – easily flexible to multiple research objectives. Despite similar goals, the current literature shows considerable differences in working definitions, including model framework, data inputs, scale, and treatment of uncertainty that limit the potential for comparison across studies. For example, among the 18 studies that used land cover as an input criterion, 13 different land cover datasets were used; among the 17 studies to use soil, 16 different soil datasets were used. When assessing modeling framework, I found that the majority of studies used simple linear GIS overlays with distinct thresholds between land categorized as marginal and land categorized as not marginal. Furthermore, the thresholds used were rarely consistent

between studies. I also found that the overwhelming majority of studies reviewed failed to address model sensitivity or uncertainty in their mapped results. These important results helped frame the rest of the dissertation.

The following chapters focused on the development of a site suitability model for two types of drought-tolerant bioenergy crops – a modified variety of switchgrass as well as two species within the *Agave* genus. Drought-tolerant bioenergy crops are of particular interest in the US because of their ability to grow on land not suitable for conventional agricultural production. Therefore these species may potentially contribute to the country's renewable energy supply while minimizing the impact on current land uses. To explore this issue, Chapters 3 and 4 both employed a site suitability model rooted in fuzzy logic. Each chapter explored different aspects of the fuzzy suitability model, including sensitivity to overlay techniques and model input parameters.

In Chapter 3, I explored the geographic potential for a modified, drought-tolerant variety of switchgrass in the US. I showed that incorporating fuzzy logic into the site suitability model provided a more progressive suitability index by synthesizing tradeoffs between multiple criteria with continuous degrees of suitability. I concluded that the Gamma fuzzy overlay function best recognizes tradeoffs between combinations of input criteria. The results highlight a range of land suitable for drought-tolerant bioenergy crops in the Great Plain States. Within this area, I focused on Kansas for a more detailed analysis and found that 96% of available land within Kansas that is suitable for switchgrass falls within a dryness index equivalent to an approximately 55-day long dry stretch. From these results, I concluded that when considering climate change and risk factors associated with crop production of marginal lands, it may be optimal to aim for higher drought-tolerance, but only to a point. Beyond a dryness index threshold equivalent to a 55-day long dry period, the data showed diminishing returns on land available. These results have very direct practical applications for plant science, as they could help guide research on drought-tolerant varieties of switchgrass.

Although switchgrass has long been discussed as a viable bioenergy species, another drought-tolerant crop – *Agave* – has recently garnered attention as a bioenergy crop. Yet prior to this research, a spatially detailed evaluation of the geographic potential for the crop in the US was unexplored. In Chapter 4, I demonstrated that the area suitable for growing *Agave* as a bioenergy feedstock in the Southwestern US is sufficient to contribute to domestic renewable energy needs (electricity and fuel). At the state-level, Arizona and Texas show the highest potential energy benefit. The sensitivity analysis revealed that the suitability for the *Agave* species studied here is most restricted by climatic variables – as opposed to physical soil factors – that prove to be major factors in determining mapped suitable area. Particularly, minimum temperature in winter months is the most influential criterion restricting land area suitable for *Agave* cultivation in these regions. I demonstrated that even in the strictest scenario tested, the suitable land area mapped for *Agave tequilana* is comparable to the land area currently planted for sorghum –

another bioenergy crop of interest. Finally, comparing these results to estimates of future production of annual energy crops and perennial grasses illustrates that *Agave* will not be a major competitor to other energy feedstocks, but rather a complementary renewable energy crop – filling the gap in areas unsuitable for conventional energy feedstocks.

*Directions for future research*

Meeting the demand for safe, reliable energy will take a diversity of renewable energy sources. Bioenergy production on marginal lands promises to become a growing sector of the renewable energy market. As such, this dissertation makes an original and timely contribution to the renewable and sustainable energy field, as research mapping marginal land for bioenergy crops is ongoing and publications continue to grow in number. While my dissertation begins to outline a framework for mapping marginal lands and optimizes the placement of drought-tolerant bioenergy crops on marginal landscapes in the US, there are key aspects that could be expanded for future research.

