DISSERTATION

INTEGRATED ASSESSMENT OF AGRICULTURAL ECOSYSTEMS USING SIMULATION-OPTIMIZATION AND MACHINE LEARNING

Submitted by

Trung H. Nguyen

Department of Soil and Crop Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2018

Doctoral Committee:

Advisor: Keith Paustian

Francesca Cotrufo
Eugene Kelly
Stephen Leisz
Christian Davies

ABSTRACT

INTEGRATED ASSESSMENT OF AGRICULTURAL ECOSYSTEMS USING SIMULATION-OPTIMIZATION AND MACHINE LEARNING

Agriculture provides many ecosystem services to human society but is also a major cause of environmental degradation. The key challenge of modern agricultural production is to meet projected increases in global demands for food, water, and energy in sustainable ways. Sustainable agricultural production requires integrated decision-support tools and rigorous assessment methods to improve the efficiency of natural resource management while minimizing its impacts to society and long-term ecosystem health. This dissertation focuses on developing methodology and modeling tools to support decision-making for sustainable agricultural resource management. The Millennium Ecosystem Assessment is used as a guiding framework for all the model development. The dissertation balances between the communication of the integrated assessment methodology and the presentation of the modeling techniques through four independent case studies. The first study links biogeochemical models with life cycle assessment (LCA) to explore the impact of regionally-specific ecosystem carbon stock changes associated with cassava cultivation for ethanol production in Vietnam. The second study couples biogeochemical models with GIS and optimization algorithms to conduct a high-resolution, spatially-explicit trade-off analysis of ecosystem services for irrigated corn production systems in the South Platte River Basin, Colorado, USA. The derived modeling platform is named the "Agricultural Ecosystem Service Optimization" (Ag-EcoSOpt). The third study integrates LCA into the Ag-EcoSOpt for a life-cycle-based optimization of feedstock landscape design for a hybrid corn grain- and stover-based ethanol production system at Front Range Energy

biorefinery, Windsor, Colorado, USA. The last study develops a surrogate-based optimization framework for Ag-EcoSOpt to reduce the computational burden of large-scale landscape analyses. The study explores the trade-offs among seven management objectives of the irrigated corn production systems in Colorado, USA at different spatial scales.

ACKNOWLEDGEMENTS

I would like to acknowledge the generous support from the Fulbright Vietnam scholarship, the US National Science Foundation, Shell Technology Center Houston, and Hue University of Agriculture and Forestry, Vietnam, without which this work would not have been possible.

I would like to express my gratitude to the five members of my committee for their encouragement and support, in particular to Eugene Kelly and Stephen Leisz for their insightful advice and career guidance; to Christian Davies for his industrial perspectives and for nudging me towards ecosystem service topics; and to Francesca Cotrufo for showing me the art of soil science, proposal writing, and more. I give special thanks to my advisor, Keith Paustian, who taught me the nuts and bolts of ecosystem modeling and provided me with a myriad of opportunities to develop my expertise. Keith's critical thinking, effective leadership, and humble personality has greatly influenced my professional attitude.

I'm grateful to receive help from so many colleagues at Colorado State University. In particular, I would like to thank Stephen Williams, Mark Easter, and Ernie Marx for teaching me how to use DayCent; my office-mate Yao Zhang for being a DayCent live dictionary; John Field for introducing me to python and providing insightful manuscript reviews; Maxwell Cook, Chen Wang, and Jialing Zheng for volunteering to help with model development and being awesome team-players; Barb Gibson, Ken Reardon, and the rest of the IGERT crew for giving me a precious opportunity to be involved with the MAS Bioenergy program.

I would also like to thank all my family and friends who provided me with emotional and financial support over my years at Colorado State University. I would not have made it this far without you.

DEDICATION

To Mom and Dad,

Who taught me to finish what I started.

To my wife Hien Dao,

Whose funny conversations always make my day

To my daughter Camellia Nguyen,

A true blessing that keeps me going

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

1.1 SUSTAINABLE AGRICULTURE: THE GUIDING CONCEPT

Agriculture provides many goods and products such as food, fuel, and fiber to sustain human society, but is also a major cause of multiple types of environmental degradation. Agriculture occupies about 40% of the Earth's land area with an additional approximately six million hectares converted from natural state every year (Swinton et al., 2007; Deininger and Byerlee, 2011). In developing countries, cultivated land is expected to increase over 47% to meet demand for food by 2050 (Fischer and Heilig, 1997). Agricultural activities and land use conversion (e.g., deforestation) contribute 19% – 29% of global anthropogenic greenhouse gases (IPCC, 2014; Glendining et al., 2009). The agricultural sector accounts for over 70% of global freshwater withdrawals, which is more than three times that of industry (23%) and nine times that of municipal use (8%) (MA, 2005a). Agriculture is also the main driving force of deforestation, biodiversity loss, and soil degradation (Clark and Tilman, 2017); and is the leading source of pollution in many countries due to the run-off and leaching of excessive agrochemical use (Sutton et al., 2013).

The key challenge of the modern agricultural production is to meet the projected increases in global demand for food, water, and energy in sustainable ways. The impacts of agricultural production on ecosystems are likely to increase in the future with ongoing increases in population growth and income-dependent dietary shifts towards more meat-based diets (Tilman et al., 2011; Springmann et al., 2016). The world population is projected to be about 9.7 billion in 2050 (UN, 2017). This will likely result in 70% - 85% increase in the demand for food crops and 30% - 85% increase in water demand over the next five decades (MA, 2005a; Brown, 2012).

According to Bruinsma, (2009), a population of nine billion people will require an additional one billion metric tons of cereals and 200 million metric tons of livestock products as compared to 2007. In addition, the global production and use of biofuels for energy security will impose further pressure on food production, land conversion, water use, biodiversity, and agricultural commodity prices (Gasparatos et al., 2011). Meeting the increasing demands for water, food, and energy will require improvement in genetic engineering of cropping species as well as the more efficient use of natural resources to achieve higher production with minimal impacts to society and long-term ecosystem health. The latter requires integrated assessment methods and decision-support tools that facilitate thorough analysis of agricultural ecosystems.

This dissertation focuses on developing methodology and modeling tools to support decision making for sustainable agricultural resource management. The concept of sustainable agriculture is used as the guiding concept for the assessments and model development in all chapters. Sustainable agriculture is defined by the 1990 U.S. Farm Bill as an "integrated system of plant and animal production practices having a site-specific application that will, over the long term: (a) satisfy human food and fiber needs; (b) enhance environmental quality; (c) make efficient use of non-renewable resources and on-farm resources and integrate appropriate natural biological cycles and controls; (d) sustain the economic viability of farm operations; and (e) enhance the quality of life for farmers and society as a whole." (U.S. Congress, 1990)

1.2 FRAMEWORK FOR AGRICULTURAL ECOSYSTEM ASSESSMENT

Conceptual frameworks are needed to facilitate systematic assessments of agricultural ecosystems, allowing fair comparisons between different scenarios and conditions to support decision-making. A well-designed framework provides a logical structure to guide the evaluation of the system of interest. Such a framework ensures that the essential components and linkages

of the system are identified and thoroughly addressed (MA, 2005b). The millennium ecosystem assessment framework (MA), prepared by 1,360 scientists from 95 countries, provides a foundation for holistic research approaches where multiple perspectives (e.g. ecological, economic, and institutional) are integrated to assess the effects of decision-making and policies on ecosystem sustainability and eventually on human well-being (MA, 2005a). This framework is employed to guide the assessments of agricultural ecosystems in this dissertation. The framework is briefly summarized below:

The MA framework emphasizes the linkages among drivers of ecosystem change, ecosystem services, and human well-being at multiple geographic and time scales (Fig. 1.1). Ecosystem services (ES) are defined as the benefits human obtained from ecosystem, and are often classified into four main categories, including supporting, provisioning, regulating, and cultural services (MA, 2005a). **Supporting services**, such as water and nutrient cycles, support the normal functioning and production of an ecosystem to provide other ecosystem services; **provisioning services** include products directly obtained from ecosystems such as food, biomass, and fresh water; **regulating services** are the benefits obtained from the regulation of ecosystem processes such as climate regulation, water purification, and pollination; and **cultural services** include non-material benefits such as spiritual enrichment, recreation, and aesthetic values. The provision mechanism of ES is often explained using a cascade model that includes the following stages: biophysical structure or process (generated by biotic/abiotic interactions in the ecosystem), function (capability to provide ES), services (flows that connect natural capital stocks produced by ecosystems to societal groups that use them), benefit(s) (e.g. nutrition, health, safety), and value (monetary and non-monetary) (Fig. 1.2) (Haines-Young and Potschin, 2010).

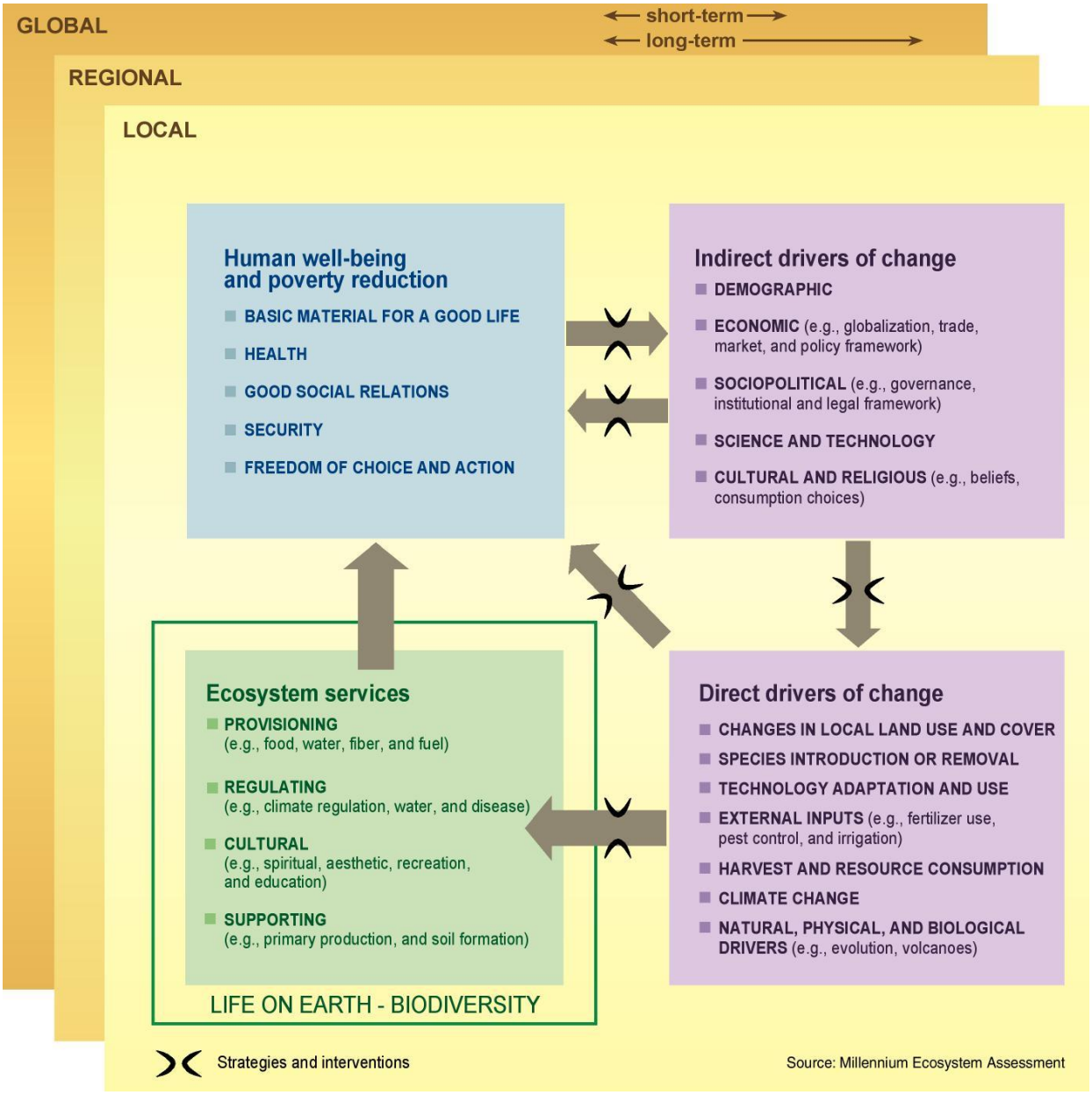


Fig. 1.1. Millennium ecosystem assessment conceptual framework. Source: MA, (2005b).

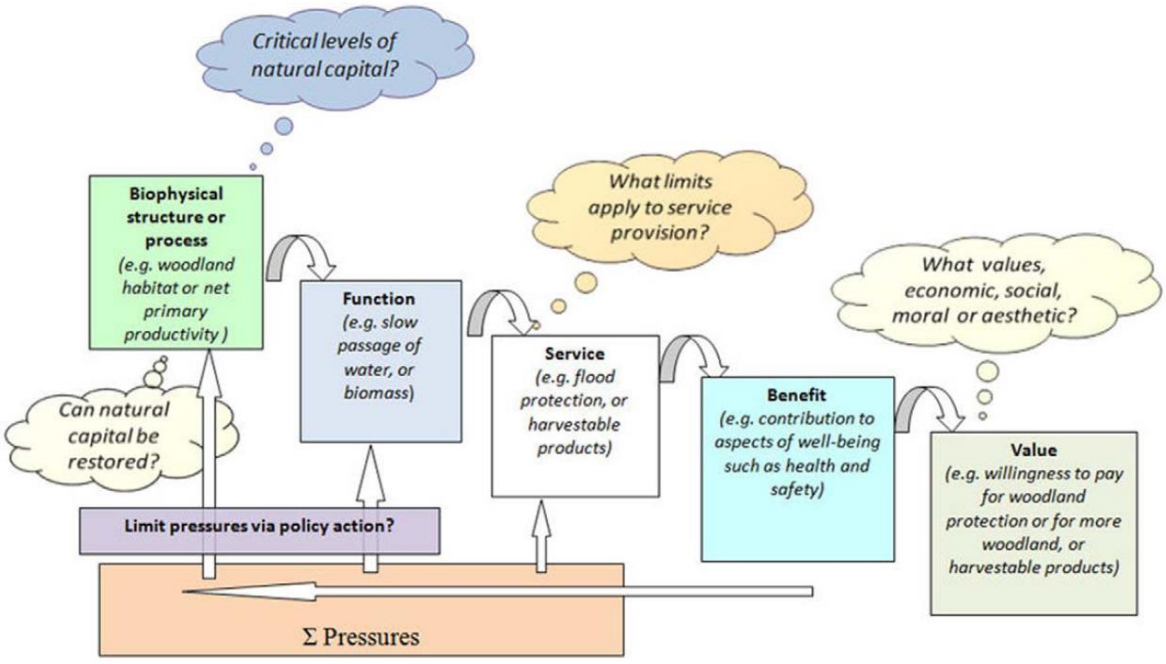


Fig. 1.2. Cascade model of ecosystem service provision. Source: Haines-Young and Potschin, (2010).

Understanding the drivers of ecosystem change is essential to designing interventions that harness positive impacts and minimize negative ones. The MA defined a 'driver' as any factor that changes an aspect of an ecosystem. A direct driver influences ecosystem processes while an indirect driver operates by altering one or more direct drivers. The influence of an indirect driver on ecosystem change is established by understanding its effect on direct drivers. For a specific decision maker, a driver is defined as endogenous if it can be influenced or as exogenous if it cannot be controlled by the decision maker. Whether a driver is exogenous or endogenous to a decision-maker is dependent upon the spatial and temporal scales.

The MA suggests assessment of ecosystems be conducted at multiple temporal and spatial scales due to the cross-scale interactions of ecosystem services and human well-being. However, it emphasizes that the scales need to be appropriate to the process or phenomenon being examined.

Coarse resolution assessments can capture bigger pictures with causality effects but may not detect fine-resolution processes (such as local patterns or anomalies), which are important to identify thresholds and nonlinearities. Finer spatial scales help identify important dynamics of the system that might otherwise be overlooked. Likewise, appropriate time scale is crucial for ecosystem assessments. If an assessment covers a shorter time period than the temporal scale of an ecosystem process, it may not adequately capture the variability associated with long-term cycles.

The MA identifies nine major tasks for conducting an integrated ecosystem assessment including: (1) identifying and categorizing ecosystems and their services; (2) identifying links between human societies and ecosystem services; (3) identifying the direct and indirect drivers of change; (4) selecting indicators of ecosystem conditions, services, human well-being, and drivers; (5) assessing historical trends and the current state of ecosystems, services, and drivers; (6) evaluating the impact of a change in services on human well-being; (7) developing scenarios of ecosystems, services, and drivers; (8) evaluating response options to deal with ecosystem changes and human well-being; and (9) analyzing and communicating the uncertainty of assessment findings.

1.3 INTEGRATED TOOLS FOR AGRICULTURAL ECOSYSTEM ASSESSMENT

Although the MA is a useful road map for ecosystem assessments, the employment of such a complex framework for integrated and operational assessments of agricultural ecosystems is challenging. First, the MA is an ecosystem-centered framework and thus it lacks focus on life cycle and supply chain analysis of provisioning services such as food and fiber, which is important for agricultural production systems. Second, the quantification of ES in highly segmented ecosystems like agriculture requires fine spatial resolutions and detailed consideration

of site-specific conditions such as soils, climate, land use history, and management practices. Such quantification is a difficult task that requires thorough understanding of fundamental physical and biological processes within the ecosystem. Third, agricultural ecosystem management often involves multiple stakeholders as well as complex trade-offs among multiple competing services, which requires thorough trade-off analysis and multi-objective optimization. Fourth, since agricultural production is highly seasonal, agricultural ecosystem assessments need to be timely to support rational decision making. Decision-support tools/platforms that integrate life cycle assessment, detailed quantification of ES at field scale, and optimization algorithms are necessary for successful implementation of the MA framework in agricultural ecosystem context. In addition, those tools need to be agile and fast for meaningful applications in resource management decision making. Several decision-support tools have been developed to support the implementation of MA framework for sustainable ecosystem management. However, as Bagstad et al., (2013) pointed out, those tools were generally designed for landscape-scale operations and not intended for site-scale analyses.

This dissertation focuses on the integration of different decision support tools and methodologies to facilitate the efficient implementation of the MA framework for agricultural ecosystem assessment (Fig. 1.3). Life cycle assessment is used to evaluate impacts of resource management decisions along the supply chains of agricultural ES; process-based biophysical models coupled with geographical information system (GIS) are used to quantify ES at fine spatial and temporal resolutions, providing continuous feedback between management decisions and ecosystem response; multi-objective optimization is used to quantify trade-offs between different management options as well as to identify the most efficient agricultural landscape designs to maximize ecosystem service provision while minimizing the negative impacts to the ecosystems;

and finally machine learning techniques are employed to reduce the computational burdens of our modeling platform, allowing its light-weight application for timely decision support. A brief description of each separate methods and tools are presented below:

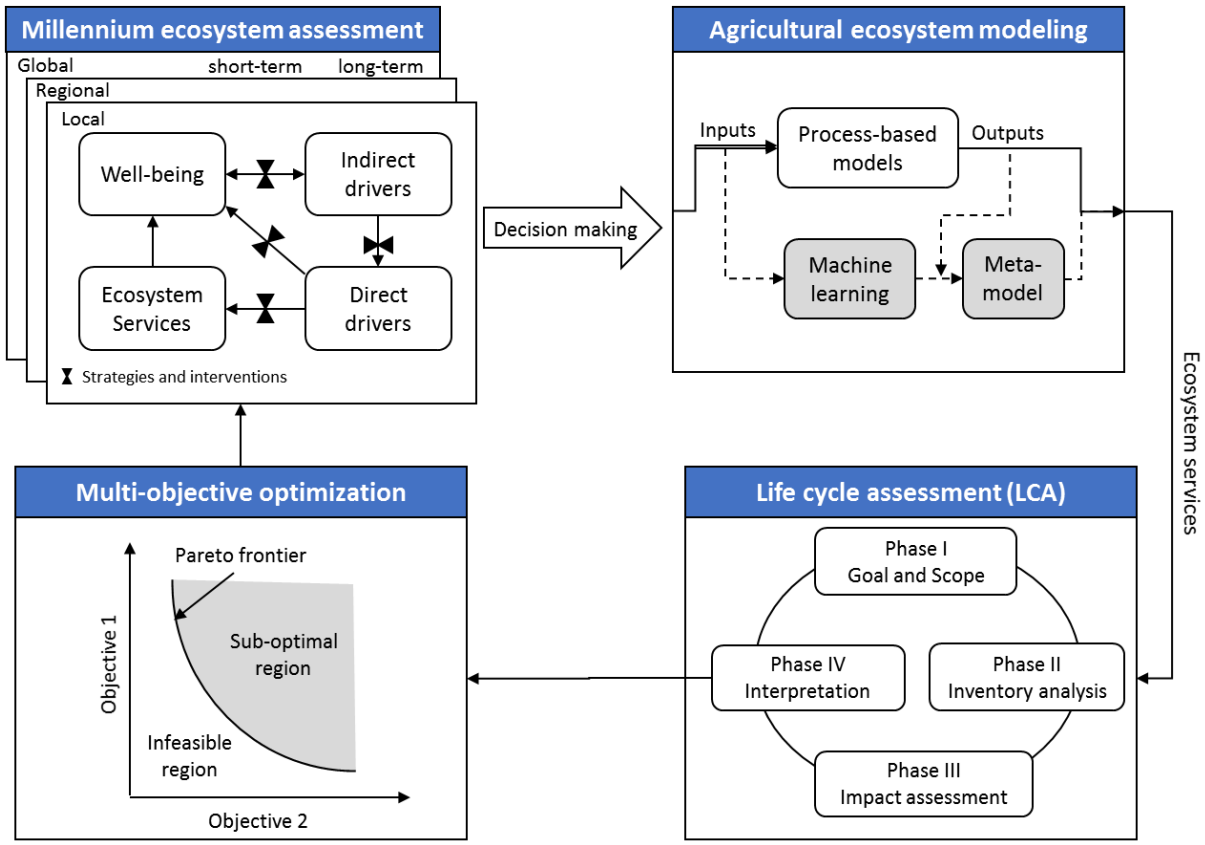


Fig. 1.3. The framework and modeling tools used in this dissertation.

Life Cycle Assessment

Life Cycle Assessment (LCA) is defined by the International Organization for Standardization (ISO) as studies of the environmental aspects and potential impacts throughout entire life cycle of a product, from raw material extraction and acquisition, through energy and material production and manufacturing, to use and end of life treatment and final disposal (Finkbeiner et al., 2006). The purpose of an LCA is to compare life cycles of products, processes, or services or identify parts of the life cycle where the greatest improvements can be made (Roy et al., 2009).

An LCA is typically carried out in four interdependent phases including goal and scope definition, life cycle inventory (LCI), life cycle impact assessment (LCIA), and interpretation (ISO 14044, 2006).

Goal and scope definition sets out the context of the study, the product or service to be assessed, the system boundaries, the reference system, the functional basis for comparison, the analysis assumptions and limitations, the allocation methods, and the impact categories. LCI involves the creating of an inventory list of all the inputs and outputs related to a "product" functional unit, collected from all activities within the system boundary. LCIA is conducted to evaluate the significance of potential environmental impacts based on the LCI results. For this, the LCI data are characterized and quantified into a limited number of impact categories (e.g., human toxicity, global warming, eutrophication) which may then be weighted for importance. The interpretation stage summarizes and evaluates the results from LCI and LCIA to draw meaningful conclusions and recommendations for decision-making. The integration of ES into LCA will likely require creating new impact categories with related calculable indicator(s) and characterization factors (Othoniel et al., 2016).

Process-based biophysical models

A process-based model is the mathematical representation of the underlying processes that characterize the functioning of well-delimited biological systems (Buck-Sorlin, 2013). Process-based biophysical models such as DayCent (Del Grosso et al., 2000), DNDC (Li et al., 1992), and APSIM (Keating et al., 2003) utilize a set of site-specific variables (e.g., biophysical conditions, cropping practices, and land use history) as inputs and produce a set of output variables (e.g., dry matter production, soil organic carbon, nitrous oxide emission) that describe the corresponding system states and functions. These models can capture the fine-scale influence

of site-specific weather conditions, soil properties, crop types, cropping practices, and land use history that determine the provision of ES (Nguyen et al., 2017). The DayCent (Del Grosso et al., 2000) and its monthly time-step version, CENTURY (Parton et al., 1994) models are used throughout this dissertation to quantify changes in ecosystem functions induced by changes in management decisions at farm level. Such results are then used for LCI and trade-off analysis at multiple scales. GIS and spatial databases are added to allow spatially-explicit assessments of agricultural landscapes.

Multi-objective optimization

An optimization includes finding the best solution for some problem given a specific objective function and a defined domain (inputs and constraints). When more than one objective is considered, the problem becomes a multi-objective optimization. A trade-off is created when these objectives conflict with each other. This is often the case of modern agricultural ecosystems where many competing objectives such as food and fiber production, carbon storage, climate change mitigation, water quality improvement, and biodiversity conservation are considered simultaneously. The multi-objective optimization typically generates a set of Pareto-optimal solutions (also called non-dominated solutions) where any improvement in one objective always impairs some others. The collection of all Pareto-optimal solutions is known as the Pareto frontier. Several multi-objective optimization algorithms are used in this dissertation including the min-max (Hwang et al., 1980), the weighted sum (Fishburn, 1967), and the ε-constraint (Haimes et al., 1971) algorithms.

Machine learning

Machine learning is a method of data analysis that automates analytical model building. Machine learning algorithms give computer systems the ability to "learn" information directly from data

without relying on a predetermined equation as a model. These algorithms are often divided into two categories including supervised and unsupervised. Supervised algorithms identified patterns or relationships between inputs and labeled outputs while unsupervised algorithms can infer a function describing a hidden structure from unlabeled data. There are many widely used Machine learning algorithms such as decision trees, kriging, polynomial functions, artificial neural networks (ANNs), radial basis functions, k-nearest neighbor, and support vector machines. The ANNs are employed in this dissertation via a metamodeling technique to create simplified surrogate models of DayCent to reduce the simulation time, expertise requirement, and data storage of our modeling platform.

1.4 ORGANIZATION OF DISSERTATION

According to Nahlik et al., (2012), although numerous ecosystem service frameworks exist in the literature, their actual penetration into the decision making process has rarely been observed due to obscure concept communication and non-practicality. McIntosh et al., (2011) pointed out that the communication of complex systems to stakeholders and decision makers as tools to support decisions is a big challenge. Therefore, all following chapters are structured as independent case studies balancing between the communication of the assessment methods/frameworks and the presentation of our modeling techniques. The chapters are written as standalone journal articles, with separate introductions and discussion sections putting the individual results and insights into a broader context.

The dissertation is organized as five separate chapters, including this introductory chapter to introduce and highlight the linkages between the following chapters. **Chapter 2** is a case study to evaluate the impact of ecosystem carbon stock change on greenhouse gas emissions and carbon payback periods of cassava-based ethanol in Vietnam. This chapter links CENTURY, a

biogeochemical model, with life cycle assessment (LCA) to explore the regionally-specific carbon stock changes associated with cassava cultivation for biofuel feedstock in Vietnam. **Chapter 3** is a trade-off analysis and optimization of ecosystem services and disservices for irrigated corn production systems in the South Platte River Basin, Colorado, USA. This chapters links DayCent, the daily time step version of the CENTURY model, with GIS and optimization algorithms for high-resolution spatially-explicit analyses of ecosystem services in agricultural landscapes. The derived modeling platform is named the "Agricultural Ecosystem Service Optimization" (Ag-EcoSOpt). **Chapter 4** identifies optimum feedstock landscape design for a hybrid corn grain- and stover-based ethanol production system at Front Range Energy biorefinery, Windsor, Colorado, USA. This chapter integrates LCA into the Ag-EcoSOpt to allow life-cycle-based optimization of ecosystem services. **Chapter 5** explores the trade-offs among seven management objectives of the irrigated corn production systems in Colorado, USA at different spatial scales using a surrogate-based optimization framework. This chapter equips the Ag-EcoSOpt with machine learning algorithms facilitating metamodeling techniques to alleviate the computational burden of large-scale optimizations of agricultural ecosystem services.

CHAPTER 2. IMPACT OF ECOSYSTEM CARBON STOCK CHANGE ON GREENHOUSE GAS EMISSIONS AND CARBON PAYBACK PERIODS OF CASSAVA-BASED ETHANOL IN VIETNAM[1]

2.1 SUMMARY

Cassava-based ethanol has been promoted in China and Southeast Asia as an effective means to reduce greenhouse gas (GHG) emissions and promote energy security. However, existing life cycle assessments of the environmental impacts of cassava ethanol have used highly-aggregated empirical methods to estimate ecosystem C stock changes, which do not capture finer-scale characteristics of different cassava growing regions within a country. We investigated carbon debts, GHG emissions, and payback periods for cassava-based ethanol in Vietnam using a life cycle assessment approach coupled with the widely-used ecosystem biogeochemical CENTURY model. The model simulated regionally-specific carbon stock changes associated with cassava cultivation for biofuel feedstock under different land use change, cassava yield and fertilization scenarios. We found that switching land use to cassava production for biofuel substantially reduced soil organic carbon in all major cassava growing regions in Vietnam. GHG emissions, carbon debts, and payback periods of Vietnam's cassava ethanol were strongly dependent on cassava yield. The mean carbon debt due to direct land use change to cassava production for ethanol ranged from 66 to 97 Mg of CO_2 per hectare, and the net carbon dioxide equivalent emission of cassava-based ethanol ranged from 36 to 95 g MJ^{-1}, depending on the range of

[1] This chapter was published in Biomass and BioEnergy Journal, Elsevier.
Nguyen, T.H., Williams, S., Paustian, K., 2017. Impact of ecosystem carbon stock change on greenhouse gas emissions and carbon payback periods of cassava-based ethanol in Vietnam. Biomass Bioenergy 100, 126–137. https://doi.org/10.1016/j.biombioe.2017.02.009

cassava fresh weight yield (from 18 to 60 Mg ha^{-1}). To repay a carbon debt from direct land use change within 25 years, the average fresh weight yield of cassava used as feedstock for ethanol production must be above 33 Mg ha^{-1}.

2.2 INTRODUCTION

Cassava (*Manihot esculenta* Crantz) has recently been promoted as a potential feedstock for ethanol production in Vietnam and many other parts of Asia, including China and Thailand, due to its high yield and low input requirements (Adelekan, 2011; Nguyen et al., 2007; Ou et al., 2009). Cassava can attain reasonably high yields with minimal irrigation and fertilization, with average fresh weight cassava yields in the region ranging from 17.2 – 60.0 Mg ha^{-1} (Le et al., 2013; Liu et al., 2013). The Vietnamese government has set a clear strategy for the development of biofuel through 2015 with a broad vision toward 2025. Pursuant to Decision No. 177/2007/QD-TTg of 2007, the biofuel output target for 2025 is 1.8 Mt $year^{-1}$, equivalent to 5% of projected total fuel demand, with 0.6 Mt $year^{-1}$ coming from cassava-based ethanol (Decision no. 177/2007/QD-TTg of Government of Vietnam, 2007).

However, questions remain about the overall benefits of cassava-based ethanol on greenhouse gas (GHG) reduction and potential negative impacts of cassava production on soil productivity due to soil erosion and high nutrient demand (Salami and Sangoyomi, 2013; Howeler, 1991; Howeler, 2000). Although GHG mitigation is not currently the main driving force for development of biofuels in Asia, the GHG footprint of cassava-ethanol and potential negative impacts of cassava as a biofuel feedstock are a concern to policymakers in developing countries and should be a consideration when implementing national biofuel expansion plans. Consequently, there have been a handful of life cycle assessment (LCA) studies carried out to assess the environmental performance of cassava-based ethanol in China, Thailand, and

Vietnam. LCA is defined by the International Organization for Standardization (ISO) as an analysis of the environmental impacts throughout the entire life cycle of a product, from raw material extraction and acquisition, to energy and material production and manufacturing, to use and end of life treatment and final disposal (Finkbeiner et al., 2006). Since LCA in the context of biofuels is still a relatively recent tool that requires further improvements and standardizations (Reap et al., 2008a; Reap et al., 2008b; Efroymson et al., 2013), the LCA results for cassava-based ethanol in the region vary dramatically between studies (Table 2.1).

Table 2.1. GHG balance for cassava-based ethanol in Vietnam: a comparison

Criteria	Gasoline (Wang et al., 2012)	Hu et al., (2004)	Liu et al., (2013)	Leng et al., (2008)	Nguyen et al., (2007)	Silalertruksa and Gheewala, (2011)	Le et al., (2013)
Country		China	Guangxi, China	China	Thailand	Thailand	Vietnam
LUC emission (g MJ^{-1})						36 – 222	33
GHG balance (g MJ^{-1})	94	73	20 – 74	734	46	63 – 313	35
% of GHG reduction (vs. gasoline)		23	79 – 22	-681	51	33 – (-233)	63

Le et al., (2013) indicated that this variation mainly stemmed from the differences in LCA assumptions, and the insufficiency in accounting for direct and indirect land use change (LUC) emissions, i.e., the emissions associated with changes in land use due to the cultivation of cassava for biofuel feedstock. Indirect land use change (iLUC) is the displacement of existing food production systems to new areas as a consequence of the land use change due to biofuel feedstock production. The quantification of iLUC is very complicated as it normally requires the linkage between the global economic equilibrium projections and the thorough analysis of land use change (Plevin et al., 2010). Therefore, it is often ignored in attributional LCAs like those conducted for cassava-based ethanol. Direct land use change (dLUC) deals with changes in carbon stock (C

stock) including vegetation carbon (Cveg) and soil organic carbon (SOC) due to the introduction of a new cropping system for biofuel feedstock production that replaces the existing land use at a specific site. Direct land use change becomes a problem when carbon-rich ecosystems, such as forests and grasslands, are converted to produce feedstock for biofuels. These conversions release a significant amount of the carbon that had previously been stored in these ecosystems into the atmosphere, creating a "carbon debt" that subsequent biofuel production will have to offset. The LUC effects are induced by the production of biofuels and thus must be carefully factored into the accounting of biofuel GHG emissions. Changes in Cveg (aboveground and belowground live carbon) are easier to measure or estimate while SOC change is more difficult and more expensive to quantify (Bandaranayake et al., 2003). Therefore, SOC has often been ignored in many early LCA studies of cassava ethanol (Nguyen et al., 2007; Liu et al., 2013; Hu et al., 2004; Leng et al., 2008).

More recent studies have attempted to address this problem by using simple (Tier 1) empirical models developed for the Intergovernmental Panel on Climate Change (IPCC) guidelines for national GHG inventories to estimate SOC change due to changes in generic land use types (Le et al., 2013; Silalertruksa and Gheewala, 2011; IPCC, 2006). Tier 1 soil C methodology uses broad averages for nine globally-defined climatic regions and six classes of soils (IPCC, 2006), and was designed to support basic national inventory accounting in countries with minimal data and scientific resources. Thus, this approach is poorly-suited for quantifying more locally-specific impacts on C stock change associated with local weather conditions, soil properties, cropping practices, and land use history of a specific location. In addition, the rates of SOC change are typically high following land use conversion, and then gradually decrease as SOC approaches a new stable state. This dynamic transition affects how GHG emissions associated

with feedstock production are analyzed. The Tier 1 method has a fixed time horizon of 20 years, over which a constant rate of change is assumed, and thus the method does not capture the more dynamic and longer-term changes in soil C stocks that are of interest for computing soil carbon debts related to land use change.

An alternative solution for estimating soil C stock changes is to employ process-based models, such as CENTURY (Parton et al., 1994), DAYCENT (Del Grosso et al., 2000), DNDC (Li et al., 1992), or RothC (Coleman and Jenkinson, 1996). Although these biogeochemical models were originally developed for research purposes at field-scale, many of them have been used effectively for soil C or GHG estimation at regional to national scales (Ogle et al., 2010; Del Grosso et al., 2010). Compared to aggregated empirical models like the IPCC Tier 1 methods, these process-based models can better represent site-specific conditions by integrating non-discrete variables, such as temperature and moisture, and the interactions of biophysical characteristics (e.g., climate and soil characteristics). In addition, these models can capture the historical influence of land use and management practices on potential future changes in soil C stocks (Paustian et al., 2010).

We investigated GHG emissions from the production stage, carbon debts, and carbon payback periods of cassava-based ethanol in Vietnam with an emphasis on soil C stock changes due to dLUC. A process-based model (CENTURY) was used to simulate these changes. The model takes into account more site-specific factors including climate, soils, cropping practices, land use history, and dLUC and produces time series of C stock changes based on dynamic simulation of carbon and nitrogen cycling processes in the soil–plant system.

2.3 METHODS

2.3.1 Modeling the emissions from C stock changes caused by dLUC

The data on dLUC to cassava cultivation for biofuel feedstock in Vietnam were obtained from the survey conducted by Le et al., (2013), and then aggregated into percentages of dLUC to cassava by main growing region (North Central, South Central, Central Highlands, and Southeast) and by baseline land use type within each region (annuals, perennials, forest, and grassland) (Table 2.2). Annuals are defined as land used for crops with a less than one-year growing cycle and which must be newly sown or planted for further production after the harvest (e.g., rice, maize, wheat, millet, sugarcane, reed, and jute). Perennials are defined as land cultivated with long-term crops which do not have to be replanted for several years (e.g., tea, coffee, rubber, and citrus). Forests are land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ, and grassland are land covered with herbaceous plants with less than 10 percent tree and shrub cover (FAO, 2014). The barren land, and denuded hills were also included in the grassland. The CENTURY model (Parton et al., 1994) was then used to model the C stock changes to cassava for the four baseline land use types in the four cassava growing regions. The results were then used to compute the mean carbon debts and annualized emissions from C stock changes caused by dLUC to cassava production for biofuel feedstock in Vietnam. The main steps of the modelling process included: (1) model inputs and parameterization; (2) model simulation; and (3) calculation of carbon debts and annualized emissions from C stock changes caused by dLUC (Fig. 2.1).

Table 2.2. Percentages of dLUC to cassava cultivation for ethanol production in Vietnam

Region	Percentage of dLUC to cassava by growing region	Percentage of dLUC to cassava by land use type in each growing region			
		Annuals	Perennials	Forest	Grassland

North Central	8	40.0	12.5	2.5	45.0
South Central	26	32.0	24.0	10.0	34.0
Central Highlands	39	16.3	17.5	20.0	46.3
Southeast	27	36.7	50.0	5.0	8.3

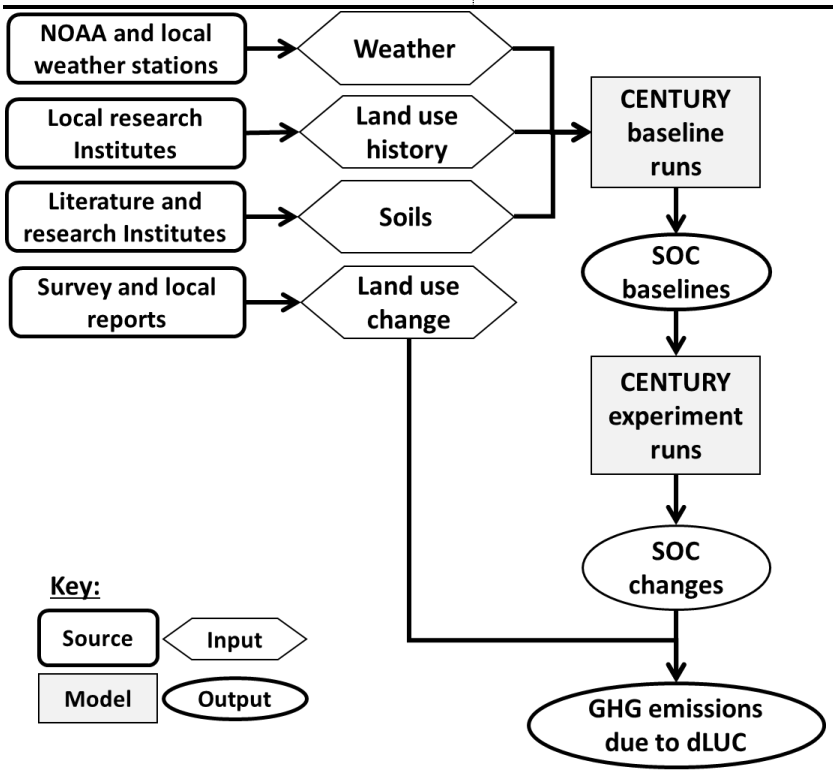


Fig. 2.1. Overview of model inputs and simulation process.

2.3.1.1 Model descriptions

The CENTURY model is a general ecosystem biogeochemical model that is widely used to simulate ecosystem carbon and nitrogen dynamics (Parton et al., 1988). It is a multi-compartmental model composed of a soil organic matter/decomposition submodel, a water budget model, a grassland/crop submodel, a forest production submodel, and routines to simulate a large suite of land management practices (e.g., plant species, planting, harvest, tillage, fertilization, irrigation, grazing, etc.) and disturbance events (e.g., land clearing, fire). CENTURY uses relatively coarse time steps (monthly for soil organic C and N processes; weekly for plant growth and water balance calculations) and thus is intended for simulating

medium- to long-term (annual to centuries) changes in SOC and other ecosystem parameters in response to changes in climate, land use, and management (Bandaranayake et al., 2003). CENTURY and DAYCENT (the daily time-step version of CENTURY (Parton et al., 1998)) have been successfully applied to various ecosystems in various locations throughout the world (Cheng et al., 2013; Cerri et al., 2007; Davis et al., 2011; Paustian et al., 1997). Relatively few data inputs are required for CENTURY, including monthly maximum/minimum air temperature and precipitation, surface soil texture class, and land cover/use data (e.g., vegetation type, cultivation/planting schedules, amount and timing of nutrient amendments), which suited the conditions of our study (e.g., limited daily weather data for the research locations and study regions).

2.3.1.2 Model inputs and parameterization

Averaged monthly precipitation, maximum, and minimum temperatures observed at stations in Vietnam were obtained from the Statistics and Food Security Database (MARD, 2015a) for 12 years (2001 - 2013) and then transformed into appropriate formats for CENTURY. We used data from five provincial weather stations to represent the four geographical regions (Fig. 2.2). There are two weather stations in the North Central region due to its location in a transitional zone between two distinct regional weather patterns.

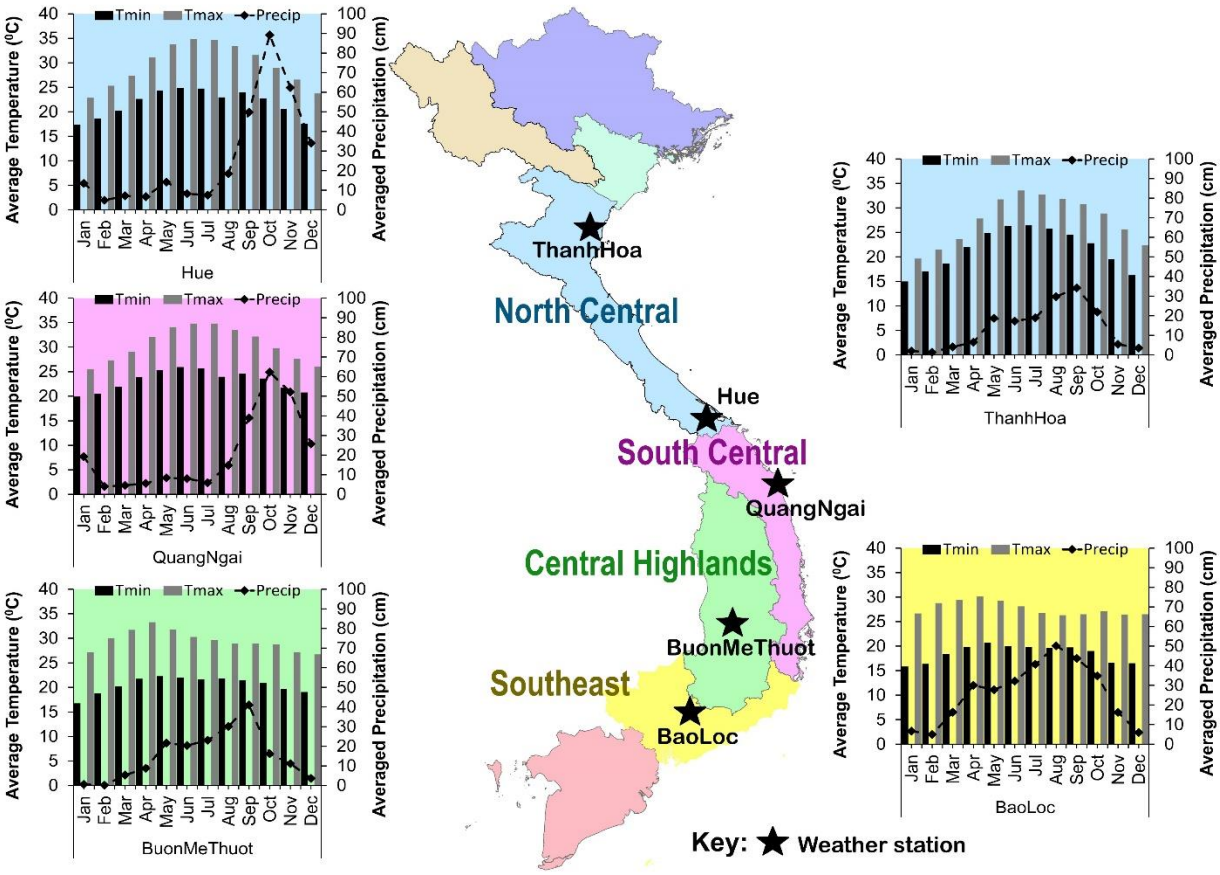


Fig. 2.2. Average monthly climate data of five weather stations in Vietnam.