First, the numerous works that map marginal land for bioenergy potential are in need of a meta-analysis to highlight shared examinations that use common data and protocols. I suggest that there is tremendous future need for spatial modeling of bioenergy, yet further work should be done to allow for comparative models across countries and scales to enhance a more accurate understanding of the cumulative global potential for bioenergy crops.

Second, some *Agave* species are known to produce even higher yields than the species studied here, however empirical data for these species that is necessary to drive the model used in the present study are not currently available. For this reason, there is a need for more research involving rigorous replicated field trials to better understand tolerance ranges of additional *Agave* species, and how varied climates affect *Agave* production and yield patterns. As additional *Agave* species are targeted for future research and as more empirical data becomes available, this working model can be modified easily to map their potential in the landscape.

Third, climate change over the next century will impact natural water cycles significantly, thereby leading to an increased frequency of extreme weather events, including floods and droughts. Therefore when considering a future with a changing climate, subsequent bioenergy research on the placement of these crops will need to take variable rainfall into consideration yet rainfall timing and location is very difficult to predict. While drought-tolerant bioenergy crops are leading the charge in water-limited conditions, even species adapted to arid conditions are likely to be affected by altered climatic conditions. Therefore future research will need to quantify the effects of climate on overall yield and production potential.

That said, the outlook is not completely dismal for bioenergy crops. Crassulacean acid metabolism (CAM) species like *Agave*, which are adapted to extreme high temperatures and drought conditions, are more likely to withstand variable temperatures and rainfall predicted in future climate. In fact, *Agave* and plantation yields are likely to benefit from increased

temperatures and $CO_2$ levels that accompany climate change because net $CO_2$ uptake in *Agave* increases as $CO_2$ levels increase. In addition, increasing temperatures, particularly increasing minimum nighttime temperatures and decreasing freezing days (both of which are currently limiting factors for *Agave* as evidenced in Chapter 4), may expand the land potential for the species.

Finally, accurate quantification of bioenergy production cannot be accomplished by remote sensing and GIS site suitability techniques alone. There are real world constraints to bioenergy cultivation that are not easily quantified in spatial databases. These mapping limitations make effective understanding of the tradeoffs between energy, food, and land elusive, yet nonetheless tremendously important and worthy of much more applied research.

# Cited Literature


Alley WM (1984) The Palmer drought severity index: limitations and assumptions. *Journal of climate and applied meteorology*, **23**, 1100–1109.

Angelis-Dimakis A, Biberacher M, Dominguez J, *et al.* (2011) Methods and tools to evaluate the availability of renewable energy sources. *Renewable and Sustainable Energy Reviews*, **15**, 1182–1200.

Arbia G, Griffith D, Haining R (1998) Error propagation modelling in raster GIS: overlay operations. *International Journal of Geographical Information Science*, **12**, 145–167.

Barlowe R (1978) Land resource economics: The economics of real estate.

Barney JN, Mann JJ, Kyser GB, Blumwald E, Van Deynze A, DiTomaso JM (2009) Tolerance of switchgrass to extreme soil moisture stress: ecological implications. *Plant Science*, **177**, 724–732.

Bartholomé E, Belward AS (2005) GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, **26**, 1959–1977.

Berndes G, Hoogwijk M, van den Broek R (2003) The contribution of biomass in the future global energy supply: a review of 17 studies. *Biomass and Bioenergy*, **25**, 1–28.

Bibby JS, Mackney D (1969) *Land use capability classification*. Rothamsted Experimental Station.

Bielecka E (2002) GIS approach for delimitation of Less-Favoured Farming Areas in Poland. *Journal of Water and Land Development*,, 73–89.

Biradar CM, Thenkabail PS, Noojipady P, *et al.* (2009) A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. *International journal of applied earth observation and geoinformation*, **11**, 114–129.

Bonham-Carter G (1994) *Geographic information systems for geoscientists: modelling with GIS*. Pergamon press.

Boryan C, Yang Z, Mueller R, Craig M (2011) Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto International*, **26**, 341–358.

Breuning-Madsen H, Reenberg A, Holst K (1990) Mapping potentially marginal land in Denmark. *Soil Use and Management*, **6**, 114–120.