Table 1.3. Properties of major soil types in Vietnam used to parameter CENTURY

Soil types	Taxonomy	pH	Sand (%)	Silt (%)	Clay (%)	Bulk density (g cm^{-3})	Drainage	Average OC (%)	Source
Sandy	Arenosols	5.3	77	12	11	1.51	0.9	0.7	(H. Hoang et al., 2010)
Alluvial	Fluvisols	4.6	23	40	37	1.31	0.8	2.2	(Klinnert and others, 2001)
Red	Ferralsols	4.0	20	27	53	0.85	0.6	5.0	(Klinnert and others, 2001; D'haeze et al., 2003)
Grey Degraded	Haplic Acrisol	4.3	64	15	21	1.34	0.8	1.0	(Nguyen and Egashira, 2008)
Ferralitic	Other Acrisol	3.9	40	22	38	0.85	0.7	2.3	(Husson et al., 2001)

Major soil types and their properties were obtained from literature (Table 2.3) and were coupled with weather data to parameterize the site conditions in CENTURY. We assumed these soil characteristics as typical for all cassava growing regions.

To model the C stock of the baseline land use types before dLUC to cassava cultivation, specific sets of soils were assigned to the different land use types for each cassava growing region (Table 2.4) (Ha, 2010). We modeled broad classes of land use such as annuals, perennials, forest, and grassland and assume that each land use type could be represented by the dominant crop/tree species. Maize, peanut, sugarcane, cassava, and rice were selected to represent annuals. Their cropping proportions in each region (Table 2.4) were calculated based on the planted area data obtained from the General Statistics Office of Vietnam (GSOV, 2014). Mixed fruit tree, secondary tropical forest, and mixed grass were used to represent perennials, forest, and grassland, respectively. Due to the unavailability of soil area proportions for the annuals land use type, modeled crops were assumed to be evenly distributed across the corresponding soil types. The SOC of the annuals was averaged based on the cropping proportions of the representing species.

Management practices for various crops used in our model simulation, such as planting and harvest dates, fertilization, and irrigation, were obtained from Vietnam's National Center for Agricultural Extension databases (MARD, 2015b), as summarized in Table 2.5.

Table 2.4. Modeled species with their cropping proportions by region

Region	Land use type	Soil type	Simulated species	Cropping proportion
Northern Central	Annuals	Sandy, Alluvial Grey degraded, Ferralitic	Maize	0.12
			Peanut	0.06
			Sugarcane	0.06
			Cassava	0.07
		Alluvial	Rice*	0.69
Southern Central	Annuals	Sandy, Alluvial, Grey degraded, Ferralitic	Maize	0.1
			Peanut	0.04
			Sugarcane	0.07
			Cassava	0.13
		Alluvial	Rice	0.66
Central Highlands	Annuals	Red, Grey degraded Ferralitic, Alluvial	Maize	0.36
			Peanut	0.03
			Sugarcane	0.07
			Cassava	0.21
		Alluvial	Rice	0.33
South East	Annuals	Alluvial, Grey degraded, Red, Ferralitic	Maize	0.15
			Peanut	0.04
			Sugarcane	0.07
			Cassava	0.18
		Alluvial	Rice	0.56
All regions	Perennials	Ferralitic/Red**	Mixed fruit tree	1.00
	Forest	Ferralitic	Secondary tropical forest	1.00
	Grassland	Ferralitic	Mixed grass	1.00

*Note: * Rice is only modeled on alluvial soil*
*** Perennials are modeled on both ferralitic and red soils in Central Highlands.*

Table 2.5. Management practices of the modeled species

Species	Management				
	Growing season[a]	Manure addition (Mg ha^{-1})	Fertilization (kg N ha^{-1})	Irrigation (%)[c]	Residue removal (%)
Maize	2	5	100	90	75
Peanut	2	3	30	90	25
Sugarcane	8-year rotation	8	100	95	95
Cassava	1	5	58	75	79
Rice	2	5	100	100	75
Mixed fruit tree	8-year rotation	5	120	95	5[b]

Note: [a] Number of growing seasons per year unless specified
[b] Mixed fruit tree is pruned annually and clearcut at the end of the rotation.
[c] Percent of field capacity irrigated

We mainly used CENTURY default parameters (Metherell et al., 1993) for each of the ecosystems and crop types in our analysis. Because of the significant impacts of vegetation growth processes on the prediction of SOC dynamics and GHG emissions (Cheng et al., 2014), we calibrated the CENTURY's growth submodels against the reported values from literature for

the studied vegetation types. PRDX is a parameter used for calculating potential aboveground monthly production as a function of solar radiation, thus representing the genetic potential production of different vegetation types or crop varieties. We adjusted the PRDX for each vegetation type and simulated growth under different regional climate, soil conditions, and standard management practices. Modeled yield or Cveg of each species at national-scale were calculated by averaging the regional values, and compared with the literature reported data (FAO, 2013; Zemek, 2009) to choose the best values for PRDX. Due to the lack of field experiments measuring SOC as a function of land management practices in Vietnam, we were not able to directly validate model predictions of SOC changes for the various crop and land use types.

2.3.1.3 Scenarios for the analysis of expanded cassava production

For a baseline scenario, we used Vietnam's average cassava yield reported by FAO in 2013 (18 Mg ha^{-1}) (FAO, 2013), and a fertilizer/pesticide application rate (per ha) of 58 kg of N, 53 kg of P_2O_5, 47 kg of K_2O, and 0.23 kg of pesticide (Le et al., 2013). Cassava fields were moderately irrigated at 75% field capacity using auto-irrigation function in CENTURY at the time of planting. The corresponding baseline Fertilizer-to-Yield ratios ($FTY_{baseline}$) were 3.22 kg Mg^{-1}, 2.94 kg Mg^{-1}, 2.61 kg Mg^{-1}, and 0.013 kg Mg^{-1} for N, P_2O_5, K_2O, and pesticide, respectively.

To represent the relation between yield increase and fertilization rates, we used a simple linear equation:

$$F_i = (\Delta Y)*FTY_i + F_{i\_baseline}, \qquad \text{Eq. 2.1}$$

where F_i, FTY_i, $F_{i\_baseline}$ are the application rate (kg ha^{-1}), Fertilizer-to-Yield ratio (kg Mg^{-1}), and the baseline application rate (kg ha^{-1}) of the fertilizer (or pesticide) *i*, respectively; ΔY (Mg ha^{-1}) is the difference between the modeled and the baseline yield.

We then performed an analysis using a range of the reported cassava yield (from 18 to 60 Mg ha^{-1}) and a range of FTY (from 50 to 150% of the $FTY_{baseline}$) to examine the effects of cassava yield improvements and changes in the corresponding fertilizer application on GHG emissions, carbon debts, and carbon payback periods. The scenarios with reduced FTY represent the effects of future cassava breeding advances and/or improved N management that enhance the yield response to nitrogen, while the scenario with the highest FTY represents potentially less efficient N utilization as yields increase. The simulated yield and fertilization ranges are specified in Table 2.6. Since most of cassava production area in Vietnam is rain-fed, we did not consider the variation in the irrigation rate. Instead, irrigation (irrigated at 75% field capacity) and other management practices used in this study such as farmyard manure application (5 Mg ha^{-1}) were kept constant throughout the sensitivity analysis. In addition, we assumed that these management practices were at average levels and were applied uniformly for all cassava growing regions in Vietnam.

Table 2.6. Modeling scenarios

	FTY[a] range	Fertilizer type	Lower bound	Upper bound	Fertilization rate equation[b]
Yield (Mg ha^{-1})			18	60	
Fertilizer inputs (kg ha^{-1})	50% $FTY_{baseline}$	N	58	126	$F_N = \Delta Y * 1.61 + 58$
		P	53	115	$F_P = \Delta Y * 1.47 + 53$
		K	47	102	$F_K = \Delta Y * 1.31 + 47$
		Pesticide	0.23	0.50	$F_{pesticide} = \Delta Y * 0.0065 + 0.23$
	150% $FTY_{baseline}$	N	58	261	$F_N = \Delta Y * 4.83 + 58$
		P	53	239	$F_P = \Delta Y * 4.41 + 53$
		K	47	212	$F_K = \Delta Y * 3.92 + 47$
		Pesticide	0.23	1.04	$F_{pesticide} = \Delta Y * 0.02 + 0.23$

Note: [a] Fertilizer-to-Yield ratio

[b] See Eq. 2.1

2.3.1.4 Model simulations

Prior to modeling SOC under the cassava system, we set up the baseline SOC for all land use types in all cassava growing regions with CENTURY. Baseline SOC reflects the SOC level under the current management practice, and is used to evaluate the magnitude and direction of SOC change due to dLUC to cassava production for ethanol. The baseline simulation is comprised of two stages: an initial model spin-up followed by a simulation of historical land use. The spin-up simulation sets the equilibrium for SOC under natural systems with little or no disturbance from human activities, with tropical forest used as the native vegetation. To run the model to equilibrium SOC, we simulated tropical forest with default CENTURY parameters for 10,000 years under locally specific soils and climate data for each region. Subsequently, the model was run from this equilibrium condition for each specific historical land use, including annuals, perennials, forest, and grassland for each region. A period of 68 years (1945 to 2013) was used for the historic and current land use period. The year 1945 was selected because it marked the beginning of broad-scale forestland conversion to agricultural use, brought on by numerous agricultural policies. The SOC stocks at the end of the historic/current period were used as a baseline for comparison with the future projection scenarios of conversion to cassava production. In all cases, the baseline SOC reached new stable states before 2013. The experimental simulation started in 2014 when cassava was grown as a replacement for each land use type. We then modeled the changes in SOC over time to identify new SOC stable states under the cassava growing system. The net changes between baseline and cassava SOC stable states, in addition to changes in Cveg between the two systems, were used to calculate the carbon debts and annualized GHG emissions due to dLUC to cassava production.

2.3.1.5 Calculation of carbon debts and annualized emissions from C stock changes caused by dLUC

The time frame (in years) chosen for our study was the longest transition time for a new SOC stable state after converting from a baseline land use to cassava land use. This provides a common time frame, beyond which there are no significant net changes in ecosystem C stocks, i.e., emissions from dLUC for all land use change scenarios are accounted for within the time frame. Since changes towards a new steady-state soil C stock (under constant climate and management conditions) are asymptotic, we defined SOC as reaching an approximate new equilibrium after land use conversion (at t=0) once a 5-year period SOC change under the new land use was less than 1%, i.e.,

$$\frac{|\,SOC(t-5) - SOC(t)\,| * 100}{SOC(t-5)} < 1\%, \qquad \text{Eq. 2}$$

where SOC(t-5) and SOC(t) are SOC at time t-5 and t, respectively. The time (t) in years for the longest time interval to a new approximate steady-state as defined in Eq. 2 (for any of the land use conversions to cassava) was chosen as the LCA accounting time frame (T).

The carbon debt due to dLUC from a specific land use to cassava is the difference in C stock between that land use and the cassava system at their SOC stable states. Our study reported carbon debts in the unit of Mg of carbon dioxide per hectare. The annualized emissions from C stock changes caused by dLUC (i.e., the annualized dLUC emissions) were calculated by dividing the mean carbon debts by the accounting time frame (T). We separately calculated the carbon debt for each baseline land use in each region, and then computed the regionally-averaged carbon debts based on these individual values, weighted by the percentages of dLUC to cassava in each growing region (Table 2.2). Finally, we calculated the mean carbon debt from

the regionally-averaged values weighted by the percentages of dLUC to cassava by growing region (Table 2.2).

2.3.2 Calculation of net GHG emissions and payback periods

2.3.2.1. Calculation of net GHG emissions

Our study used a standard life cycle assessment approach (Gnansounou et al., 2009a) to calculate the net GHG emissions for cassava-based ethanol in Vietnam. We considered three GHGs including carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (CH_4), and reported the GHG emissions in CO_2 equivalent (CO_2eq) by aggregating the individual GHG using their 100-year global warming potential values (GWP_{100}) (IPCC, 2006). The functional unit used to compare the environmental performance of ethanol and conventional gasoline (CG) was one megajoule (MJ). The CO_2eq emission of CG used in our study was 94 g MJ^{-1} (Wang et al., 2012). Our analysis employed the cassava-based ethanol production system and associated life cycle inventory values described in Le et al., (2013) (Fig. 2.3).

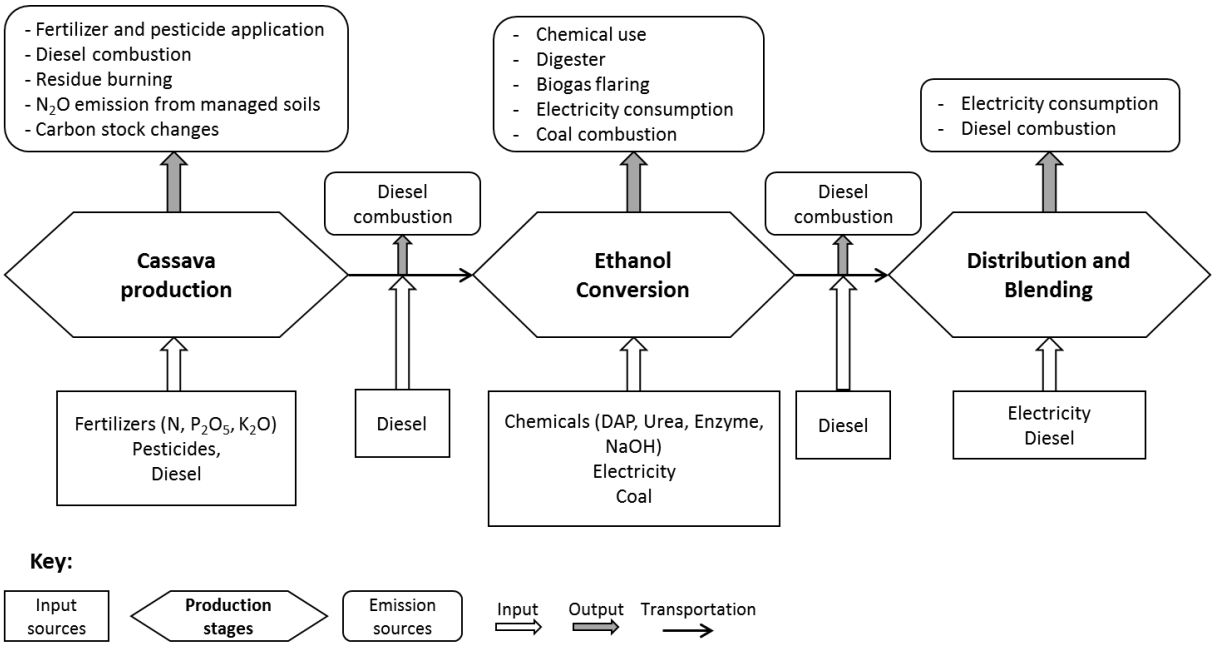


Fig. 2.3. Cassava-based ethanol production system in Vietnam.
Figure adopted from Le et al., (2013).

Net GHG emissions (E) from production and utilization of cassava-based ethanol were comprised of three components, stemming from: a) cassava cultivation (E_{cp}), b) ethanol conversion (E_c), and c) transport and distribution (E_{td}). Since determining the emissions from processing and transport (E_c and E_{td}) was not the focus of our study, we used the results reported in Le et al., (2013). The authors employed the guidelines from the IPCC (IPCC, 2006), the Harmonized Calculations of Biofuel Greenhouse Gas Emissions in Europe (BioGrace) (BioGrace, 2015), and Renewable Energy Directive (RED) (Carré and Intergovernmental Panel on Climate Change, 2010) to calculate net CO_2eq emissions of 129 g L^{-1} for the conversion process and 31 g L^{-1} for the distribution and blending process of cassava-based ethanol in Vietnam.

We estimated GHG emissions from the cassava cultivation stage (E_{cp}) according to equation (3).

$$E_{cp} = E_{fertilizer} + E_{pesticide} + E_{diesel} + E_{burning} + E_{N2O} + E_{dluc}, \quad \text{Eq. 3}$$

where $E_{fertilizer}$ is the CO_2eq emission from fertilizer use (g L^{-1}), $E_{pesticide}$ the CO_2eq emission from pesticide use (g L^{-1}), E_{diesel} the CO_2eq emission from diesel consumption (g L^{-1}), $E_{burning}$ the CO_2eq emission from burning of cassava residue (g L^{-1}), E_{N2O} the CO_2eq emission of N_2O from managed soils (g L^{-1}), and E_{dluc} the annualized CO_2eq emission from C stock changes caused by dLUC (g L^{-1}).

For $E_{fertilizer,}$ $E_{pesticide}$, E_{diesel}, and $E_{burning}$ we used emission factors (EFs) reported in the IPCC (IPCC, 2006) and BioGrace guidelines (BioGrace, 2015). E_{dluc} calculations were described above (section 2.1.5). Direct N_2O emissions from managed soils were calculated from the amount of

organic and synthetic N applied as fertilizers, the N in crop residues returned to soils, and any additional mineralized N associated with net decreases in total soil organic N stocks (IPCC, 2006). Indirect N_2O emissions were computed from both fractions of volatized N and leaching and runoff of N estimated in the CENTURY simulations. Direct and indirect N_2O emissions were computed using emission factors in the BioGrace GHG calculation excel tool (version 4c) (BioGrace, 2015).

We allocated total GHG emissions between the main product (ethanol) and coproducts and calculated the GHG offset credits (if any) from the utilization of the coproducts. The coproducts of the cassava-based ethanol production are distiller's dried grains with solubles (DDGS), biogas, cassava processing residues, and CO_2 (Nguyen et al., 2007; Le et al., 2013; Silalertruksa and Gheewala, 2011; Leng et al., 2008). According to Leng et al., (2008), 18% of total GHG emissions from the production of cassava-based ethanol can be allocated to DDGS. The biogas from the ethanol conversion process is utilized for combustion as a supplemental input energy source (Le et al., 2013). Due to the high moisture content (86%) of the cassava processing residues, the energy and emission credits of this coproduct used for the replacement of a farm fertilizer application was offset by the emissions generated during the transportation of this bulky material from ethanol plants to farms (Nguyen and Gheewala, 2008). Le et al., (2013) reported that a CO_2eq emission of 490 g L^{-1} are being collected from the fermentation process for further use in other industries, such as those for food and beverage. We argue that this is not a real GHG emission offset because CO_2 is only temporarily captured and will be released back into the atmosphere as soon as the food and beverage products are consumed. Furthermore, there is no information regarding the compression and transportation of the collected CO_2 to other industries. We, therefore, did not consider any GHG offset credit for the CO_2 coproduct.

2.3.2.2 Calculation of carbon payback period

Carbon payback period for cassava-based ethanol represents the number of years it takes for the carbon savings from avoided fossil fuel combustion (if any) to offset the losses in ecosystem carbon from clearing land to grow cassava for biofuel feedstock (Gibbs et al., 2008). Carbon payback period (R) was computed according to equation (2.4):

$$R = \frac{D}{S*P*LHVe}, \qquad \text{Eq. 2.4}$$

where D is the mean carbon debt (Mg ha^{-1}) calculated in section 2.1.5, S is the avoided CO_2eq emission due to replacing conventional gasoline (CG) with ethanol (Mg MJ^{-1}), P is the annual ethanol productivity (L ha^{-1} $year^{-1}$), and LHVe is the low heat value of ethanol (21.1 MJ L^{-1}). S is calculated as the difference between GHG emission of CG and that of ethanol, and must be greater that zero for carbon payback period to be realized. S*P* LHVe is the annual emission reductions by production and utilization of ethanol. The conversion ratio of fresh cassava root (kg) to ethanol (L) used in our study was 6:1 (Sriroth et al., 2012).

2.4 RESULTS

2.4.1 Calibration of CENTURY'S predicted yield and biomass production

Table 2.7 shows the CENTURY modeled yields and Cveg together with reported yields from literature (in Mg of carbon per hectare) that were used for model calibration. The average modeled yields of maize, peanut, rice, cassava, and sugarcane were compared with the average yields reported by FAO (2013) for Vietnam (FAO, 2013). FAO reports crop yields in the marketed forms of the crops in each country, which are clean and dry weight for cereals (e.g. rice, maize) and oil-bearing crops (e.g. peanut), and clean weight with relative high moisture contents for roots and tubers (e.g. cassava) and sugar crops (e.g. sugarcane). We reported yields

in term of carbon fraction (yield-C) by assuming the carbon content of 40%, and the moisture contents of 65%, 70% and 15% for cassava root, sugarcane stalk, and other annual crops (including maize, peanut, and rice), respectively (Sriroth et al., 2012; Bakker, 2012). The modeled Cveg of mixed fruit tree (representing perennials), secondary tropical forest (representing forest), and mixed grass (representing grassland) were compared with those of mixed fruit tree, grazed secondary forest, and perennial grass (*Chromolaena odorata L.)* reported by Zemek, (2009) for the Chieng Khoi watershed in Vietnam. The results showed that CENTURY was able to adequately represent plant productivity of the chosen species.

Table 2.7. Comparison between modeled yield-C and Cveg with calibrating literature data

Species	CENTURY's modeled outputs (Mg ha^{-1})*				Calibrating data (Mg ha^{-1})	
	North Central	South Central	Central Highlands	Southeast	FAO, (2013)	Zemek, (2009)
Maize	1.3	1.2	1.8	1.6	1.5	
Peanut	0.5	0.3	1.2	1.3	0.7	
Rice	1.5	1.1	2.6	2.5	1.9	
Cassava	2.4	2.1	2.7	2.9	2.5	
Sugarcane	7.5	7.8	7.6	7.5	7.8	
Mixed grass	5.1	4.4	5.7	4.4		4.6
Mixed fruit tree	12.4	11.6	16.9	16.0		13.0
Secondary tropical forest	59.4	58.2	62.4	53.1		60.0

*Note: *All the unit are reported in term of Mg of carbon per hectare.*

2.4.2 Changes in C stock due to cassava production for ethanol in Vietnam

Simulated trends in SOC changes for a 40-year time period following conversion to cassava are presented in Fig. 2.4. The SOC before conversion to cassava varied by different growing regions and baseline land use types. In general, the SOC levels were highest in the southeast and lowest in the north central region for all the baseline land uses. For the different land use types, baseline SOC stocks generally differed in the order forest ≈ perennials > grassland > annuals. Difference between the forest and grassland vs annual crop baseline values were least for the central highlands region.

The conversion of land use to cassava for ethanol decreased SOC for most baseline land uses in all regions (Fig. 2.4). The main driving factor is the low carbon inputs from cassava biomass residuals (annual crop residuals are often burned onsite), particularly for the conversions from non-annual crops where the decrease in C inputs, along with higher decomposition rates with tillage, caused the greatest decline in SOC. Only the conversion from other annual crops to cassava slightly increased SOC at cassava yields of 57 Mg ha^{-1} and above.

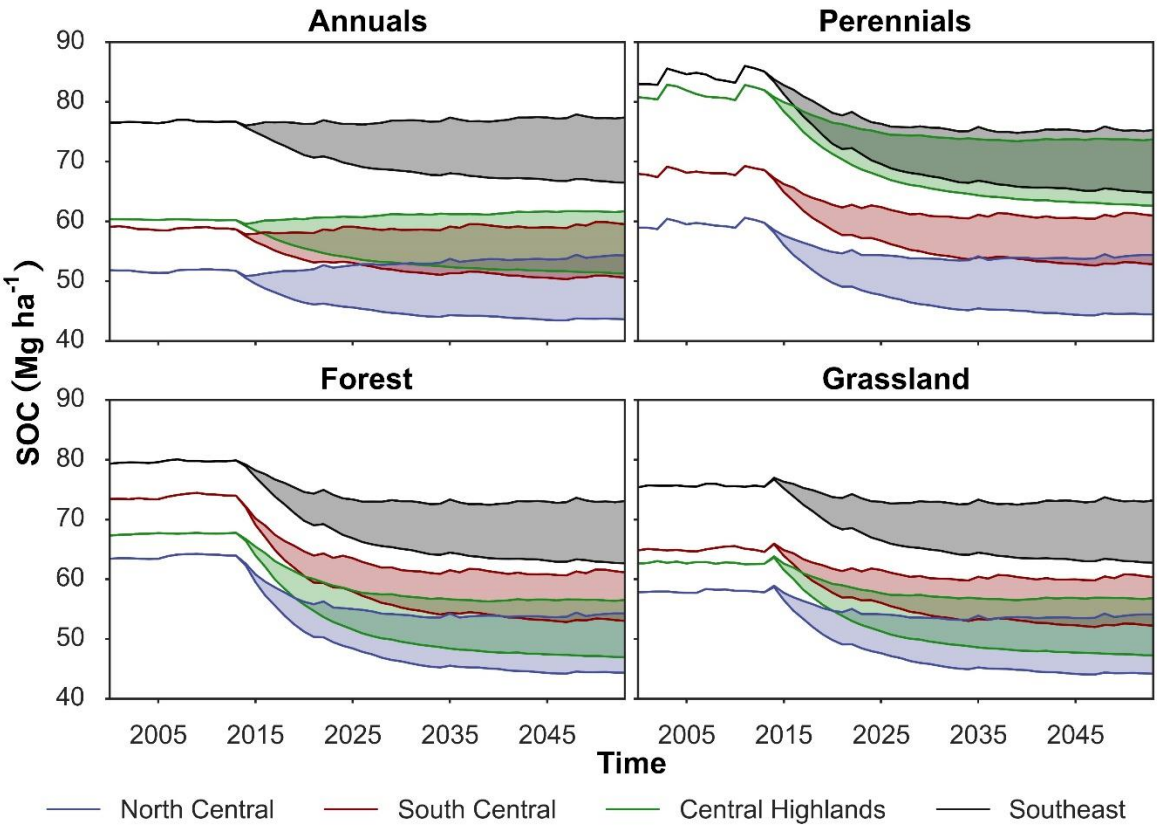


Fig. 2.4. Changes in SOC due to dLUC to cassava cultivation for bioethanol in Vietnam. The filled areas are the SOC difference between the cassava fresh wt. yield of 18 Mg ha^{-1} and 60 Mg ha^{-1}.

The predicted rates and magnitudes of SOC changes due to dLUC varied by different cassava growing regions and varied according to baseline land uses, due to a number of factors including

cassava yield, management practices, soil type, and weather. The rates of SOC change tended to be highest in forest, followed by grassland and perennial systems. Most SOC losses occurred during the first 5 - 10 years after conversion to cassava. Forest, grassland, and perennials require more time than annuals to approach a new SOC stable state after land clearing for cassava cultivation. In all scenarios, SOC tended to reach new approximate stable states within the maximum of 25 years. We, therefore, selected 25 years as the time frame for our analysis. This time frame is also appropriate for the discussion of biofuels policy or budget planning. The difference in SOC between the baseline land uses and cassava cultivation at 25 years after conversion ranged from 7 – 20 Mg ha^{-1} at cassava yields of 18 Mg ha^{-1}, and (-2) to 12 Mg ha^{-1} at the maximum considered yield of 60 Mg ha^{-1} (Fig. 2.5). In most regions (except the Southeast), the highest carbon loss occurred when clearing forest to grow cassava, followed by perennials, grass, and annuals. SOC stocks under cassava tended to increase as projected yields (hence also plant residues) of cassava increased. The difference in SOC stocks between minimum and maximum considered cassava yields ranged from 8 – 11 Mg ha^{-1}.

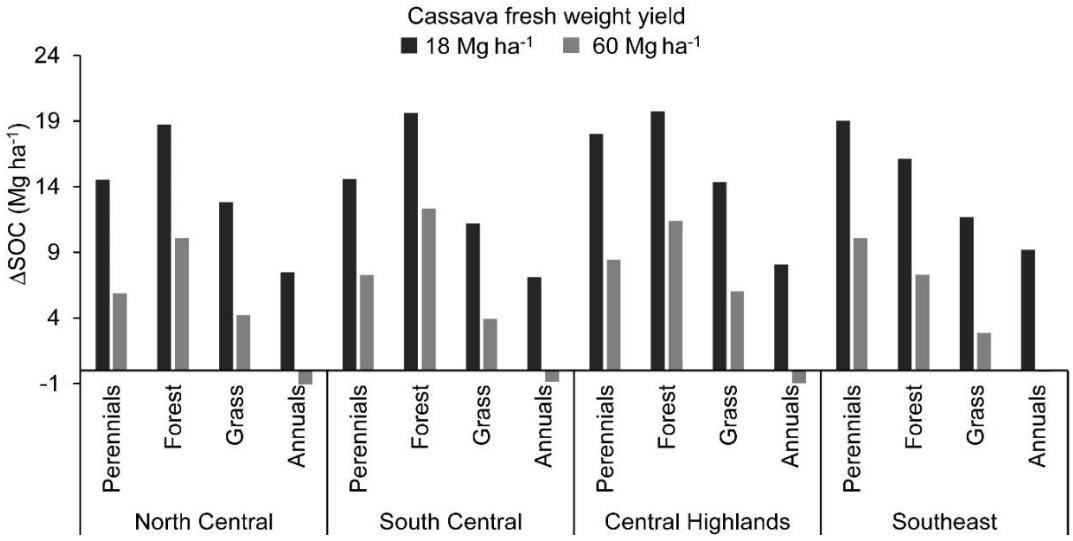


Fig. 2.5. Difference in SOC between baseline land use and cassava at 25 years after conversion. Negative values of ΔSOC indicate soil carbon increases.

The mean carbon debt of cassava-based ethanol in Vietnam strongly depended on cassava yield (Fig. 2.6). Higher yields reduced the carbon debts due to decreases in SOC loss. The mean carbon debt due to dLUC to cassava production for ethanol declined linearly from 97 to 66 Mg ha^{-1} corresponding to the increase in cassava yield from 18 to 60 Mg ha^{-1}.

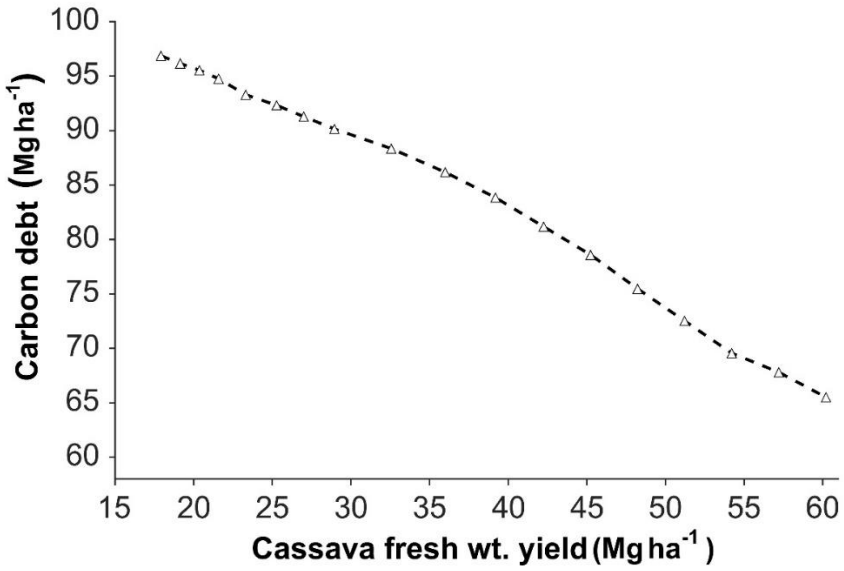


Fig. 2.6. Mean carbon debt as a function of yield for cassava-based ethanol in Vietnam.

2.4.3 Greenhouse gas emission of cassava-based ethanol in Vietnam

Annualized dLUC GHG emissions of cassava-based ethanol in Vietnam ranged from 10 to 50 g MJ^{-1}, and constitutes about 23 - 53% of total net GHG emissions, depending on cassava yield and fertilizer application (Fig. 2.7). The net GHG emission of cassava-based ethanol decreased in conjunction with an increase in cassava yield and a decrease in fertilization. At the cassava yield level of 18 Mg ha^{-1}, the net GHG emission was 95 g MJ^{-1}, indicating no avoided GHG emission when replacing CG with ethanol. In fact, the displacement of CG by cassava-based ethanol increased the total emission by 1 g MJ^{-1} at this yield level. With the least fertilizer addition (50% $FTY_{baseline}$), increasing yield from 18 to 60 Mg ha^{-1} decreased the net GHG emission from 95 to 33 g MJ^{-1}, and increased the avoided GHG emission from (-1) to 61 g MJ^{-1}. The GHG emission

difference between 50% $FTY_{baseline}$ and 150% $FTY_{baseline}$ was higher at higher cassava yield levels and reached a maximum at 13 g MJ^{-1}.

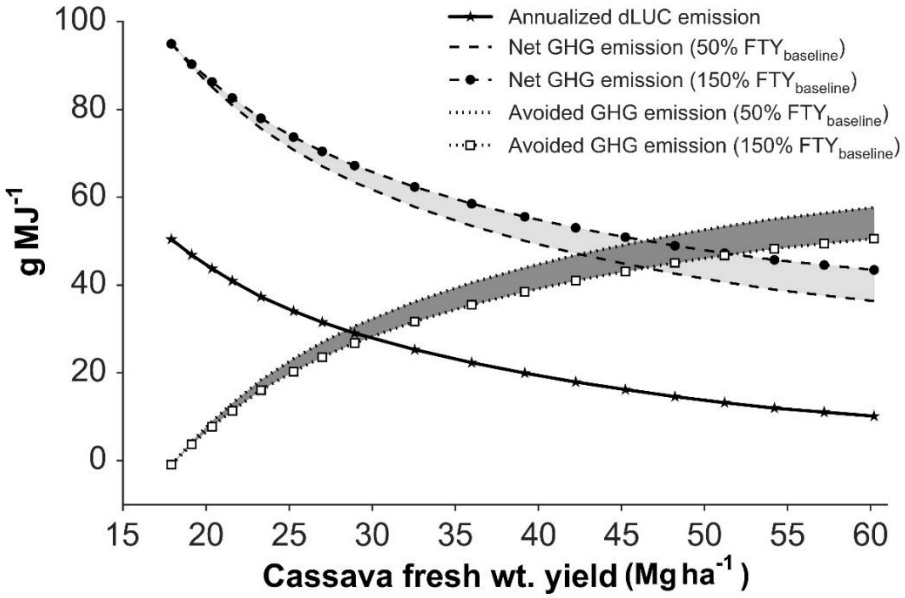


Fig. 2.7. Greenhouse gas emissions of cassava-based ethanol in Vietnam under different cassava yield and fertilization scenarios. The filled areas are the difference between 50% $FTY_{baseline}$ and 150% $FTY_{baseline}$.

2.4.4 Carbon payback period for cassava-based ethanol in Vietnam

Our results showed that payback period declined significantly as cassava yield increased (Fig. 2.8). This was because of lower carbon debts and greater fossil fuel offset at higher yields. The vertical asymptote of payback period is the yield level where the net GHG emission of ethanol is equal to that of CG (i.e., payback period reaches infinity). The vertical asymptote of payback period for cassava-based ethanol in Vietnam was 18.3 Mg ha^{-1}. There was no payback period below this yield level because replacing CG with ethanol increases net GHG emissions. The GHG mitigation benefit of cassava-based ethanol starts to be realized above this yield level. Payback period was equal to the study time frame (25 years) at the cassava yield of 31 Mg ha^{-1}

and 33 Mg ha^{-1} for 50% $FTY_{baseline}$ and 150% $FTY_{baseline}$, respectively. The most optimistic scenario with the highest considered cassava yield (60 Mg ha^{-1}) and 50% $FTY_{baseline}$ resulted in the lowest payback period of 5 years. The effect of fertilization on payback period was substantial at lower cassava yields and decreased as yield increased. The difference in payback period between 50% $FTY_{baseline}$ and 150% $FTY_{baseline}$ ranged from 95 years to 1 year as yield varied from 19 to 60 Mg ha^{-1}.

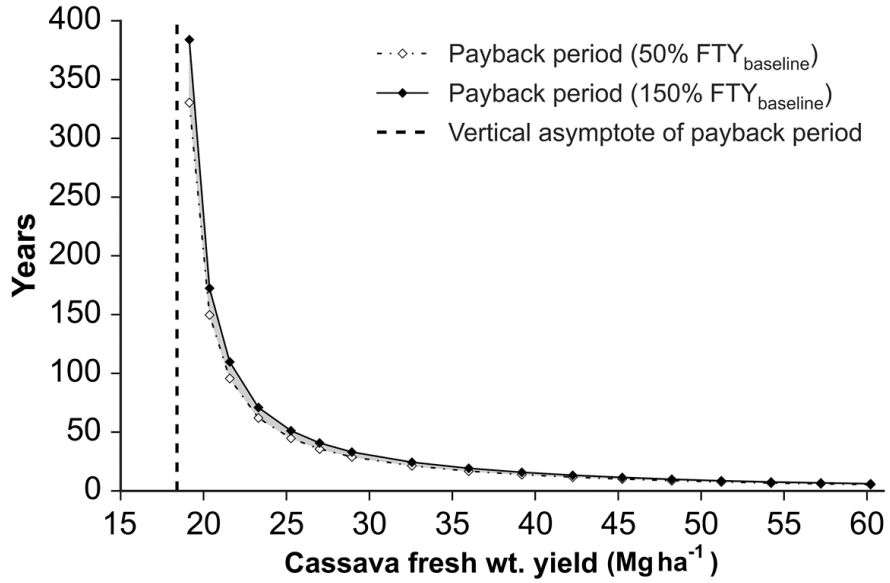


Fig. 2.8. Carbon payback period of cassava-based ethanol in Vietnam under different cassava yield and fertilization scenarios.

2.5 DISCUSSION

The carbon debt, net GHG emission, and payback period of cassava-based ethanol in Vietnam computed based on CENTURY's modeled dLUC were compared with those calculated based on the results from Le et al., (2013) at the same yield level. Using Tier 1 IPCC's guideline, Le et al., (2013) reported an annualized dLUC CO_2 emission of 3.89 Mg ha^{-1} corresponding to a carbon debt of 77.84 Mg ha^{-1} over a 20-year period. To compare the results between the two approaches, we recalculated the net GHG emission and payback period using the method presented in section

2.2.1 and 2.2.2 with the annualized dLUC emission obtained from Le et al., (2013) at the cassava yield of 33 Mg ha^{-1}. The comparison revealed that using Tier 1 IPCC's guidelines to estimate the annualized dLUC emissions for cassava-based ethanol resulted in a reduction of 10.5 Mg ha^{-1} in mean carbon debt, but an increase of 2.56 g MJ^{-1} in the net GHG emission due to higher annualized dLUC emission caused by shorter time frame (20 years). This combination yielded a slight reduction in the payback period of 1 year (assuming 150% $FTY_{baseline}$) and 1.1 year (assuming 50% $FTY_{baseline}$).

Cassava-based ethanol has been internalized in Vietnam's energy policy for partial replacement of conventional fossil fuels, and increased environmental protection (Decision no. 177/2007/QD-TTg of Government of Vietnam, 2007). However, as shown in our results, the GHG mitigation effects of cassava-based ethanol depended strongly on the average cassava feedstock yield. Higher yields significantly reduced carbon debts and payback periods. The mean carbon debt was paid off within the study time frame of 25 years only at the cassava yields above 33 Mg ha^{-1} (Fig. 2.8). This yield level is only attained on farm trials using practices that currently have low adoption rates (<10%) (K. Hoang et al., 2010). To maximize the environmental benefit of cassava-based ethanol, research and policies should be implemented to improve the average cassava yield used for ethanol production.

Additional factors that affect the estimation for carbon debts and LUC emissions of cassava-based ethanol include soil erosion and iLUC. Although they were beyond the scope of our study, it is worthwhile to discuss their potential impacts on the environmental performance of cassava-based ethanol. Studies have shown that growing cassava can severely erode the soils, especially those planted on hilly terrain (Howeler, 2000). The average soil erosion rate under cassava cultivation in the region is 50 Mg ha^{-1} $year^{-1}$ (Howeler, 2000; Hill and Peart, 1998; Newson,

1997). While erosion could potentially impact crop productivity, and thus lower C inputs from residues, its impact as a net source or sink of C with respect to the atmosphere at large scales (landscape, region, globe) is a topic of active debate (Harden et al., 2008; Berhe et al., 2007). Soil erosion results in the removal of SOC from one site and its subsequent deposition at other sites. The mineralization during transport and landscape deposition of eroded SOC were previously thought to be a source of CO_2 to the atmosphere (Lal et al., 2004). However, the burial of eroded SOC in some depositional environments like lakes, rivers, or oceans may also result in a long-term C sequestration due to inhibited C decomposition. The fundamental controls on the magnitude of the erosion-induced sink or source are the extremely complex and combined effects of erosion and deposition, changes in the rates of SOC decomposition and C inputs in eroded or deposited soil (Oost et al., 2007). Thus, erosion impacts of cassava production are highly site-specific and require a more detailed analysis to account for the net impact on C stock change.

Indirect LUC could occur when existing agricultural land is converted to cassava production for ethanol, causing new land conversions elsewhere to displace the food and feed production. These conversions could produce large GHG emissions if carbon-rich ecosystem such as forests are converted. In the case of cassava production for biofuel in Vietnam, the conversion of annuals and perennials to cassava could potentially induce partial conversion of other land uses such as forest and grassland to annuals and perennials, intensification of existing crop production systems, increase in import, and demand adjustments by price regulation in agricultural commodity market to make up for the reduced supply of annual and perennial products. The proportion of each factor and its associated emissions contributing to iLUC effects are highly uncertain. Our simulations showed that land conversion from forest to annuals resulted in an

average SOC loss of 11 Mg ha^{-1} and an average Cveg loss of 58 Mg ha^{-1}, yet conversion from forest to perennials resulted in no change in SOC and a Cveg loss of 44 Mg ha^{-1}. In the worst-case scenario, if 100% of the conversion from annuals and perennials to cassava was replaced by forestland conversion, iLUC could add a significant amount of 116 Mg ha^{-1} to the carbon debt, and shift the asymptote for payback period to between 31 Mg ha^{-1} (assuming 50% $FTY_{baseline}$) and 35 Mg ha^{-1} (assuming 150% $FTY_{baseline}$) (Fig. 2.9).

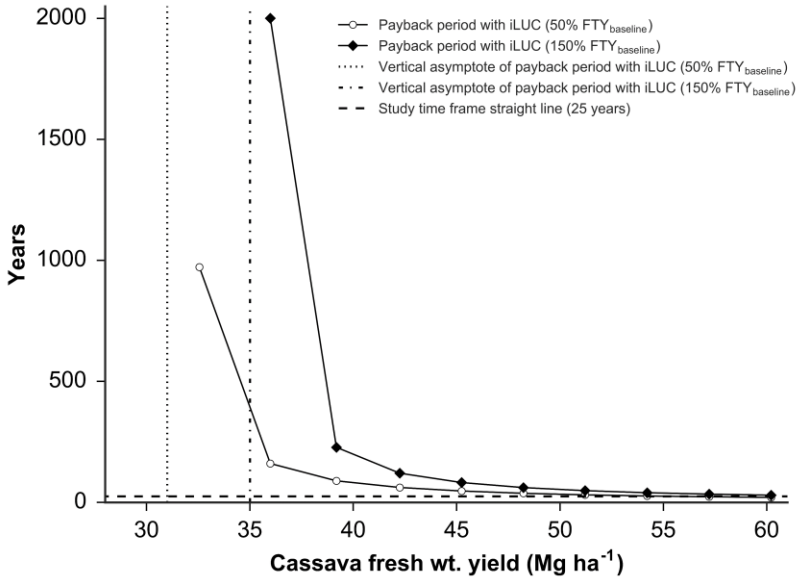


Fig. 2.9. Carbon payback period of cassava-based ethanol in Vietnam with iLUC (assuming 100% of the conversion from annuals and perennials to cassava was replaced by forestland conversion).

2.6 CONCLUSION

We investigated carbon debts, GHG emissions, and carbon payback periods for cassava-based ethanol in Vietnam under different yield and fertilization scenarios for a time frame of 25 years. The employment of a biogeochemical model enabled us to do a more spatially-explicit qualification of C stock changes due to dLUC to cassava production that accounted for regional differences in climate, soil, and baseline management conditions. The fresh weight yield range used in our analysis (18 – 60 Mg ha^{-1}) is the realistic yield levels reported in literature. Our

results showed that the mean carbon debt due to dLUC to cassava production ranged from 66 to 97 Mg ha^{-1}. The net GHG emission of cassava-based ethanol with the inclusion of dLUC emissions ranged from 33 to 95 g MJ^{-1}, varying as a function of cassava yield (from 18 to 60 Mg ha^{-1}) and fertilizer inputs. The cassava yield level above which the GHG mitigation benefit of cassava-based ethanol is realized was 18.3 Mg ha^{-1}. To repay the carbon debt from direct land use change within 25 years, the average cassava yield used as feedstock for ethanol production must be above 33 Mg ha^{-1}. It is recommended that research and policies be implemented to increase average feedstock yield to improve the environmental performance of cassava-based ethanol.