Bryngelsson DK, Lindgren K (2013) Why large-scale bioenergy production on marginal land is unfeasible: A conceptual partial equilibrium analysis. *Energy Policy*, **55**, 454–466.

Cai X, Zhang X, Wang D (2011) Land availability for biofuel production. *Environmental Science and Technology-Columbus*, **45**, 334.

Calvert K, Pearce JM, Mabee WE (2013) Toward renewable energy geo-information infrastructures: Applications of GIScience and remote sensing that build institutional capacity. *Renewable and Sustainable Energy Reviews*, **18**, 416–429.

Campbell JE, Lobell DB, Genova RC, Field CB (2008a) The global potential of bioenergy on abandoned agriculture lands. *Environmental science & technology*, **42**, 5791–5794.

Campbell JE, Lobell DB, Genova RC, Field CB (2008b) The global potential of bioenergy on abandoned agriculture lands. *Environmental Science & Technology*, **42**, 5791–5794.

Cassman KG, Liska AJ (2007) Food and fuel for all: realistic or foolish? *Biofuels, Bioproducts and Biorefining*, **1**, 18–23.

Caver SJ (1991) Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information System*, **5**, 321–339.
Cedeño MC (1995) Tequila production. *Critical reviews in biotechnology*, **15**, 1–11.
Charabi Y, Gastli A (2011) PV site suitability analysis using GIS-based spatial fuzzy multi-criteria evaluation. *Renewable Energy*, **36**, 2554–2561.
Congalton RG (2009) Accuracy Assessment of Spatial Datatsets. In: *Manual of geographic information systems* pp225–233. Bethesda, MD, ASPRS.
Crosetto M, Tarantola S, Saltelli A (2000) Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture, ecosystems & environment*, **81**, 71–79.
Dai A, Trenberth KE, Qian T (2004) A global dataset of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming. *Journal of Hydrometeorology*, **5**, 1117–1130.
Dale VH, Kline KL, Wiens J, Fargione J (2010) *Biofuels: implications for land use and biodiversity*. Ecological Society of America Washington, DC.
Davis SC, Dohleman FG, Long SP (2011) The global potential for Agave as a biofuel feedstock. *GCB Bioenergy*, **3**, 68–78.
Downing M, Eaton LM, Graham RL, *et al.* (2011) US Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry.
Drennan PM, Nobel PS (2001) Responses of CAM species to increasing atmospheric CO2 concentrations. *Plant, Cell & Environment*, **23**, 767–781.
Elobeid A, Hart C (2007) Ethanol Expansion in the Food versus Fuel Debate: How Will Developing Countries Fare? *Journal of Agricultural & Food Industrial Organization*, **5**.
Energy Information Administration, U.S. Department of Energy, Washington D.C. (2013) Petroleum & Other Liquids - Prices, Sales Volumes & Stocks by State.
Energy Information, U.S. Department of Energy, Washington D.C. Administration (2013) State Electricity Profiles.
Escamilla-Treviño LL (2012) Potential of Plants from the Genus Agave as Bioenergy Crops. *BioEnergy Research*, **5**, 1–9.
Eschenbach TG (1992) Spiderplots versus tornado diagrams for sensitivity analysis. *Interfaces*, **22**, 40–46.
ESRI (2012) *ArcGIS 10.1*. Redlands, CA.
Fahd S, Fiorentino G, Mellino S, Ulgiati S (2012) Cropping bioenergy and biomaterials in marginal land: The added value of the biorefinery concept. *Energy*, **37**, 79–93.
Field CB, Campbell JE, Lobell DB (2008) Biomass energy: the scale of the potential resource. *Trends in ecology & evolution*, **23**, 65–72.
Fiorese G, Guariso G (2010) A GIS-based approach to evaluate biomass potential from energy crops at regional scale. *Environmental Modelling & Software*, **25**, 702–711.
Friedel MJ (2011) Climate change effects on ecosystem services in the United States–issues of national and global security. In: *Climate Change and its Effects on Water Resources* pp17–24. Springer.
Fritz S, See L (2005) Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, **19**, 787–807.
Fritz S, See L (2008) Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, **14**, 1057–1075.

Fritz S, See LM, Van der Velde M, *et al.* (2013) Downgrading Recent Estimates of Land Available for Biofuel Production. *Environmental Science & Technology*,.

Fuller RM, Smith GM, Sanderson JM, Hill RA, Thompson AG (2002) Land Cover Map 2000: a general description of the UK's new vector GIS based on classification of remotely sensed data. *Cartogr. J*, **39**, 15–25.

Garcia-Moya E, Romero-Manza A, Nobel PS (2011) Highlights for Agave productivity. *GCB Bioenergy*, **3**, 4–14.

Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Gross KL, Robertson GP (2013a) Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, **493**, 514–517.

Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Gross KL, Robertson GP (2013b) Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, **493**, 514–517.

Gelfand I, Zenone T, Jasrotia P, Chen J, Hamilton SK, Robertson GP (2011) Carbon debt of Conservation Reserve Program (CRP) grasslands converted to bioenergy production. *Proceedings of the National Academy of Sciences*, **108**, 13864–13869.

Gill-Austern D (2011) The Impact of Rising Corn Prices on the Conservation Reserve Program: An Empirical Model. *Undergraduate Economic Review*, **7**, 22.

Gobeille A, Yavitt J, Stalcup P, Valenzuela A (2006) Effects of soil management practices on soil fertility measurements on *Agave tequilana* plantations in Western Central Mexico. *Soil and Tillage research*, **87**, 80–88.

Goldewijk KK (2001) Estimating global land use change over the past 300 years: the HYDE database. *Global Biogeochemical Cycles*, **15**, 417–433.

Gopalakrishnan G, Cristina Negri M, Snyder SW (2011) A novel framework to classify marginal land for sustainable biomass feedstock production. *Journal of Environmental Quality*, **40**, 1593–1600.

Gu Y, Brown JF, Verdin JP, Wardlow B (2007) A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters*, **34**, L06407.

Hall DO, Rosillo-Calle F, Williams RH, Woods J (1993) *Biomass for energy: supply prospects.* Earthscan.

Hall GB, Wang F, Subaryono J (1992) Comparison of Boolean and fuzzy classification methods in land suitability analysis by using geographical information systems. *Environment and Planning A*, **24**, 497–516.

Han W, Yang Z, Di L, Mueller R (2012) CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, **84**, 111–123.

Hansen HS (2005) GIS-based multi-criteria analysis of wind farm development. *Proceedings of the 10th Scandinavian Research Conference on Geographical Information Science, Royal Institute of Technology*,, 75–87.

Hardie IW, Parks PJ, van Kooten GC (2004) Land use decisions and policy at the intensive and extensive margins. *International Yearbook of Environmental and Resource Economics. UK and MA: Edward Elgar*,, 101–138.

Hargreaves GH (1994) Defining and using reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, **120**, 1132–1139.

Holtum JA, Chambers DON, Morgan T, Tan DK (2011) Agave as a biofuel feedstock in Australia. *GCB Bioenergy*, **3**, 58–67.

Hoogwijk M, Faaij A, van den Broek R, Berndes G, Gielen D, Turkenburg W (2003) Exploration of the ranges of the global potential of biomass for energy. *Biomass and Bioenergy*, **25**, 119–133.

Houghton RA, Unruh JD, Lefebvre PA (1993) Current land cover in the tropics and its potential for sequestering carbon. *Global biogeochemical cycles*, **7**, 305–320.

Ignaciuk A, Vöhringer F, Ruijs A, Van Ierland EC (2006) Competition between biomass and food production in the presence of energy policies: a partial equilibrium analysis. *Energy policy*, **34**, 1127–1138.

Jackson RD, Idso SB, Reginato RJ, Pinter Jr PJ (1981) Canopy temperature as a crop water stress indicator. *Water Resour. Res*, **17**, 1133–1138.

Janke JR (2010) Multicriteria GIS modeling of wind and solar farms in Colorado. *Renewable Energy*, **35**, 2228–2234.

Jankowski P (1995) Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Systems*, **9**, 251–273.