CHAPTER 3. HIGH-RESOLUTION TRADE-OFF ANALYSIS AND OPTIMIZATION OF ECOSYSTEM SERVICES AND DISSERVICES IN AGRICULTURAL LANDSCAPES[2]

3.1 SUMMARY

Agricultural land management often involves trade-offs between ecosystem services (ES) and disservices (EDS). Balancing these trade-offs to achieve low-impact production of agricultural commodities requires rigorous approaches for quantifying and optimizing ES and EDS, reconciling biophysical constraints and different management objectives. In this study, we demonstrate a high-resolution spatially-explicit analysis of ES and EDS trade-offs for irrigated corn production systems in the South Platte River Basin, Colorado, USA, as a case study. The analysis integrated a biogeochemical model (DayCent) with optimization algorithms to assess trade-offs between multiple ES and EDS indicators, including net primary production, soil organic carbon, water use, nitrogen leaching, and greenhouse gas emissions. Our results show a large fraction of total potential system productivity (up to 21 Mg ha^{-1} $year^{-1}$) can be realized at minimal ecosystem impacts through careful land management decisions. Our analysis also explores how different land management objectives imply different landscape configurations.

3.2 INTRODUCTION

Although the provision of food and fiber has always been the primary objective of agricultural production, agricultural ecosystems can also be managed for other benefits such as climate mitigation, water quality improvement, and biodiversity conservation. The collection of these

[2] This chapter was published in Environmental Modelling & Software Journal, Elsevier.
Nguyen, T.H., Cook, M., Field, J.L., Khuc, Q.V., Paustian, K., 2018. High-resolution trade-off analysis and optimization of ecosystem services and disservices in agricultural landscapes. Environ. Model. Softw. 107, 105–118. https://doi.org/10.1016/j.envsoft.2018.06.006

benefits are referred as “ecosystem services” (ES) (MA, 2005a). Agricultural ecosystems are affected by a variety of human activities involving land use decisions and specific land management practices. The negative impacts of humans on these ecosystems can directly reduce productivity (e.g. reduced soil fertility and loss of habitat for biodiversity conservation) or impose detrimental off-site effects on other ecosystems and human society such as ground water pollution from nutrient leaching, pesticide poisoning of non-target species, and increased greenhouse gas (GHG) emissions. These negative impacts are known as “ecosystem disservices” (EDS) (Zhang et al., 2007). Due to resource limitations (e.g., land, water, nutrients, technology, and labor) in agricultural production, there are often trade-offs between and among ecosystem services/disservices (hereafter referred as ES-EDS trade-offs). For instance, increasing food and fiber production tends to come with higher GHG emissions and nitrogen leaching (Power, 2010). The questions of interest are “what is the magnitude of the trade-off, i.e., how much change in one ES or EDS would lead to change in other ES and/or EDS?” and “how do we optimize ES-EDS trade-offs for the most efficient agricultural production?”. Answering these questions is a context-dependent exercise that necessitates quantifying agricultural ES and EDS and their spatial-temporal dynamics at different scales and levels (de Groot et al., 2010), and integrating those results with optimization procedures for trade-off analyses.

Earlier studies on ES assessments used land cover types as indicators to infer potential values of ES and EDS in different landscapes (Maynard et al., 2010; Kershner et al., 2011; Schneiders et al., 2012; Bagstad et al., 2013). Subsequently, ES-EDS trade-offs were examined using scenario analysis, which facilitates the investigation of drivers of change and the impacts of certain land use and land management options on ES and EDS indicators under specifically defined scenarios (Volk, 2013). The use of land cover types as ES indicators, while useful for the rapid and cost-

effective analysis of aggregated watersheds and natural landscapes, falls short of describing the underlying ecosystem processes, the temporal and spatial dynamics of the ES and EDS provision, and the changes in ES and EDS as a function of varying external factors such as management decisions, policy, and market price (Villa et al., 2014). Therefore, it does not allow the finer measurements of the ES-EDS trade-off nor support the integration of optimization procedures.

Quantifying ES and EDS is a difficult task that requires thorough understanding of fundamental physical and biological processes within the ecosystem. In practice, such understanding is often challenged by limited, incomplete, and/or costly field measurements. Furthermore, ecosystem responses to external disturbances are highly heterogeneous due to variability in soils, climate, land use history, and other site-specific attributes, making interpolation between existing field trials with statistical models very difficult. More recent studies overcome these issues by using process-based models coupled with geographic information system (GIS) to quantify ES and EDS associated with variations in crop rotation schemes and management practices, in a spatially-explicit manner (Lautenbach et al., 2013; Kragt and Robertson, 2014; Balbi et al., 2015). A process-based model is the mathematical representation of the underlying processes that characterize the functioning of well-delimited biological systems (Buck-Sorlin, 2013). Models such as DAYCENT (Del Grosso et al., 2000), DNDC (Li et al., 1992), and APSIM (Keating et al., 2003) can capture the finer-scale influence of site-specific weather conditions, soil properties, crop types, cropping practices, and land use history that determine the provision of ES and EDS (Nguyen et al., 2017). Although this approach requires calibration, validation, setup, and implementation of complex dynamic models, as well as in-field expertise to carry out the analysis, it provides more insights into the fundamental physical and biological processes that

determine ES and EDS, allowing a continuous feedback between decision-making and the corresponding changes in different ES and EDS at multiple spatial scales.

The use of process-based modeling approaches for ES and EDS quantification, coupled with optimization procedures (simulation-optimization) for trade-off analyses, can make decision-making in natural resource management more effective, efficient, and defensible (Volk, 2013). Process-based models can be employed for exploratory quantification of ES and EDS to investigate the potential production of an agricultural landscape based on a set of well-defined scenarios. The results can then be optimized with mathematical algorithms like the non-dominated sorting genetic algorithm (NSGA-II) (Lautenbach et al., 2013), simulated annealing (Chan et al., 2006), or goal programming (Aldea et al., 2014). ES-EDS trade-offs are often presented via simulated Pareto frontiers (also called 'production possibility frontiers'), which define the set of solutions that maximize ES while minimizing EDS given finite available resources (i.e. biophysical constraints). Decision makers can then decide on the optimal solutions on the Pareto frontier that meet their specific management objectives.

Although this simulation-optimization approach has been adopted in previous ecosystem service studies, we found that most studies focused on the aggregation of ES-EDS trade-offs at regional, national, or sub-global levels to inform strategic policy-making (e.g., Lautenbach et al., 2013; Lester et al., 2013; Kragt and Robertson, 2014; Balbi et al., 2015; Ewing and Runck, 2015; King et al., 2015; Kennedy et al., 2016). Other spatially-explicit studies zeroed in on the tactical optimizations of biofuel supply chain and/or biofuel crop production systems at coarse resolutions such as land resource unit level (5 square-mile hexagons) (Yu et al., 2014), hydrological response units (HRU) (> 204 ha) (Lautenbach et al., 2013), county-level (P.W. Tittmann et al., 2010), watershed and sub-basin level (Parish et al., 2012). Only few studies

could quantify ES-EDS at finer spatial resolutions, such as field-level (Zhang et al., 2010). Besides, these studies often considered a limited number (between 1 - 4) of ES-EDS objectives to ease the optimization and visualization. Thus, there is a lack of higher-dimensional trade-off analyses at more local scales (i.e., finer spatial, temporal, and management resolutions) to inform the direct decision-makers of agricultural ecosystems (e.g., farmers, ranchers, and forest landowners) on how they could manage their farms (the principle decision unit in the agricultural landscape) for optimal ES and EDS provision. As implied by Zhang et al., (2007) and Power, (2010), when it comes to ES-EDS trade-offs in spatially heterogeneous ecosystems like agriculture, the devil is really in the details.

Our study aims at demonstrating the linkages among different components, including the field-scale and detailed quantification of management-induced ES changes, the landowner's management preferences, and the multiobjective optimizations for rigorous trade-off analyses of multiple ES and EDS in agricultural ecosystems. We present this research with a case study on high-resolution quantification of ES-EDS tradeoffs and optimization of fertilizer and irrigation decisions for irrigated corn production in the South Platte River Basin, Colorado, United States. A biogeochemical model (DayCent) was employed for exploratory quantification of five ES and EDS, including biomass production, soil carbon storage, water provision, water quality, and climate regulation, at the field scale (1 ha). The model simulations considered the effects of site-specific factors such as soil properties, weather data, and historical (dated back to the 1880s) land use and current management practices on the ES and EDS quantification. The simulated outputs of ED and EDS were linked with a non-dominated sorting algorithm to construct Pareto frontiers quantifying the best possible basin-scale outcomes. We then used linear programming

to identify the optimum fertilizer and irrigation rates for each land unit in the basin based on different predefined land management objectives.

3.3 CASE STUDY AND METHOD

3.3.1 Study site

Our study focused on the irrigated corn growing area in the South Platte River Basin located within north-central Colorado, USA (Fig. 3.1). The region is among the most productive irrigated agricultural areas in the state with a majority of fine-loamy soils, average growing season evapotranspiration of irritated corn crop of 65 cm reported by the Colorado Agricultural Meteorological Network (http://www.coagmet.com/), average growing season precipitation of 31 cm, and average minimum and maximum growing season temperature of 11.4 and 24.7 ^{o}C, respectively (Mesinger et al., 2006). The total area of the study region is 116,959 ha, comprising 33% of all irrigated area in the basin (CDSS, 2010).

The South Platte River Basin is facing many issues such as water pollution from excessive fertilizer run-off and large-scale diversion of limited water resources away from irrigated agriculture (dry-up) in order to meet future municipal and industrial (M&I) needs (Water Conservation Board, 2010). The South Platte agricultural area ranked first in nitrate contamination and second in phosphorus contamination among the 20 major rivers in the US (Strange et al., 1999). This is due to the basin's low capacity of contaminant dilution, which is 10 times below national average level (Mueller et al., 1995) and the lack of riparian vegetation to filter irrigation return flows and feedlot run-off (Loomis et al., 2000). According to the Colorado's Statewide Water Supply Initiative 2010 report (Water Conservation Board, 2010), under medium economic development assumptions, the population of the South Platte Basin is projected to grow from 3.5 million people in 2008 to 6.0 million people by 2050. This would

result in a 136-million-cubic-meter gap in water supply for M&I uses and will likely trigger a permanent dry-up of 73,000 to 108,000 ha of irrigated farmland in the basin (Water Conservation Board, 2010). The large-scale dry-up of irrigated agriculture land will likely cause significant negative economic, social, and environmental impacts to the basin and to the whole state. These issues necessitate integrated assessments and better landscape designs to improve the efficiency of resource allocation and minimize detrimental impacts on the environment while meeting future demands.

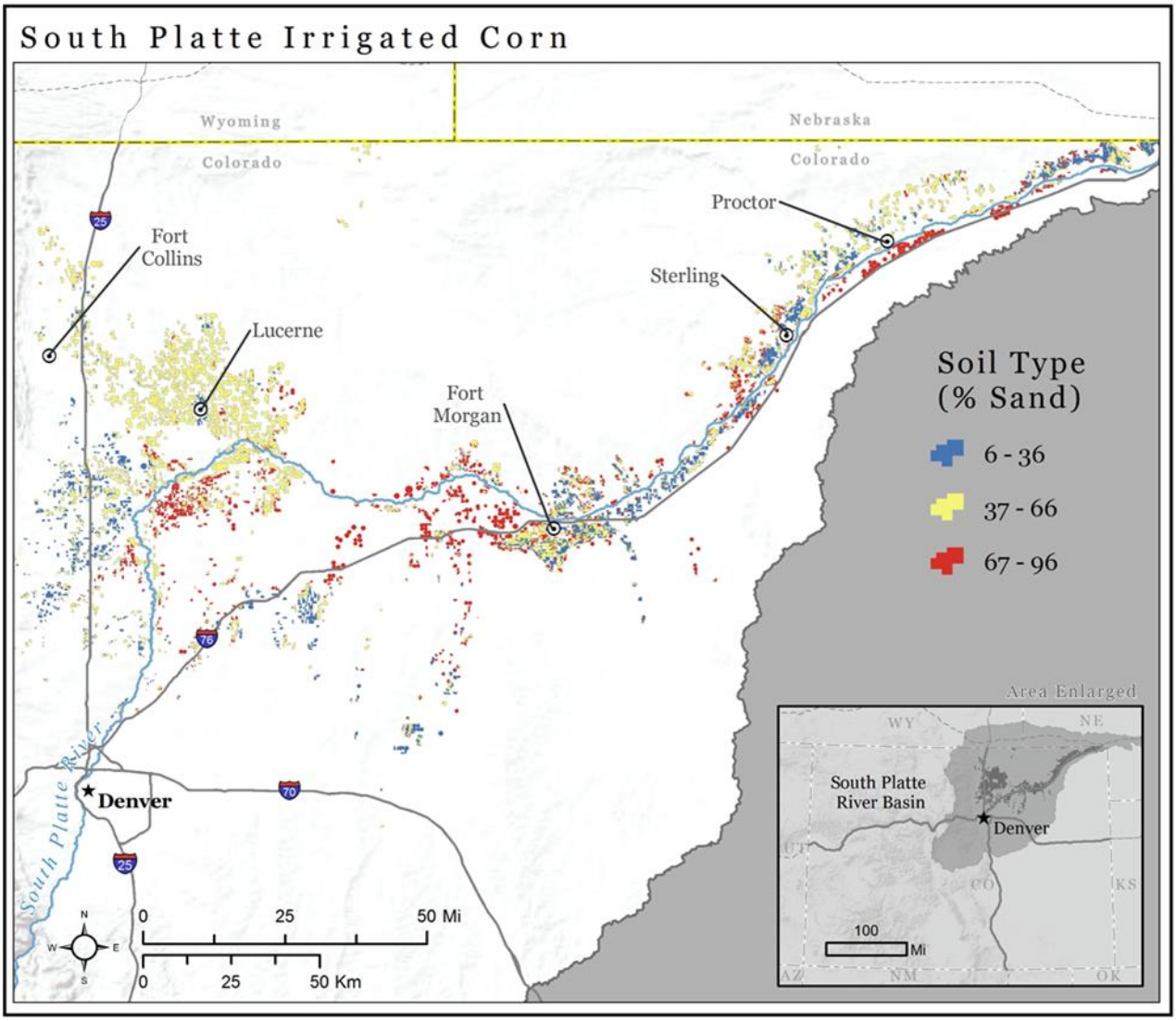


Fig. 3.1. Irrigated corn land units with their sand percentage within the South Platte River Basin.

The small window on the lower, right corner shows the location of South Platte River Basin in Colorado, USA.

3.3.2 DayCent model simulation

3.3.2.1 DayCent model

DayCent (Parton et al., 1998) is a daily time step, process-based model that represents biogeochemical flows of carbon, nitrogen, and water. It simulates the dynamics of many ecosystem processes including changes in soil organic matter, soil water, nutrient cycling, and trace gas fluxes (Del Grosso et al., 2002).

Plant growth (net primary productivity, or NPP) is simulated as a function of nutrient availability, soil water and temperature, shading, vegetation type, and plant phenology (Metherell et al., 1993). The resulting carbon is then partitioned to different plant components (e.g., roots vs. shoots) based on plant type, phenology, soil water content and nutrient availability.

Soil organic matter (SOM) dynamics are simulated for surface litter pools and the top soil layer (0-20 cm). SOM is divided into two litter pools (structural and metabolic) and three mineral-associated organic matter pools (active, slow, and passive) with different potential decomposition rates. Microbially-mediated decomposition of litter and SOM are represented as functions of substrate availability, substrate quality (lignin content, C/N ratio), soil texture, temperature, water availability, and tillage intensity.

DayCent's water flow submodel simulates the daily flow of water through the plant canopy, litter, and soil layers. Soil water content is simulated for each soil layer throughout the soil profile depth (Parton et al., 1998). DayCent also simulates water loss through sublimation of snowpack, interception of rainfall by surface litter and the plant canopy, water runoff from

infiltration excess, water leaching from the bottom of the soil profile or deep storage, soil evaporation, and plant transpiration.

The soil nitrogen submodel includes processes such as N addition, atmospheric deposition, volatilization, leaching, plant uptake, and N mineralization and immobilization. DayCent calculates N leaching as a function of soil nitrate (inorganic N leaching) and active soil SOM pool decomposition (organic N leaching), soil texture, and the amount of water moving through the soil profile. The trace gas submodel of DayCent simulates soil NO_x and N_2O gas emissions from nitrification and denitrification as well as N_2 emissions from denitrification. Methane (CH_4) efflux from soil are also modeled by DayCent.

The model's primary inputs are daily maximum and minimum air temperature and precipitation, soil texture for each horizon in the soil profile, soil (rooting) depth, land cover/use data (e.g., vegetation types, land use history), and management practices (e.g., irrigation, tillage, fertilization). The nominal spatial scale for DayCent simulations is one square meter, though model results are typically scaled up across larger areas (ha-sized) for which model inputs are assumed as uniform (Del Grosso et al., 2008a). The DayCent model has been used previously for the simulation of agricultural system productivity, soil carbon, and trace gas emissions accounting in a wide variety of US agricultural systems. For example, the model has been calibrated and validated across all principal US cropping systems in the process of producing the annual Inventory of U.S. Greenhouse Gas Emissions and Sinks (US EPA, 2014). For this study region in particular, Zhang, (2016) reported a coefficient of determination (R^2) of 0.89 and a root mean squared error (RMSE) of 1343 kg ha^{-1} when comparing the 15-year average DayCent's simulated corn yield with those reported by the National Agricultural Statistics Service (NASS).

3.3.2.2 Spatial input databases

An overview of our methodological approach is presented in Fig. 3.2. Several spatial databases were used to create input data for DayCent. Soil and climate data were obtained from the Natural Resource Conservation Service Soil Survey Geographic database (SSURGO, NRCS-USDA, 2014), and the North American Regional Reanalysis database (NARR, Mesinger et al., 2006), respectively. For each climate grid cell, a table of daily precipitation, maximum temperature and minimum temperature is constructed in a format described in Easter et al., (2005). Soil texture, rock fraction, and pH for different soil profile layers of the dominant soil component for each map unit were taken directly from the SSURGO database. The bulk density, field capacity, wilting point, and saturated hydraulic conductivity were computed using the Saxton equations (Saxton et al., 1986). Irrigated land cover data of the South Platte River Basin were obtained from the GIS database of Colorado's Decision Support Systems (CDSS, 2010). These three GIS layers (soil, climate, and land use) were then intersected and any small slivers (<1 ha) were merged into their longest shared-edge neighbors. The intersection created 16,651 polygons of various sizes for irrigated corn in the South Platte River Basin, including 1,263 unique combinations of soil type and weather data (DayCent input ‘strata’). Each unique stratum was simulated with DayCent, and the results were then linked back to their associated landscape polygons for further calculation and spatial visualization. To facilitate similar landscape studies, we created an ArcGIS add-in toolbox to automate the geospatial data processing for DayCent landscape simulation.

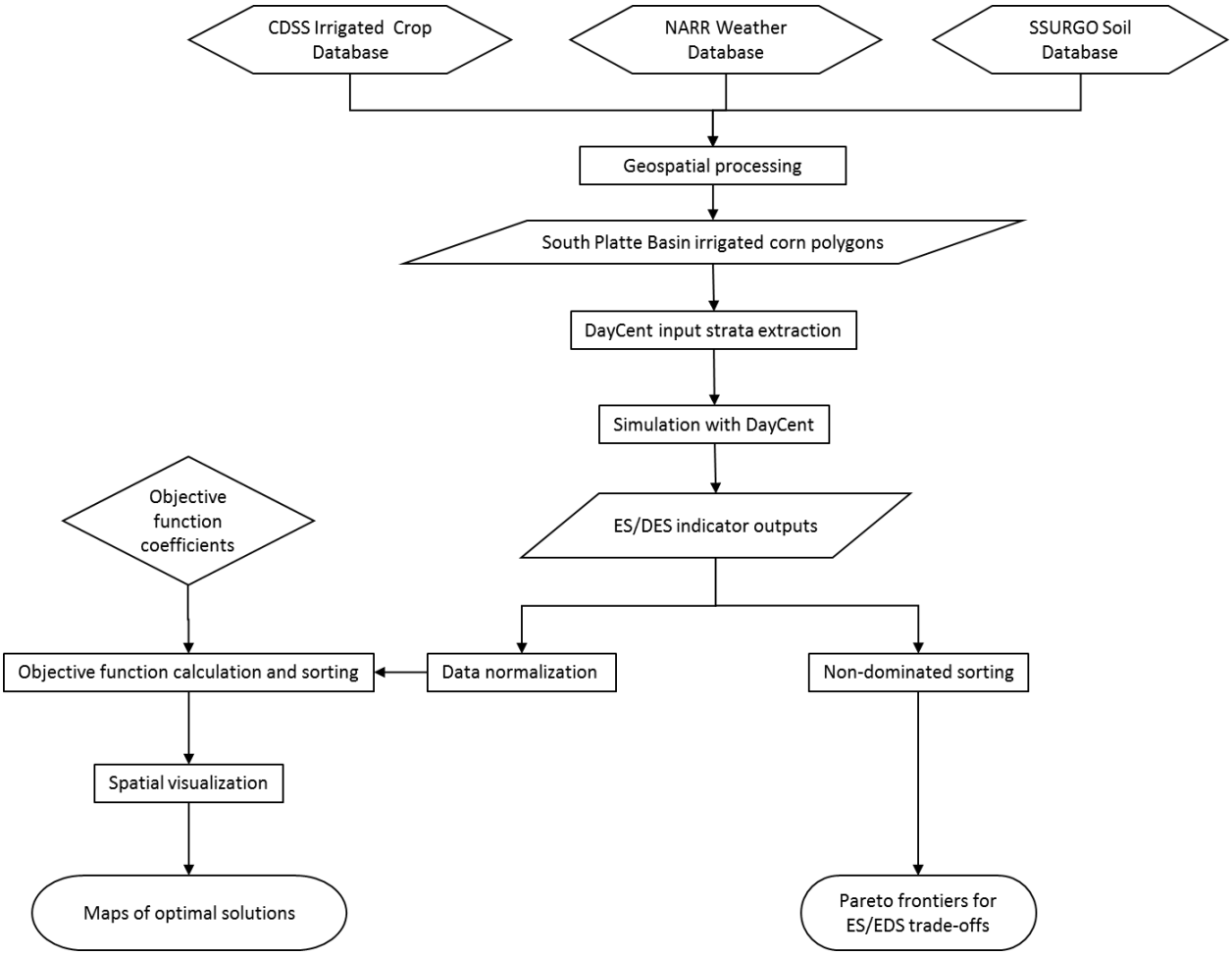


Fig. 3.2. Overview of our study approach.