Jensen K, Clark CD, Ellis P, English B, Menard J, Walsh M, de la Torre Ugarte D (2007) Farmer willingness to grow switchgrass for energy production. *Biomass and Bioenergy*, **31**, 773–781.

Jiang H, Eastman JR (2000) Application of fuzzy measures in multi-criteria evaluation in GIS. *International Journal of Geographical Information Science*, **14**, 173–184.

Johansson DJ, Azar C (2007) A scenario based analysis of land competition between food and bioenergy production in the US. *Climatic Change*, **82**, 267–291.

Kang S, Gu B, Du T, Zhang J (2003) Crop coefficient and ratio of transpiration to evapotranspiration of winter wheat and maize in a semi-humid region. *Agricultural water management*, **59**, 239–254.

Kang S, Post WM, Nichols JA, Wang D, West TO, Bandaru V, Roberto C. Izaurralde (2013a) Marginal Lands: Concept, Assessment and Management. *Journal of Agricultural Science*, **5**, p129.

Kang S, Post W, Wang D, Nichols J, Bandaru V, West T (2013b) Hierarchical marginal land assessment for land use planning. *Land Use Policy*, **30**, 106–113.

Karl TR, Trenberth KE (2003) Modern global climate change. *science*, **302**, 1719–1723.

Kartesz JT (2010) The Biota of North America Program (BONAP). North American Plant Atlas (http://www.bonap.org/MapSwitchboard.html). Chapel Hill, N.C. [maps generated from Kartesz, J.T. 2010. Floristic Synthesis of North America, Version 1.0. Biota of North America Program (BONAP). (in press)].

Kim S, Dale BE (2011) Indirect land use change for biofuels: testing predictions and improving analytical methodologies. *biomass and bioenergy*, **35**, 3235–3240.

Klingebiel AA, Montgomery PH (1961) *Land-capability classification*. Soil Conservation Service, US Department of Agriculture.

Kukk L, Astover A, Muiste P, Noormets M, Roostalu H, Sepp K, Suuster E (2010) Assessment of abandoned agricultural land resource for bio-energy production in Estonia. *Acta Agriculturae Scandinavica Section B–Soil and Plant Science*, **60**, 166–173.

Lam MK, Tan KT, Lee KT, Mohamed AR (2009) Malaysian palm oil: surviving the food versus fuel dispute for a sustainable future. *Renewable and Sustainable Energy Reviews*, **13**, 1456–1464.

Larson GA, Roloff G, Larson WE (1988) A new approach to marginal agricultural land classification. *Journal of soil and water conservation*, **43**, 103–106.
Liu T, Ma Z, McConkey B, *et al.* (2012a) Bioenergy production potential on marginal land in Canada. In: *Agro-Geoinformatics (Agro-Geoinformatics), 2012 First International Conference on* pp1–6.
Liu L, Zhuang D, Jiang D, Huang Y (2012b) Assessing the potential of the cultivation area and greenhouse gas (GHG) emission reduction of cassava-based fuel ethanol on marginal land in Southwest China. *African Journal of Agricultural Research*, **7**, 5594–5603.
Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Yang L, Merchant JW (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, **21**, 1303–1330.
Lovett AA, Sünnenberg GM, Richter GM, Dailey AG, Riche AB, Karp A (2009) Land use implications of increased biomass production identified by GIS-based suitability and yield mapping for Miscanthus in England. *Bioenergy Research*, **2**, 17–28.
Lu L, Jiang D, Zhuang D, Huang Y (2012) Evaluating the Marginal Land Resources Suitable for Developing Pistacia chinensis-Based Biodiesel in China. *Energies*, **5**, 2165–2177.
Lubowski RN, Bucholtz S, Claassen R, Roberts MJ, Cooper JC, Gueorguieva A, Johansson R (2006) *Environmental effects of agricultural land-use change*. US Department of Agriculture, Economic Research Service Washington, DC, USA.
Malczewski J (2004) GIS-based land-use suitability analysis: a critical overview. *Progress in planning*, **62**, 3–65.
Malczewski J (2006) Ordered weighted averaging with fuzzy quantifiers: GIS-based multicriteria evaluation for land-use suitability analysis. *International Journal of Applied Earth Observation and Geoinformation*, **8**, 270–277.
McHarg I (1998) *To heal the earth: selected writings of Ian L. McHarg*. island Press.
McHarg Ian L (1969) *Design with nature*. Natural History Press. New York.
McHarg IL, Mumford L (1969) *Design with nature*. American Museum of Natural History New York.
Mendelsohn R, Dinar A (2003) Climate, water, and agriculture. *Land economics*, **79**, 328–341.
Milbrandt A, Overend RP (2009) *Assessment of biomass resources from marginal lands in APEC economies*. National Renewable Energy Laboratory (NREL), Golden, CO.
Miller DA, White RA (1998) A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions*, **2**, 1–26.
Mishra V, Cherkauer KA, Shukla S (2010) Assessment of drought due to historic climate variability and projected future climate change in the midwestern United States. *Journal of Hydrometeorology*, **11**, 46–68.
Mo KC (2011) Drought onset and recovery over the United States. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **116**.
Moran MS, Clarke TR, Inoue Y, Vidal A (1994) Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote Sensing of Environment*, **49**, 246–263.
Munn LC (1986) The Canada Land Inventory. In: *Land and its Uses—Actual and Potential* pp391–406. Springer.
Nakicenovic N, Alcamo J, Davis G, *et al.* (2000) *Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change*.

Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US).
Niblick B, Monnell JD, Zhao X, Landis AE (2012) Using geographic information systems to assess potential biofuel crop production on urban marginal lands. *Applied Energy*,.
Nijsen M, Smeets E, Stehfest E, Vuuren DP (2012) An evaluation of the global potential of bioenergy production on degraded lands. *GCB Bioenergy*, **4**, 130–147.
Nobel PS, GARCÍA-MOYA E, Quero E (1992) High annual productivity of certain agaves and cacti under cultivation. *Plant, Cell & Environment*, **15**, 329–335.
Nobel PS, Hartsock TL (1986) Short-term and long-term responses of crassulacean acid metabolism plants to elevated CO2. *Plant physiology*, **82**, 604–606.
Nobel PS, Valenzuela AG (1987) Environmental responses and productivity of the CAM plant, Agave tequilana. *Agricultural and forest meteorology*, **39**, 319–334.
Noorollahi Y, Itoi R, Fujii H, Tanaka T (2008) GIS integration model for geothermal exploration and well siting. *Geothermics*, **37**, 107–131.
NRCS-USDA General Soil Map (STATSGO2). Accessed August 2012. Available online at: http://soildatamart.nrcs.usda.gov
Odeh IOA, Tan DKY, Ancev T (2011) Potential suitability and viability of selected biodiesel crops in Australian marginal agricultural lands under current and future climates. *BioEnergy Research*, **4**, 165–179.
OECD/FAO (2012) *OECD-FAO Agricultural Outlook 2012-2021*. OECD Publishing and FAO.
Ostrowski J (1999) Identification of the occurrence of marginal agricultural lands with the GIS. *Journal of Water and Land Development*,, 7–20.
Pedrycz W (1994) Why triangular membership functions? *Fuzzy sets and Systems*, **64**, 21–30.
Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions Discussions*, **4**, 439–473.
Perez R, Ineichen P, Moore K, Kmiecik M, Chain C, George R, Vignola F (2002) A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, **73**, 307–317.
Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC (2005) *Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply*. DTIC Document.
Perrin RK, Fretes NF, Sesmero JP (2009) Efficiency in Midwest US corn ethanol plants: A plant survey. *Energy Policy*, **37**, 1309–1316.
Peterson B (2012) North American CTA Railroad Network. 2012 August edition. Available online at: http://cta.ornl.gov/transnet/RailRoads.html.
Peterson GM, Galbraith JK (1932) The concept of marginal land. *Journal of Farm Economics*, **14**, 295–310.
Pimienta-Barrios E, Robles-Murguia C, Nobel PS (2001) Net CO2 Uptake for Agave tequilana in a Warm and a Temperate Environment. *Biotropica*, **33**, 312–318.
Ramachandra TV, Shruthi BV (2007) Spatial mapping of renewable energy potential. *Renewable and Sustainable Energy Reviews*, **11**, 1460–1480.
Rathmann R, Szklo A, Schaeffer R (2010) Land use competition for production of food and liquid biofuels: An analysis of the arguments in the current debate. *Renewable Energy*, **35**, 14–22.