3.3.2.3 Simulation of baseline conditions

Before conducting experimental simulations for each stratum we initialized DayCent using pre-settlement and historical agricultural land use assumptions at the Major Land Resource Area (MLRA) scale as described in Ogle et al., (2010). This is the same procedure used in the Inventory of U.S. Greenhouse Gas Emissions and Sinks (US EPA, 2015) and the COMET-Farm^{TM} system (http://cometfarm.nrel.colostate.edu/). The model initialization included two stages: an initial spin-up followed by a simulation of historical land use. The spin-up simulation defines steady-state (i.e., equilibrium) soil carbon stocks, assuming little or no disturbance from

human activities prior to the conversion to agriculture. Grassland was chosen as the native vegetation and was simulated with DayCent for 5000 years. Subsequently, the model run was extended from this equilibrium condition with historical land use assumptions for each MLRA. There were five MLRAs within the study region with different areal coverage percentage, including G67B (84.2%), H72 (15.3%), G67A (0.3%), and E49 (0.17%), and E48A (0.03%). Based on data compiled by Ogle et al., (2010), a two-year rotation between bare fallow and winter wheat was assigned to G67A, G67B, E49, and E48A for the period from 1881 to 1979, where plow-out of the native grassland was set to occur in 1880. As for H72, a four-year rotation of corn – corn – spring small grain – winter wheat was assigned from 1881 to 1920 and a three-year rotation of corn – spring small grain – alfalfa was assigned from 1921 to 1979. Typical management practices for each crop over time were compiled from various sources (e.g., USDA-ERS, 1997; USDA-NASS, 2004; Iback and Adams, 1967; Smalley and Engle, 1942).

To complete the model initialization, we extended the historical baseline conditions from 1980 to 2015 with a 'business as usual' scenario (BAU) for irrigated corn production in Colorado. The BAU reflects a system state under current management practice, and is used to evaluate the magnitude and direction of change due to subsequent changes in management practices. Corn in northern Colorado is usually planted between April 28 and May 20, and harvested between October 8 and November 13 (NASS, 2010). The average BAU management practices reported by the Economic Research Service's Agricultural Resource Management Survey Tailored Reports (ERS-ARMS, 2010) for Colorado's irrigated corn crop were 170 kg of nitrogen per hectare, 34 cm of irrigated water per growing season, 75% residue removal, and conventional tillage.

3.3.2.4 Management practice scenarios

To examine the impacts of N fertilization and irrigation on the provision of different agricultural ES and EDS, we ran our forward simulations with N fertilizer and irrigation rates set to 20%, 40%, 60%, 100%, 120%, 140%, and 160% of the BAU level (68, 102, 136, 170, 204, 238, and 272 kg of N per ha, and 13.6, 20.4, 27.2, 34.0, 40.8, 47.6, and 54.4 cm water $year^{-1}$, respectively). The permutation of these fertilizer and irrigation rates yielded 49 management practice scenarios. We simulated these 49 management practice scenarios for each polygon for a 30-year period extending from 2016 to 2046. The percentage of stover removal was kept constant at the baseline level (75%) and only conventional tillage was considered. We did not consider the effects of changes in CO_2 concentration and climate on crop production and water use efficiency. The combination of 1,263 DayCent input strata and 49 management practice scenarios yielded 61,887 DayCent simulations. These simulations were executed in parallel on an 18-node, 216-processor cluster computing system at Colorado State University's Natural Resources Ecology Laboratory.

3.3.2.5 Ecosystem service and disservice indicators

DayCent simulation outputs were used to estimate multiple ES and EDS indicators for each simulation. Soil organic carbon (SOC) and net primary productivity (NPP) were chosen as ES indicators representing ecosystem carbon storage and biomass production services, respectively. Water use, nitrogen (N) leaching, and greenhouse gas emissions (GHG) were chosen as EDS indicators representing water provision, water quality, and climate regulation disservices.

Since SOC dynamics are a long-term effect of management practices, soil properties, and weather conditions, we reported the net change in SOC in Mg of carbon per hectare at the end of the simulation period to capture the cumulative effects of changing management practices. Due

to the large seasonal variations, other indicators were reported as flow variables averaged over the entire analysis, including Mg of carbon dioxide equivalent (CO_2e) per hectare per year for GHG, kg of nitrogen per hectare per year for N leaching, Mg of dry weight biomass per hectare per year for NPP (assuming 43.5 % carbon content (Gesch et al., 2010)), and cm of water per year for water use.

The values of SOC and NPP indicators were taken as direct DayCent outputs while soil GHG emission estimates included modeled outputs from DayCent as well as additional models (e.g., for indirect N_2O emissions) from the U.S. Department of Agriculture entity-scale greenhouse gas inventory guidelines (Eve et al., 2014). Annual net CO_2 emission was assumed to come mainly from soil microbial respiration of SOC and was computed by taking the difference of SOC levels between two consecutive years. Annual N_2O emission comprised of direct and indirect N_2O emissions. The direct N_2O emission was reported by DayCent in term of N_2O efflux, while the indirect N_2O emission was computed from NOx efflux, volatilized NH_3, and nitrogen leaching with the emission factors (EF) of 0.01, 0.01, and 0.075, respectively. Annual N_2O and CH_4 were converted into CO_2 equivalents by using 100-year global warming potential values (GWP_{100}) of 298 and 25, respectively (IPCC, 2006). N leaching was calculated from the amount of inorganic and organic N leached out of the soil profile. Since we did not model the effect of terrain slope and stream water flow, the N leaching values only reflected the potential amounts of nitrate that can enter the ground water. The ecosystem water use over a growing season was calculated as the total amount of water used to sustain all ecosystem processes including evapotranspiration (the water utilized by the crop to produce biomass) and other processes such as sublimation (the vaporized water from snow and ice), interception by the crop canopy, run-off, and deep drainage.

To simplify the indicator comparisons between different management scenarios, we computed the area-weighted average value of each indicator using the following equation:

$$V_i = \sum_{j=1}^{16,651} \frac{a_j v_j}{A}, \qquad \text{Eq. 3.1}$$

where:

V_i = area-weighted average value of an ES or EDS indicator i,

j = the j^{th} polygon within the landscape,

a_j = the area of the j^{th} polygon within the landscape (ha),

v_j = the value of ES or EDS indicator i of the j^{th} polygon within the landscape,

A = the total area of the study region (A = 116,959 ha).

3.3.3 Calculation of Pareto frontiers and optimization of landscape designs

To examine the magnitude of trade-offs between NPP and other ES and EDS indicators in the study region, we used a simple non-dominated sorting algorithm to identify the Pareto frontier for each pair of NPP and ES or EDS indicators (see Appendix A.1). These Pareto frontiers illustrate macro trade-offs between NPP and other indicators assuming that management practice scenarios were uniformly applied to the landscape ('uniform management'). In reality, it is of greater importance to identify and visualize the optimum landscape design under variable management (i.e., the optimum management practice scenario for each polygon within the landscape) corresponding to a specific management objective. To do this, we used a linear optimization algorithm that seeks to maximize the following objective function:

$$\text{maximize } \sum_i^n c_i X_i, \qquad \text{Eq. 3.2}$$

where:

X_i = the normalized value of ES (or EDS) indicator i. $X_i = \frac{x_i - P_{i\,min}}{P_{i\,max} - P_{i\,min}}$,

x_i = the value of ES or EDS indicator i,

$P_{i\,min}$ and $P_{i\,max}$ = the population minimum and maximum of ES or EDS indicator i, respectively.

c_i = the objective function coefficient (a.k.a., weighting factor) corresponding to X_i, and

n = total number of ES and EDS indicators considered in the optimization (n <= 5).

Since all the ES and EDS indicators were expressed in different units and scales, we first normalized the indicators to a uniform scale (0 – 1) and used the normalized values in the objective function for a fair comparison (Hwang et al., 1980). The optimization algorithm was applied to every polygon in the landscape to identify the optimum fertilizer and irrigation rates for each individual polygon.

Seven management objective scenarios representing four different hypothetical management objectives by the landowner were chosen for the optimization (Table 3.1). For each scenario, a set of objective function coefficients was assigned to the ES and EDS indicators. We limited the coefficients to integer numbers that ranged from -2 to 2 in our analysis. The coefficient absolute values indicated the importance of one indicator relative to the others in the objective function (Eq. 3.2) and the sign of the coefficients reflected the indicator's direction of change that we wanted to optimize. For example, in scenario 2a, increasing NPP and decreasing water use (expressed in the normalized unit term) were set to be twice as important as increasing or decreasing other indicators. Except for scenario 1, where we equally considered all ES and EDS indicators (coefficients = ± 1), other management objective scenarios were biomass-centric (i.e., higher positive NPP coefficients) as biomass production is the primary provisioning goal of the

study region. In addition, scenarios 2, 3, and 4 focused on reducing EDS (coefficients = -2), including water use, N leaching, and GHG, respectively. Among these management objective scenarios, the 'a' scenarios put additional constraints on the remaining ES and EDS indicators while the 'b' scenarios excluded them from the optimization (Table 3.1).

We further compared the EDS footprints between different management objectives. The EDS footprint is defined as the amount of EDS indicator produced per unit of NPP. The EDS footprints for a polygon under a specific management objective were calculated using the following equation:

$$F_{i,j,k} = \frac{x_{i,j,k}}{NPP_{j,k}}, \qquad \text{Eq. 3}$$

where:

$F_{i,j,k}$ = footprint of EDS *i* for polygon j under management objective scenario k,

$x_{i,j,k}$ = the value of ES or EDS indicator *i* for polygon j under management objective scenario k,

$NPP_{j,k}$ = the value of NPP indicator for polygon j under management objective scenario k.

The optimization results were mapped for spatial visualization using ArcGIS software (ESRI, 2014). We performed descriptive statistical analysis (e.g., distribution, central tendency, and dispersion) to explore ES and EDS indicators data (see Appendix A.2) and one-way ANOVA analysis to test the effects of fertilizer and irrigation on each indicator for better presentation of the results (see Appendix A.3). The simulation processes, data analyses, and map generation routines were automated using the Python programming language (https://www.python.org/). Our derived model was named the "Agricultural Ecosystem Service Optimization" (Ag-EcoSOpt) for the sake of future studies and development.

Table 3.1. Management objective scenarios and the corresponding objective function coefficients

Management objective	Management objective scenario	Objective function coefficient					Description
		SOC	NPP	Water use	N leaching	GHG	
Equal optimization of all indicators	1	1	1	-1	-1	-1	An equal consideration of all ES and EDS indicators where one normalized unit increased or decreased of a ES is as important as that of other ES or EDS.
Maximizing NPP and minimizing water use	2a	1	2	-2	-1	-1	Maximizing NPP and minimizing water use **with** constraints on other ES or EDS.
	2b	0	2	-2	0	0	Maximizing NPP and minimizing water use **without** constraints on other ES or EDS.
Maximizing NPP and minimizing N leaching	3a	1	2	-1	-2	-1	Maximizing NPP and minimizing N leaching **with** constraints on other ES or EDS.
	3b	0	2	0	-2	0	Maximizing NPP and minimizing N leaching **without** constraints on other ES or EDS.
Maximizing NPP and minimizing GHG emissions	4a	1	2	-1	-1	-2	Maximizing NPP and minimizing GHG emission **with** constraints on other ES or EDS.
	4b	0	2	0	0	-2	Maximizing NPP and minimizing GHG emission **without** constraints on other ES or EDS.

Note: *red color represents objective function coefficient of 0 (i.e. no consideration in the optimization), white color represents objective function coefficient of 1 or -1, green color represents objective function coefficient of 2 or -2.*

3.4 RESULTS

3.4.1 Effects of fertilizer and irrigation on ES and EDS indicators

There were various combined effects of fertilizer and irrigation on SOC, NPP, N leaching and GHG emission (Fig. 3.3). The rate of change of an indicator at a specific irrigation rate is defined as the difference in indicator values between the two fertilizer levels (rise/run). Higher fertilizer rates resulted in higher SOC, NPP, N leaching, and GHG emission. As the fertilizer rate increased, SOC and NPP responded with a decreasing rate of change until a maximum value was reached asymptotically, while N leaching and GHG emission rose linearly or with an increasing rate of change. Irrigation rate was proportional with the steepness of the slopes of the indicator's response lines/curves. The effects of irrigation on increasing SOC, NPP, and N leaching diminished at the irrigation rates above 47.6 cm $year^{-1}$. We also found that reducing irrigation rates amplified net GHG emission within any specific fertilizer level below 238 kg ha^{-1}. This is due to the increasing contribution of net CO_2 flux from soil C stock change in the total GHG emission (see Appendix A.4).

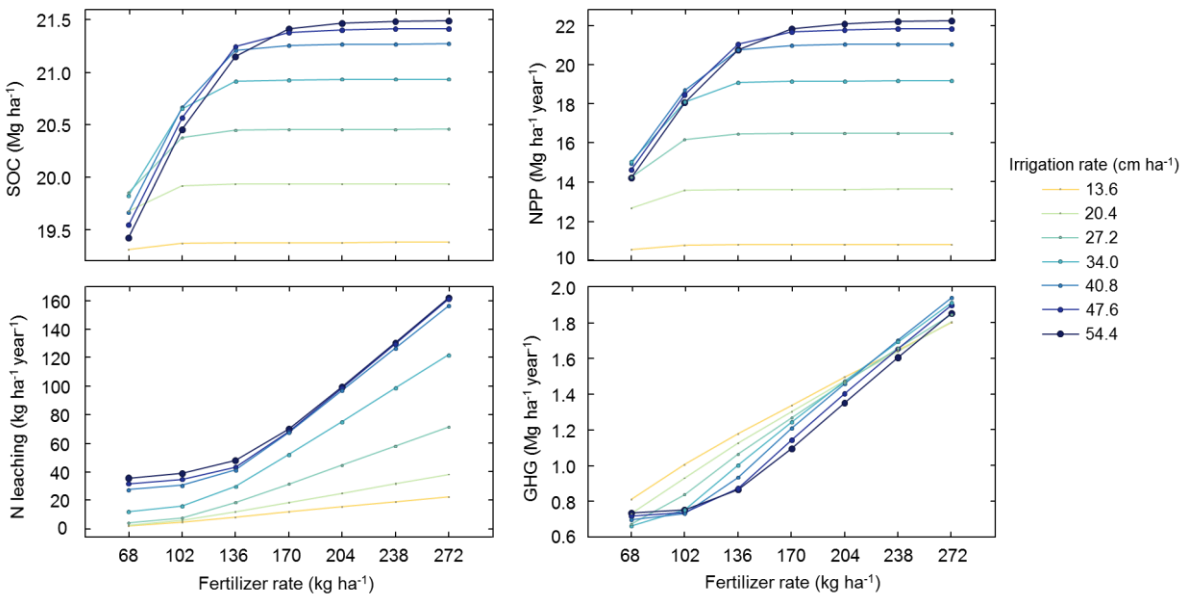


Fig. 3.3. Ecosystem services and disservices indicators at different irrigation and fertilizer rates. SOC is reported for top soil (0 – 20 cm) at the end of the 30-year simulation period. NPP is the 30-year average of total above and below ground dry mass at harvest. GHG is the 30-year average of all soil-related emissions. N leaching is the 30-year average of nitrate that leached out of the soil profile.

Since we did not find statistically significant impacts of fertilizer application rate on total water use as determined by one-way ANOVA ($F(6, 27502) = 0.00$, $p = 1.0$) (see Appendix A.3), we only reported total water use against the irrigation rate averaged across all fertilizer levels (Fig. 3.4). While water use was linearly correlated with irrigation, different water use components responded differently to increasing irrigation rate. In particular, the evapotranspiration increased with a decreasing rate and plateaued at the irrigation rate of 34 cm year^{-1}, whereas the water use from other processes rose slowly at lower irrigation rates and soared as soon as the evapotranspiration leveled off.

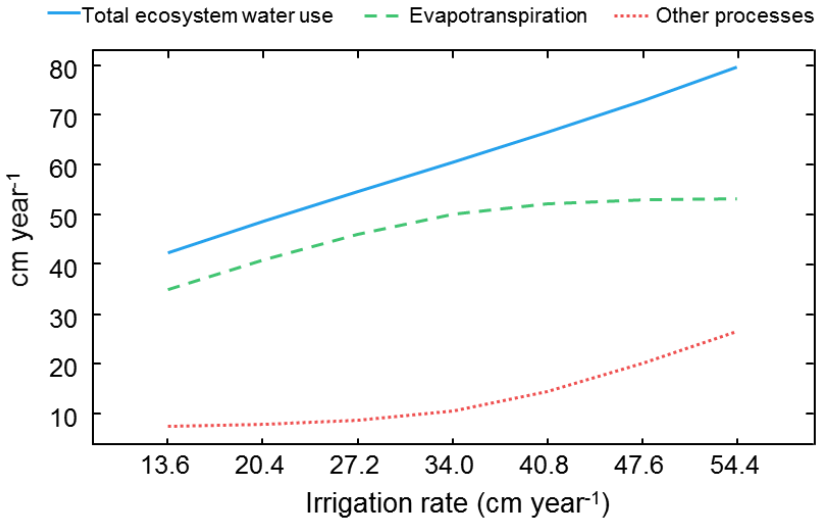


Fig. 3.4. Average growing season water use by the irrigated corn ecosystem at different irrigation rates. The water use was averaged across all fertilizer levels. Other ecosystem processes include sublimation, interception by the crop canopy, run-off, and deep drainage.

3.4.2 Trade-offs between NPP and other indicators

The trade-off between the ES indicators (i.e. NPP and SOC) was a "win-win" situation as increasing NPP led to a linear increase in SOC, whereas those between NPP and EDS indicators were "win-loss" situations since higher NPP also came with higher EDS (Fig. 3.5). The 'uniform management' Pareto frontiers for NPP and other EDS revealed convex patterns. This means that the marginal improvements of NPP at the higher end will come with exponential penalties on environmental performance. While trade-offs between water use and NPP were almost linear, frontiers were increasingly convex for N leaching and GHG emissions. All trade-off frontiers showed a marked increase in EDS to achieve NPP greater than 21 Mg ha^{-1} $year^{-1}$, corresponding to average N fertilizer and irrigation rates above 136 kg ha^{-1} and 40.8 cm $year^{-1}$, respectively. This indicates that a relatively large fraction (>90%) of total potential system productivity can be realized, while incurring only a relatively small fraction of total potential disservices through proper management practice decisions.

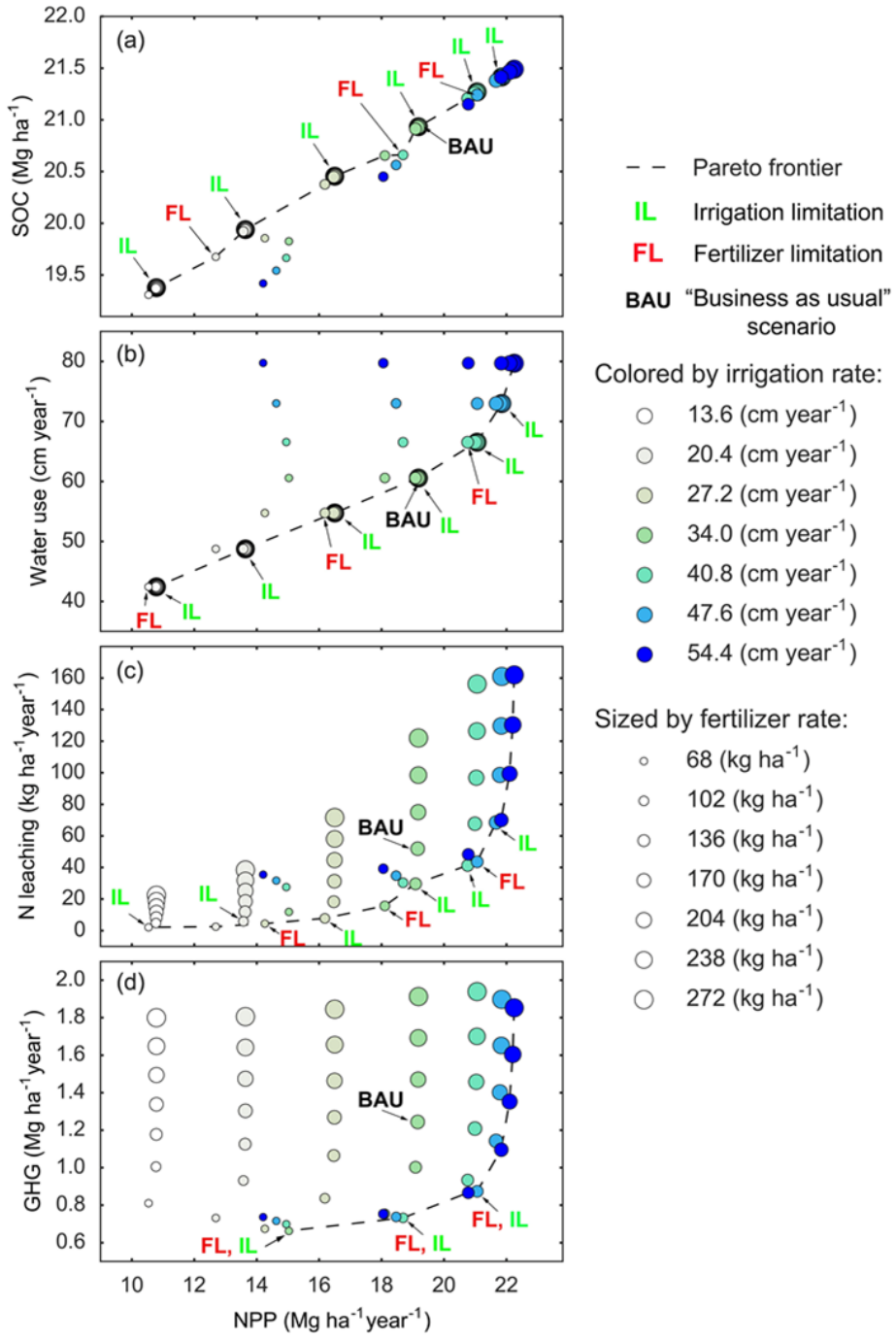


Fig. 3.5. Trade-offs between NPP and other ecosystem service and disservice indicators. The circles represent management practice scenarios. Pareto frontier lines are based on the uniform simulation of management practice scenarios ('uniform management'). Fertilizer limitation (FL) and irrigation limitation (IL) are states on a Pareto frontier at which no further improvement of the optimum NPP can be made without adding more fertilizer and irrigation, respectively.

Our results also showed the influence of fertilizer and irrigation in achieving certain optimal solutions (i.e. scenarios that lie on the Pareto frontiers). The trade-offs between NPP and other indicators were affected by fertilizer limitation (FL) and irrigation limitation (IL) which are states on a Pareto frontier at which no further improvement of the optimum NPP can be made without adding more fertilizer and irrigation, respectively. In the N leaching and GHG optimization cases, optimal solutions with NPP below this level were achieved by modulating irrigation rates while leaving fertilizer application at a minimum level; additional NPP gains beyond this point required increasing N fertilizer application, causing a sharp increase in N leaching and N_2O-driven GHG emissions. Inappropriate applications of fertilizer and irrigation could lead to sub-optimality in ES-EDS trade-offs (i.e., scenarios that do not lie on the Pareto frontiers). This sub-optimality was due to either excessive application of irrigation at the FL states (FL sub-optimality) or excessive application of fertilizer at the IL states (IL sub-optimality). FL sub-optimality reduced NPP as irrigation increased while IL sub-optimality improved NPP with enormous costs on other ES and EDS with increasing fertilizer. Although we found no significant IL sub-optimality in the trade-offs between NPP and SOC (Fig. 5a) and between NPP and water use (Fig. 5b), the addition of fertilizer at the IL states demonstrated a waste in resource investments.

3.4.3 Trade-offs between management objectives

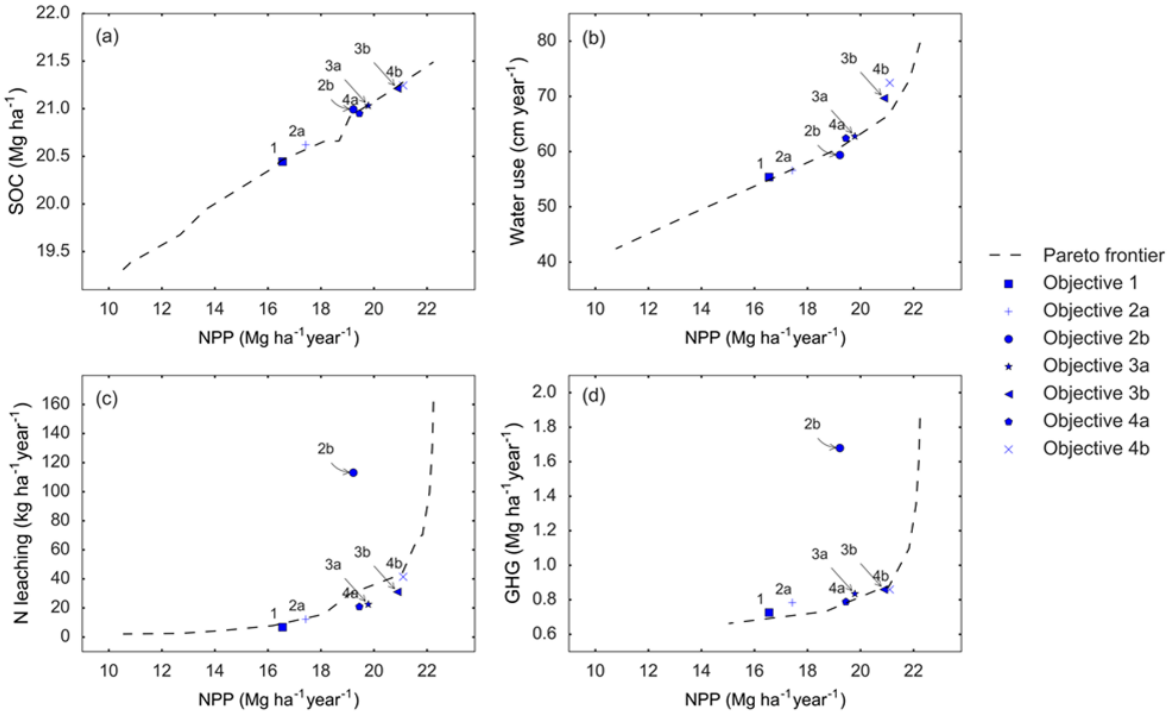


Fig. 3.6. Trade-offs in ecosystem service and disservice indicators between management scenarios. Pareto frontier lines are based on uniform management scenarios in section 3.2. The labeled points show optimization results for different management objective scenarios defined in table 1, where management practices for each land unit in the landscape were optimized.

Fig. 3.6 presents the ES-EDS trade-offs for the optimization of the management practices under seven management objective scenarios defined in Table 3.1. In general, we observed slight modulations of ES and EDS around the 'uniform management' Pareto frontiers, except for the objective 2b (unconstrained optimization of NPP and water use) where high sub-optimality in N leaching and GHG emissions occurred. For this scenario (2b), our model allowed higher fertilizer application (Table 3.2) because there was no constraint placed on either reducing N leaching or GHGs. Our results also showed that allowing variable management during

optimization obviously improved landscape performance in terms of N leaching as compared to the 'uniform management' Pareto frontiers (i.e., objectives 3a, 3b, and 4a in Fig. 3.6c).

Table 3.2. Landscape-averaged fertilizer and irrigation rates for the management objective scenarios

Management practice	**Management objective scenario**						
	1	2a	2b	3a	3b	4a	4b
Fertilizer (kg ha^{-1})	90	101	228	121	135	115	135
Irrigation (cm $year^{-1}$)	28	30	33	37	44	36	47

3.4.4 The spatial distributions of management practices and EDS footprints

The optimum spatial designs of fertilizer and irrigation rates and the spatial distribution of EDS footprints of irrigated corn polygons in the South Platte River Basin were visualized for the management objective scenarios with constraints (1, 2a, 3a, and 4a) (Fig. 3.7 & 3.8). Those for management objective scenarios 2b, 3b, and 4b can be obtained from Appendix A.5. The distributions of optimal fertilizer and irrigation rates were heterogeneous throughout the landscape, with many visible clusters of high and low values. While the optimum rates of irrigation varied widely from 13.6 to 47.6 cm $year^{-1}$, those of fertilizer only varied below the BAU fertilizer level of 170 kg ha^{-1}. The EDS footprint values also varied greatly within the study region depending on the management objective and the EDS indicator. We observed some clusters of high EDS footprint values in the areas north of Denver and around the town of Fort Morgan.

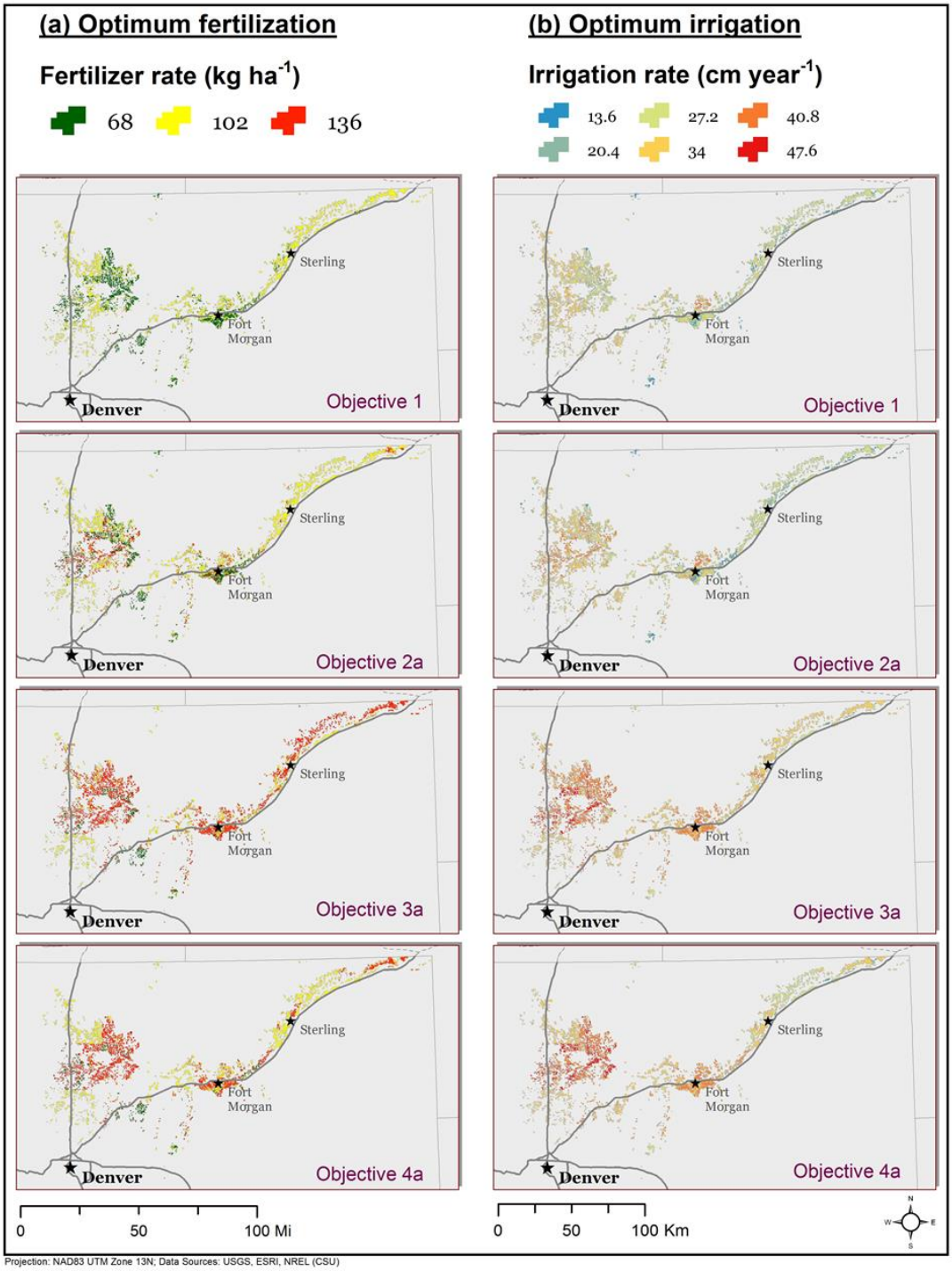


Fig. 3.7. Optimum fertilizer and irrigation rates for irrigated corn area in South Platte River Basin. The results are presented for management scenario 1 and the 'a' scenarios.

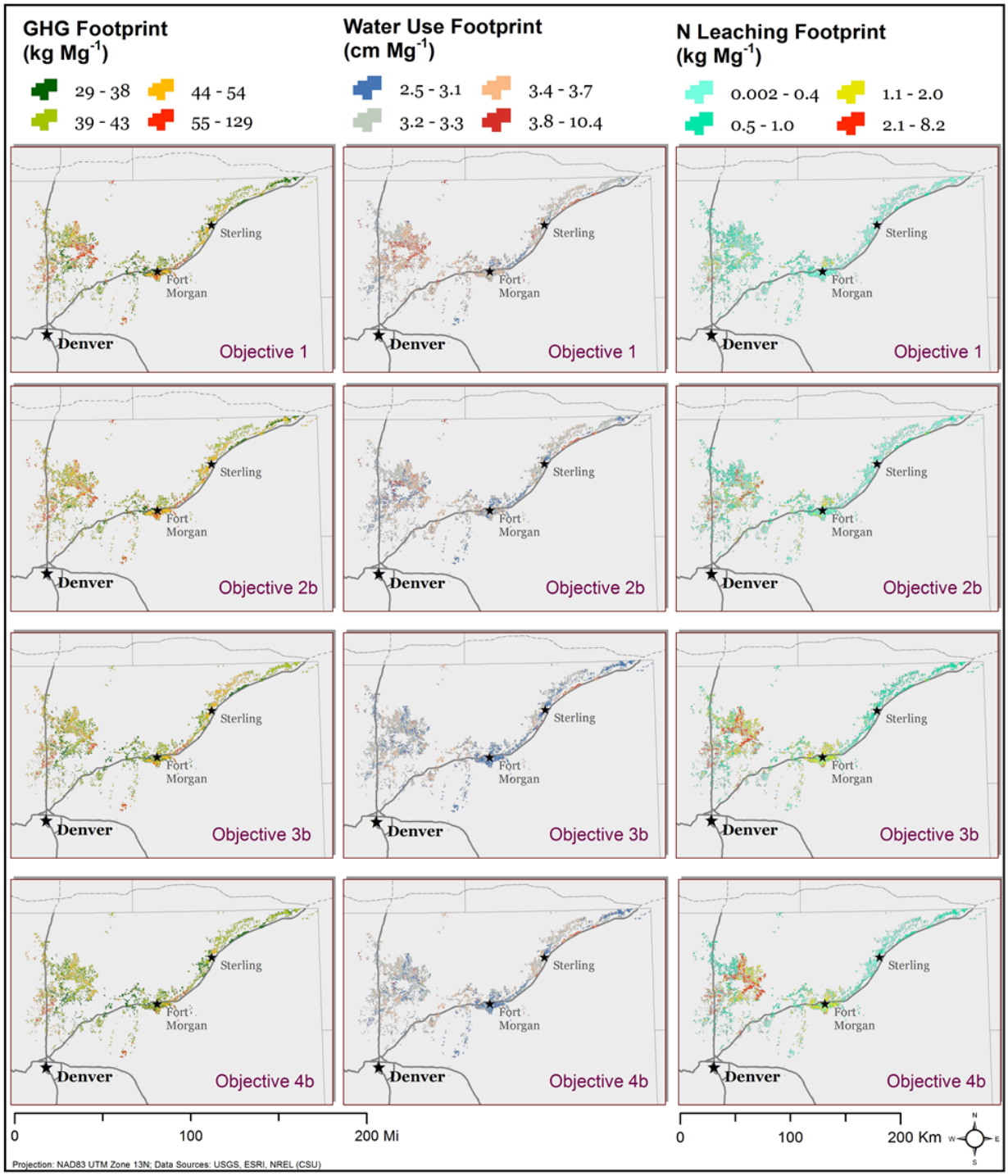


Fig. 3.8. Ecosystem disservice footprints for irrigated corn area in South Platte River Basin. The results are presented for management objective scenario 1 and the 'a' scenarios.

3.5. DISCUSSION

Pressure from the population and industrial expansion in our study region will result in higher demands for food, biomass, and water. The question is "how much of a viable productive agriculture is possible with declining water supply and what are the environmental impacts?". To answer this, we compared the area-weighted landscape sum of ES and EDS indicators and the corresponding management practices under the BAU scenario (Table 3.3) with those from the landscape optimization of the seven management objective scenarios (Fig. 3.9). We found that ecosystem water use can be reduced by 1.9 – 8.6 % relative to the BAU scenario level through lowering irrigation by 3.4 - 17.3%, respectively (objective 1, 2a, and 2b). However, these reductions came with either declines in SOC and NPP (objective 1 and 2a) due to reduced N fertilization or increases in GHGs and N leaching (objective 2b) due to increased N fertilization. This indicated that it was not possible to improve both water use and environmental performance at the current and higher corn production level by simply adjusting fertilizer and irrigation application rates. However, our results implied that these management practices can be modulated to attain higher food and biomass provision at the lowest costs on ecosystems for the study region.

Table 3.3. Landscape sum of ES, EDS, and management practices for the BAU management scenario

Scenario	**Landscape sum of ***						
	SOC (10^6 Mg)	NPP (10^6 Mg)	Water use (10^6 m^3)	GHG (10^3 Mg)	N leaching (10^6 kg)	Fertilizer (10^6 kg)	Irrigation (m^3)
BAU	2.45	2.24	708	146	6.07	19.88	398

Note: * All values are reported on an annual basis.

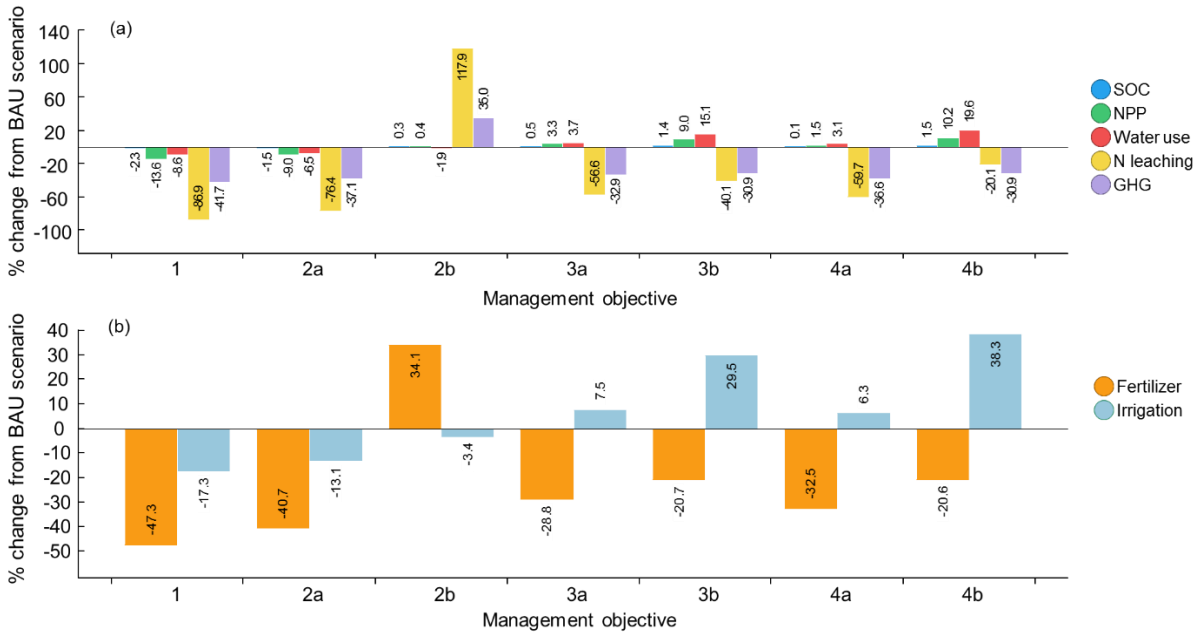


Fig. 3.9. Comparison of the landscape sum of ecosystem service and disservice indicators (a) and management practices (b) between different management objective scenarios and the Business-as-usual (BAU) scenario.

One advantage of the spatially-explicit approaches for ES and EDS quantification is the ability to visualize the fine-scale details and the hidden spatial patterns of landscapes (DeAngelis and Yurek, 2017). Detailed maps such as Fig. 3.7 and Fig. 3.8 have great utility for research and policy applications since they can be used in combination with other spatial layers for more complex quantitative spatial analyses and/or to identify potential hotspots (i.e. significant spatial clusters of high values) of ecosystem threats for specific policy foci.

While our study involved a simple optimization problem in a geographically-limited corn production landscape, the approach and the resulting model would be equally applicable for more complex problems such as those with many constraints, broader scopes, or multiple stakeholders. More constraints can be formulated to reflect biophysical barriers of the landscape, policy regulation of environmental impacts, and regional or local targets/demands for ecosystem

service production. The system boundary can be expanded to include the optimization of multiple land cover scenarios or the supply chain optimization of multiple agricultural products. The potential for the incorporation of economic optimization in this type of analysis allows not only the identification of the economically optimal solutions but also a better representation of the farmer's perspective and influences of the market economy (Lautenbach et al., 2013). The involvement of stakeholders in the analysis would help to identify the most preferred problem solutions (Miettinen et al., 2008). For instance, our optimization results could be used as the Pareto optimal starting points to show to the decision maker. The decision maker would then be expected to express preference for ES-EDS trade-offs. The preference information would be used to adjust the coefficients of the objective function for a new optimization simulation. This process could be iterated until the most desirable solution was attained.

Our analysis might also have practical implications for direct decision makers like farmers. If results of such analysis were made available to farmers, they could adjust their management decisions on their own farms accordingly for the most efficient production. Farmers could also report ES-EDS trade-offs in their farm planning and budgeting. We believe that making the information more available and accessible to stakeholders would be the best way to propagate the ES and EDS concepts, boost their communication, and facilitate the sprouting of ES and EDS market mechanisms (Paustian et al., 2009).

Despite the potential advantages of the approach, we do realize some drawbacks in the methodology including, but not limited to, the expertise challenge and computational burden of the simulation, the discretization of management practice variables, the ambiguousness in the use of the objective function coefficients, and the curse of dimensionality. Using process-based models like DayCent for detailed quantification of ES and EDS requires some level of model

literacy and is computationally expensive for large area simulations. In our case, the analysis necessitated much effort for initial model setup and simulation time took several hours on a cluster of 216 high-performance computer cores. The infrastructure requirements for this type of analysis could impede the adoption of this approach for similar or more complex studies. A possible solution to this disadvantage is the integration of the methodology into infrastructure-ready tools or platforms to enable the flexible replications of ES and EDS analyses. Furthermore, in order to alleviate the computational burden, advanced meta-modeling techniques such as support vector machine and neural network can be applied to create cost-effective simplified surrogates of the detailed process-based models (Wu et al., 2016). Based on these solutions, further efforts are being made to prepare the infrastructure for our model (Ag-EcoSOpt).

The analysis and optimization of ES and EDS based on fixed intervals of the continuous fertilizer and irrigation variables (i.e., discretization) might not guarantee actual optimal solutions since the solutions were forced to some discretized options. However, it significantly reduced the simulation time and supported the quality control of the analysis. Our method also used the normalized values of ES and EDS in the objective function with arbitrary objective function coefficients. Although, this approach was good enough for the scope of our study (finding the optimal landscape designs for pre-defined sets of objective function coefficients based on discretized values of fertilizer and irrigation), the unclear relationships between the coefficients and actual ES and EDS outcomes are a challenge for more complex analyses. For example, one might need to conduct a brute-force search through all possible combinations of objective function coefficients to find the specific set that results in a certain level of ES-EDS trade-offs. Potential solutions for this problem are to involve the decision maker in interactive multi-objective optimization (Wierzbicki et al., 2000), to integrate more sophisticated algorithms

that are capable of coefficients optimization such as Levenberg-Marquardt backpropagation (Levenberg, 1944; Marquardt, 1963), or to utilize different optimization solvers that allow the consideration of multiple objective functions with constraints and bounds configuration such as NSGA-II (Deb, 2011). Last but not least, the employment of this approach for more complex studies could be impeded with increasing dimensions of the analyses such as the addition of management practice types (e.g. tillage and residue harvest), considered levels for each management practice, ES and EDS, management objectives, different land use/cover types, and land use/cover change scenarios. Therefore, a careful consideration of the potential trade-offs between the scope of analysis and modeling capacity is always recommended for this type of study.