Regonda SK, Rajagopalan B, Clark M, Pitlick J (2005) Seasonal cycle shifts in hydroclimatology over the western United States. *Journal of Climate*, **18**, 372–384.
Renewable Fuels Association (2013) *Battling for the Barrel: 2013 Ethanol Industry Outlook*.
Ricardo D (1817) *On the principles of political economy and taxation*. London, U.K., S. M. Dent & Sons ltd.
Robertson GP, Dale VH, Doering OC, Hamburg SP, Melillo JM, Wander MM, Parton W (2008) Agriculture-sustainable biofuels redux. *Science*, **322**.
Rodman LC, Meentemeyer RK (2006) A geographic analysis of wind turbine placement in Northern California. *Energy Policy*, **34**, 2137–2149.
Romm J (2011) The next dust bowl. *Nature*, **478**, 450–451.
Ruiz-Corral JA, Pimienta-Barrios E, Zanudo-Hernández J (2002) Optimal and marginal thermal regions for the cultivation of Agave tequilana on the Jalisco State. *Agrociencia*, **36**, 41–53.
Runge CF, Larson WE, Roloff G (1986) Using productivity measures to target conservation programs: A comparative analysis. *Journal of soil and water conservation*, **41**, 45–49.
Schlenker W, Roberts MJ (2006) Nonlinear effects of weather on corn yields. *Applied Economic Perspectives and Policy*, **28**, 391–398.
Schmer MR, Vogel KP, Mitchell RB, Perrin RK (2008) Net energy of cellulosic ethanol from switchgrass. *Proceedings of the National Academy of Sciences*, **105**, 464–469.
Schweers W, Bai Z, Campbell E, *et al.* (2011) Identification of potential areas for biomass production in China: Discussion of a recent approach and future challenges. *Biomass and Bioenergy*, **35**, 2268–2279.
Searchinger T, Heimlich R, Houghton RA, *et al.* (2008) Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, **319**, 1238–1240.
Sissine F (2007) Energy Independence and Security Act of 2007: a summary of major provisions
Solomon BD (2010) Biofuels and sustainability. *Annals of the New York Academy of Sciences*, **1185**, 119–134.
Somerville C, Youngs H, Taylor C, Davis SC, Long SP (2010) Feedstocks for lignocellulosic biofuels. *Science*, **329**, 790–792.
Steinitz C, Parker P, Jordan L (1976) Hand drawn overlays: their history and prospective uses. *Landscape Architecture*, **66**, 444–455.
Swinton SM, Babcock BA, James LK, Bandaru V (2011) Higher US crop prices trigger little area expansion so marginal land for biofuel crops is limited. *Energy Policy*, **39**, 5254–5258.
Taheripour F, Zhuang Q, Tyner WE, Lu X (2012) Biofuels, cropland expansion, and the extensive margin. *Energy, Sustainability and Society*, **2**, 1–11.
Tang Y, Xie J-S, Geng S (2010) Marginal Land-based Biomass Energy Production in China. *Journal of Integrative Plant Biology*, **52**, 112–121.
Tenerelli P, Carver S (2012) Multi-criteria, multi-objective and uncertainty analysis for agro-energy spatial modelling. *Applied Geography*, **32**, 724–736.
Thünen J von, Heinrich J (1966) The isolated state. *Trans. Carla M. Wartenberg. Oxford: Pergamon*,.
Tian Y, Zhao L, Meng H, Sun L, Yan J (2009) Estimation of un-used land potential for biofuels development in (the) People's Republic of China. *Applied Energy*, **86**, S77–S85.

Tilman D, Hill J, Lehman C (2006) Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science*, **314**, 1598–1600.
Tilman D, Socolow R, Foley JA, *et al.* (2009) Beneficial biofuels—the food, energy, and environment trilemma. *Science*, **325**, 270.
USDA N (2008) *National soil survey handbook - Part 622*. Available online at: http://soils.usda.gov/technical/handbook/contents/part622.html
USDA-FSA (2011) CRP Acreage Expiring by 2017. Accessed October 2011. Available online at: Available online at: http://fsa.usda.gov/.
USDA-NASS (2007) Census of Agriculture: Market land value per acre. Accessed July 2011. Available online at: http://www.agcensus.usda.gov/Publications/2007/index.php.
USDA-NASS (2012) Cropland Data Layer. Accessed July 2011. Available online at: http://nassgeodata.gmu.edu/CropScape.
USGS (1996) *Geological Survey Center for Earth Resources Observation and Science (EROS). HYDRO1k Elevation Derivative Database. LP DAAC, Sioux Falls, South Dakota*.
USGS GAP (GAP) (2011) Protected Areas Database of the United States (PAD-US), version 1.2. Accessed February 2011. Available online at: http://gapanalysis.usgs.gov/padus.
Voivontas D, Assimacopoulos D, Mourelatos A, Corominas J (1998) Evaluation of renewable energy potential using a GIS decision support system. *Renewable Energy*, **13**, 333–344.
Wang H, Fu R, Kumar A, Li W (2010) Intensification of summer rainfall variability in the southeastern United States during recent decades. *Journal of Hydrometeorology*, **11**, 1007–1018.
Wehner M, Easterling DR, Lawrimore JH, Heim Jr RR, Vose RS, Santer BD (2011) Projections of future drought in the Continental United States and Mexico. *Journal of Hydrometeorology*, **12**, 1359–1377.
Wei M, Nelson JH, Greenblatt JB, *et al.* (2013) Deep carbon reductions in California require electrification and integration across economic sectors. *Environmental Research Letters*, **8**, 014038.
Wickham JD, Stehman SV, Gass L, Dewitz J, Fry JA, Wade TG (2013) Accuracy assessment of NLCD 2006 land cover and impervious surface. *Remote Sensing of Environment*, **130**, 294–304.
Williams JH, DeBenedictis A, Ghanadan R, *et al.* (2012) The technology path to deep greenhouse gas emissions cuts by 2050: the pivotal role of electricity. *Science*, **335**, 53–59.
Wu W, Huang J, Deng X (2010) Potential land for plantation of Jatropha curcas as feedstocks for biodiesel in China. *Science in China Series D: Earth Sciences*, **53**, 120–127.
Wullschleger SD, Davis EB, Borsuk ME, Gunderson CA, Lynd LR (2010) Biomass production in switchgrass across the United States: database description and determinants of yield. *Agronomy Journal*, **102**, 1158–1168.
Yan X, Tan DK, Inderwildi OR, Smith JAC, King DA (2011) Life cycle energy and greenhouse gas analysis for agave-derived bioethanol. *Energy & Environmental Science*, **4**, 3110–3121.
Yi Z, XiFeng F, JuYing W, LiuSheng D, XinCun H (2012) Effects of water stress on growth and biomass quality of switchgrass. *Journal of China Agricultural University*, **17**, 59–64.
Young DR, Nobel PS (1986) Predictions of soil-water potentials in the north-western Sonoran Desert. *The Journal of Ecology*,, 143–154.

Youngs H, Somerville C (2012) Development of feedstocks for cellulosic biofuels. *F1000 Biology Reports*, **4**.

Yousefi H, Ehara S, Noorollahi Y (2007) Geothermal potential site selection using GIS in Iran. In: *Proceedings of the 32nd Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California* pp174–182.

Zadeh LA (1965) Fuzzy sets. *Information and Control*, **8**, 338–353.

Zhuang D, Jiang D, Liu L, Huang Y (2011) Assessment of bioenergy potential on marginal land in China. *Renewable and Sustainable Energy Reviews*, **15**, 1050–1056.

The PRISM Climate Group at Oregon State University (2004) PRISM 30-arcsec (800m) NORMALS 1971-2000. Available online at: http://www.prism.oregonstate.edu.