3.6 CONCLUSION

We demonstrated a robust analysis of ecosystem services (ES) and disservices (EDS) trade-offs for the irrigated corn growing area in South Platte River Basin of Colorado. The analysis integrated a high-resolution quantification of ES and EDS associated with different fertilization and irrigation applications through the use of a process-based biogeochemical model (DayCent), the calculation of Pareto frontiers for the trade-offs among multiple ES and EDS considering uniform management practices, and the identification of optimum fertilizer and irrigation rates of each polygon in the landscape based on different hypothetical management preferences. We assessed five ES and EDS indicators including net primary production, soil organic carbon, water use, nitrogen leaching, and greenhouse gas emissions. Our results showed the influence of management practices and site-specific factors on the provision of the ES and EDS and various trade-offs among ES and EDS, management practices, EDS footprints, and management objectives. Inappropriate applications of fertilizer and irrigation could lead to sub-optimal trade-

offs of ES and EDS. Although, the analysis of ES and EDS trade-offs is case-specific, we believe that our approach can be adopted for other locations as well as for more complex studies for decision support in sustainable agriculture. Further research is encouraged to improve this approach and to prepare the infrastructure for its widespread use in public- and private-sector decision making.

CHAPTER 4. HIGH-RESOLUTION MULTI-OBJECTIVE OPTIMIZATION OF FEEDSTOCK LANDSCAPE DESIGN FOR HYBRID FIRST AND SECOND GENERATION BIOREFINERIES.

4.1 SUMMARY

Biofuels have been proposed as a potential solution for climate change mitigation. However, there exist several barriers, such as "food vs fuel" issues and technological constraints, restricting the sustainable commercialization of both first- and second-generation biofuels. Combining arable crops and their residues for hybrid first- and second-generation biofuels productions provides opportunities to overcome these barriers. This study presents a high-resolution quantitative tool to support decision-making in feedstock planning for hybrid biofuel supply chains. We demonstrate this work with a case study on optimizing feedstock landscape design for a hybrid corn grain- and stover-based ethanol production system at Front Range Energy biorefinery, Windsor, Colorado, USA using a life-cycle approach. The case study considered three competing design objectives including the minimization of production costs, farm-to-refinery greenhouse gas emissions (GHG), and nitrogen (N) leaching, subject to constraints in land use and biofuel feedstock demand. Social costs of carbon (SC-CO_2) and nitrogen leaching (SC-NL) were used as weighting factors for GHG and N leaching in the objective function. Our results showed a broad win-lose 3D Pareto surface among the three design objectives and various modulating patterns of the required feedstock area, management input investments, and corn grain- and stover-based ethanol ratio associating with Pareto optimum solutions. Changes in feedstock landscape design were most sensitive to the variations of the SC-CO_2 between $400 -

$800 per Mg CO_2e, SC-NL between $0 - $50 per kg N leaching, and their ratio between 0 - 350, respectively.

4.2 INTRODUCTION

Biofuels have received considerable attention as part of a solution to reduce a nation's dependency on petroleum and reduce net CO_2 emissions from fuel use. In particular, the European Union has targeted a 10% share of energy from renewable sources, including biofuels, by 2020 in the transport sector while the United States Renewable Fuels Standard 2 mandated an annual production target of 136 billion liters of biofuels by 2022 (EPA-RFS2, 2010). Biofuels can be derived from many sources such as food crops like corn grain and sugarcane (first-generation) or ligno-cellulosic materials like crop and forest residues or dedicated grasses (second-generation). Comparing to the second-generation, first-generation technology is more commercially mature. However, its sustainability is currently debated due to the potential threats on food supplies and biodiversity, and consequential impacts of indirect land use change on total greenhouse gas (GHG) emissions (Searchinger et al., 2008; Gnansounou et al., 2009; Farrell et al., 2006; Reijnders and Huijbregts, 2007; Liska et al., 2009). Second-generation biofuels were proposed to solve these problems since they can use non-food sources as feedstock and marginal land for feedstock production. Nevertheless, due to the immature market structure and conversion technology, commercially viable production of second-generation biofuels is still challenging (Awudu and Zhang, 2012).

Combining arable crops and their residues such as corn grain and corn stover for a hybrid first- and second-generation biofuels production provides a potential solution to the large-scale and sustainable production and use of biofuels. This solution takes advantage of first-generation infrastructure but requires less productive land to secure the biomass feedstock, thus imposing

less pressure on food commodities and biodiversity. However, this coupling necessitates further efforts to design and develop more sustainable and robust biofuel supply chains, emphasizing multiple feedstocks production and management.

Although there has been a relatively large number of works on analyzing and optimizing biomass energy supply chains (Yue et al., 2014), we found that most of the studies focused on the sole production of first-generation (Zamboni et al., 2009a; Zamboni et al., 2009b; Mele et al., 2011; Mele et al., 2009; Dal-Mas et al., 2011; Corsano et al., 2011; Akgul et al., 2011) or second-generation biofuels (Santibañez-Aguilar et al., 2011; Čuček et al., 2010; Leduc et al., 2010; Huang et al., 2010; Parker et al., 2010; Tittmann et al., 2010). There is limited research that attempted to optimize biomass supply chain for hybrid first- and second-generation biofuel systems (Giarola et al., 2011). In addition, the optimization problems were mainly directed towards supply chain logistics, especially the design of transportation network and facility siting, using mathematical-programming-based tools such as Mixed Integer Programming. There is an obvious dearth of focus on optimizing feedstock cultivation and land management in a spatially-explicit manner with high spatial and temporal resolutions. We believe that this is important to harness the "management swing potential" (Davis et al., 2013) between different biofuel feedstocks cultivation, especially in hybrid biofuel systems, for the most cost-effective and sustainable biofuel production.

Formulating optimization problems using mathematical-programming approaches in previous studies has advantages in terms of model generalization, simulation efficiency, and exact solution search. These approaches also fit many supply chain logistic problems such as transportation or process design and operation at strategic and tactical levels. However, for higher-resolution optimization of biofuel feedstocks cultivation and land management, these

approaches are poorly-suited since it is inadequate to use a set of simplified equations to depict the dynamic ecosystem changes induced by different land management. Such changes are the results of complex and non-linear interactions among the site-specific weather conditions, soil properties, crop types, cropping practices, and land use history (Nguyen et al., 2017). For example, high biomass removal for biofuel not only increases the feedstock yield but also affects the subsequent farming input requirements, CO_2 emissions from carbon stock change, susceptibility to erosion and other ecosystem services (Davis et al., 2013). The extent of these subsequent impacts depends on topography, feedstock types, weather and soil conditions, and many other factors. Accurately capturing these changes in the optimization is the key to finding the true optimum solutions for biofuel supply chain design.

The use of biophysical models offers more detailed quantification of cropping system changes associated with the changes in land management (Kragt and Robertson, 2014; Balbi et al., 2015). These models such as CENTURY (Parton et al., 1994), DayCent (Del Grosso et al., 2000), DNDC (Li et al., 1992), RothC (Coleman and Jenkinson, 1996), and APSIM (Keating et al., 2003) can better represent site-specific interactions of biophysical characteristics (e.g., climate and soil characteristics). They can also capture the current and historical influence of land use and management practices on potential future changes in many ecosystem functions such as soil carbon stocks (Paustian et al., 2010). The coupling of biophysical models with geographic information system (GIS) and optimization algorithms such as the "Agricultural Ecosystem Service Optimization" (Ag-EcoSOpt) system (Nguyen et al., 2018) provides useful platforms for the spatially-explicit quantification and optimization of agricultural ecosystem service provisioning as well as the more specific feedstock production for biofuel supply chains at very high spatial resolutions. Ag-EcoSOpt uses a process-based biogeochemical model (DayCent) for

exploratory landscape simulations to investigate the production potential of a cropping system. The simulated outcomes can then be optimized with mathematical algorithms for different design and management objectives. The resulting solutions can be shown and analyzed via maps and graphs to support decision-making. Although, this approach is a discrete optimization, fine-scale exploratory simulations of the target systems (i.e., better inputs for optimization) will likely result in approximate solutions that are very close to the global optimum for any biofuel supply chain optimization problems.

Multi-objective optimization has also been used in many studies to optimize for competing objectives within biofuel supply chain design and management (Zamboni et al., 2009a; Corsano et al., 2011; Santibañez-Aguilar et al., 2011; Bernardi et al., 2012; Yue et al., 2013). This type of optimization typically generates a set of Pareto-optimal solutions (also called non-dominated solutions) where any improvement in one objective always impairs some others. The goal of multi-objective optimization is to help the decision maker find a solution that is preferable to them. We found that most previous studies have focused on the bi-objective optimizations of costs and greenhouse gas emissions or water footprints for biofuel supply chain. Only a few studies considered more than two design objectives (You et al., 2012; Bernardi et al., 2013). When it comes to ecosystem-level impact analysis, Čuček et al., (2012) argued the importance of considering other environmental criteria in the supply chain optimization such as nitrogen, land use, and biodiversity footprints.

The main goal of this study was to develop an optimization extension for Ag-EcoSOpt to facilitate the spatially-explicit and multi-objective optimization of hybrid biofuel supply chains, focusing on the design of feedstock landscape and the identification of optimum land management. We present this work with a case study on a tri-objective optimization of feedstock

landscape design for a hybrid first- and second-generation ethanol production at Front Range Energy biorefinery (FRE), Windsor, Colorado, USA. The case study considered three competing objectives including the minimization of production costs, farm-to-refinery greenhouse gas emissions, and nitrogen leaching with a life-cycle approach.

4.3 METHOD

4.3.1 Study site

The Front Range Energy (FRE) biorefinery is a local fuel ethanol supplier with a capacity of 1.51×10^8 L of ethanol per year (~ 40 million U.S. gallons per year). Our analysis assumed that FRE can produce both first generation and cellulosic biofuels and that the feedstock comes mainly from corn grain and corn stover in the study region. Corn stover is the residue of stalk, leaf, husk, and cob left behind following a corn harvest. The targeted feedstock production area is comprised of irrigated corn fields in 6 counties within an 80 km-radius (~50 miles) surrounding the biorefinery, including Adams, Boulder, Broomfield, Larimer, Morgan, and Weld (Fig. 4.1). These counties are located in north-central Colorado, and are among the most productive irrigated agricultural regions in the state with a majority of fine-loamy soils, an average annual rainfall of 38 cm, an average July high temperature of 30.6 ^{o}C, and an average January low temperature of -8.5 ^{o}C (Mesinger et al., 2006).

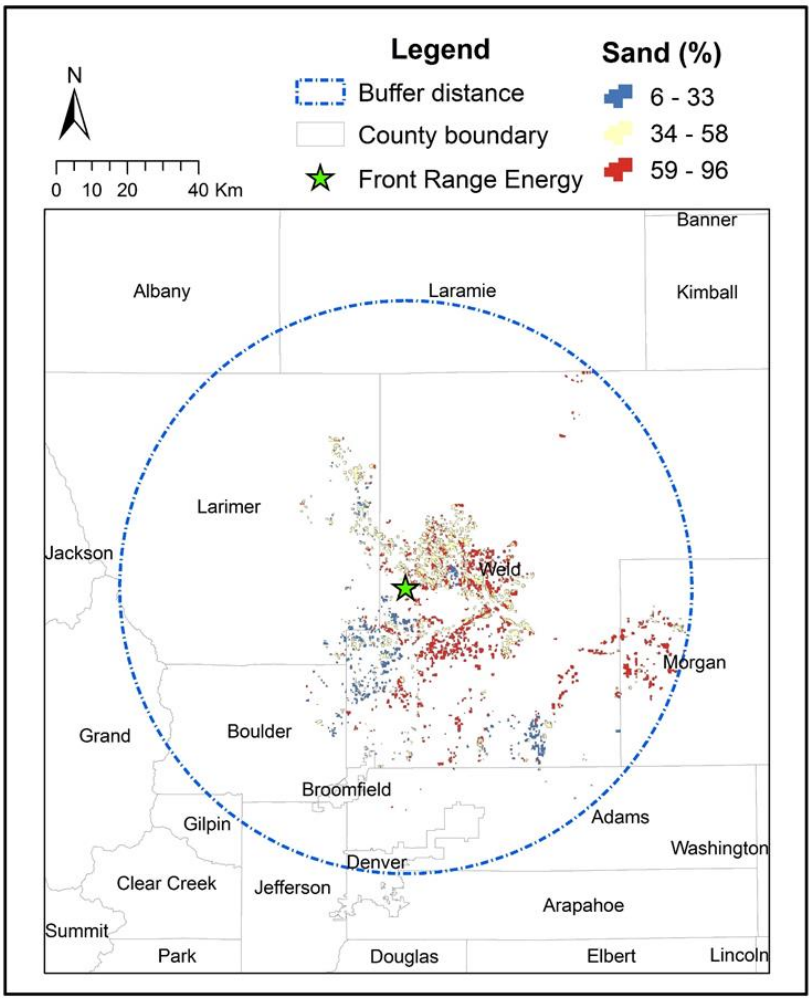


Fig. 4.1. Irrigated corn land units with the percent sand content of the soils within the study region. The blue circle is the 80km-radius buffer around Front Range Energy.

4.3.2 Scope of the analysis

The main objective of this case study was to identify the Pareto surface for the trade-offs among three competing design objectives including minimization of production costs, farm-to-refinery greenhouse gas emissions (GHG), and nitrogen (N) leaching. For each Pareto-optimal solution on the surface, we also identified the spatial configuration of the feedstock landscape as well as the optimum management practices (e.g., rates of N fertilization, irrigation, stover removal, and

tillage intensity) for each corn field in the landscape. The corresponding ratio between grain-based and stover-based ethanol was also calculated.

Our study used the life cycle optimization framework outlined in Yue et al., (2013). This framework follows the first three phases (goal and scope definition, inventory analysis, and impact assessment) of a classical life cycle assessment (LCA) (Finkbeiner et al., 2006), and performs the last phase (LCA interpretation) by coupling optimization tools with impact assessment. The system boundary for our optimization analysis was from field to biorefinery gate. The functional unit used to compare the GHG intensity of ethanol and conventional gasoline (CG) was one megajoule (MJ). Farm-to-refinery emissions for grain- and stover-derived ethanol were calculated using detailed, high resolution data in our analysis and then extended using refinery-to-wheel emissions reported in the literature to estimate the full attributional GHG reductions of biofuels relative to CG. The reference life-cycle emission of CG used in our study was 94 g MJ^{-1} (Wang et al., 2012).

Since corn grain and corn stover were treated as two feedstocks for biofuel production in our analysis, we applied the process-level co-product handling method to allocate cost and environmental burdens between the two products (Wang et al., 2015). This method assigns the burdens of individual process steps within the system boundary to the product that is responsible for the existence of those steps. The burdens of the shared steps and components can be allocated between co-products based on their energy contents or market values. We compared both energy content and market value allocation in our analysis. The energy contents in terms of lower heating values (LHV) used for corn grain and corn stover in our study were 15.01 and 16.22 MBTU Mg^{-1}, respectively (UChicago Argonne, 2015). For market value allocation, we used a 10-year averaged U.S. corn grain price of \$180 Mg^{-1} reported by Ycharts (https://ycharts.com/)

for the period from 2007 to 2017 and an average stover price of \$70 Mg^{-1} reported by Thompson and Tyner, (2011).

The Agricultural Ecosystem Service Optimization (Ag-EcoSOpt) tool (Nguyen et al., 2018) was employed to simulate changes in ecosystem indicators at a high spatial resolution under variations in management practices. The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation Model – commonly known as GREET (UChicago Argonne, 2015) - was used to calculate emissions associated with farm operations and feedstock transportation. The supply chain budgeting was carried out using empirical data. The optimization problem was formulated in PuLP (Mitchell et al., 2011), a Python's linear programing (LP) modeler, and solved by CPLEX solver (IBM®, 2017). We used the weighted sum method to iteratively solve the multi-objective optimization problem and employed the natural neighbor interpolation method to approximate the Pareto surface representing trade-offs among the three design objectives.

4.3.3 Landscape modeling of the study site

4.3.3.1 Agricultural Ecosystem Service Optimization tool

The Ag-EcoSOpt tool (Nguyen et al., 2018) includes a core biogeochemical model (DayCent), a database of historical land use for U.S. crop land, a geospatial component for the spatially-explicit processing of inputs and outputs from DayCent's landscape simulations, and an optimization component for land use and land management decision making. The DayCent model (Parton et al., 1998; Del Grosso et al., 2000) is the daily time step process-based model that simulates the dynamics of biogeochemical flows of carbon and nitrogen among the atmosphere, vegetation, and soil. The model has been widely used to study soil carbon and nitrogen dynamics, and soil greenhouse gas (CO_2, N_2O, CH_4) emissions especially for

agricultural ecosystems (Del Grosso et al., 2002; Delgrosso et al., 2005; Del Grosso et al., 2008a; Del Grosso et al., 2008b; Kim et al., 2009; David et al., 2008; US EPA, 2015).

The site-specific input data needed to initialize and run DayCent include climate (i.e. maximum and minimum air temperature, precipitation), soil properties, vegetation types, land use history, and management practices (e.g., irrigation, tillage, fertilization). These input data were acquired for the study region through the Ag-EcoSOpt's automated geospatial component from several spatial databases (Table 4.1). The shapefiles of these databases were then intersected to create 4,782 polygons of various sizes representing irrigated corn fields in the study site, including 673 unique combinations of soil and weather data (DayCent input strata) (Fig. 4.1). Each unique stratum was simulated with DayCent and the results were linked back to their associated landscape polygons for optimization and spatial visualization.

For each input stratum, DayCent was initialized using historical land use information from the Ag-EcoSOpt database. The database contains pre-settlement and historical agricultural land use assumptions for Major Land Resource Areas (MLRA) up to 1979 (Ogle et al., 2010). The stratum was then extended with the simulation of a business-as-usual scenario (BAU) for corn production from 1980 to 2016. The BAU scenario for Colorado's irrigated corn crop was 170 kg of nitrogen per hectare, 34 cm of irrigated water per growing season, 75% residue removal, and conventional tillage (ERS-ARMS, 2010). The exploratory landscape simulation was continued from the BAU for a 30-year period (from 2017 – 2047) under the management practice scenarios defined in section 4.3.3.2. These simulations were executed in parallel on an 18-node, 216-processor cluster computing system at the Colorado State University's Natural Resources Ecology Laboratory. The detailed modeling process was described by Nguyen et al., (2018)

where the authors used Ag-EcoSOpt to simulate corn production in the South Platte River Basin, Colorado, USA.

Table 4.1. Summary of spatial data inputs

Spatial database	Data type	Year	Native Resolution	Source
SSURGO	Soils	2014	1:12,000 to 1:63,360	(NRCS-USDA, 2014)
NARR	Daily weather	1979 - 2010	32km	(Mesinger et al., 2006)
MIrAD-US	Irrigation extent	2007	250m	(Pervez and Brown, 2010)
Crop Data Layer	Crop types	2016	30m	(NASS-CDL, 2016)

4.3.3.2 Scenarios for exploratory landscape simulations

Besides the BAU scenario, we performed a combinatorial set of simulations for every corn field in the landscape, for a total of 180 management scenarios permuting five N fertilization rates, four irrigation rates, three stover removal rates, and three levels of tillage intensity. The five N fertilization rates, comprising 70, 110, 150, 190, and 230 kg N ha^{-1}, were chosen to reflect a common range of fertilization for irrigated continuous corn in the area (Halvorson et al., 2006). Irrigation rates were dynamically calculated from soil-specific field capacity (FC). We simulated four levels of irrigation representing full irrigation to 100% FC, and three limited irrigations at 80%, 60%, and 40% FC. Our limited irrigation scenarios were characterized by the reduction in the amount of irrigated water applied. The three rates of stover removal were 22, 52, and 83% of the total residue. These rates represented low, moderate and high stover harvest scenarios, respectively (Muth Jr. et al., 2013). We also simulated three levels of tillage intensity, including conventional tillage (CT), reduced tillage (RT), and no-till (NT). Conventional tillage was defined as multiple tillage operations every year with significant soil inversion through moldboard plowing. Reduced tillage lessened soil inversion by using chisel plowing. No-till was

defined as not disturbing the soil except through the use of fertilizer and seed drills (Ogle et al., 2010). The outputs from the combinatorial simulation of 181 management scenarios on 4,782 corn fields in the landscape were used in the optimization.

4.3.4 Feedstock landscape design objectives

4.3.4.1 Production costs

The economic data used for calculating production costs were obtained from different sources (Thompson and Tyner, 2011; Vadas and Digman, 2013; Russell et al., 2016; Ibendahl et al., 2015), as reported in Appendix B1. All the monetary values acquired from the literature were adjusted for inflation to 2017 US dollars. The production costs, including farm production costs and feedstock transport costs, were calculated for every corn field in the study region. Farm production costs included tillage-based fixed costs and variable costs of N fertilization and irrigation, land rent, and feedstock-dependent costs for harvest, grain drying, storage, and handling. Feedstock transport costs were calculated based on feedstock yields and the transport distance from each field to the FRE refinery. The Euclidean distances from fields to the biorefinery were calculated through Ag-EcoSOpt's geospatial component. DayCent reports corn grain and harvested stover as grams of carbon per square meter. We converted corn grain and harvested stover to Mg of dry weight biomass per hectare per year (averaged over entire simulation period) assuming 43.5% carbon content and moisture content of 15.5% for corn grain and 20% for harvested stover (Gesch et al., 2010). All commodity prices were treated as exogenous and static regardless of management practice changes in the analysis.

4.3.4.2 Farm-to-refinery greenhouse gas emissions

The farm-to-refinery GHG emissions associating with biofuel feedstock production included soil, farm operation, and feedstock transport emissions. Soil emissions were calculated based on

soil-related emission sources modeled with DayCent, including annual CO_2, N_2O, and CH_4 emissions (Eve et al., 2014). Annual CO_2 emission was assumed to come mainly from soil microbial respiration of SOC and calculated by taking the difference between SOC levels during two consecutive years. Annual N_2O emission consisted of direct and indirect N_2O emissions. The direct N_2O emission was reported by DayCent in term of N_2O efflux, while the indirect N_2O emissions were computed from volatilized N ($NOx + NH_3$) fluxes and nitrogen leaching with the emission factors (EF) of 0.01 and 0.075, respectively (Eve et al., 2014). Annual N_2O and CH_4 were converted into carbon dioxide equivalent (CO_2e) by using their 100-year global warming potential values (GWP_{100}) (IPCC, 2006).

To calculate GHG emissions associated with farm operations and feedstock transportation, we used the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation Model (GREET) (UChicago Argonne, 2015). The GREET's default pathway for Integrated Corn/Stover Ethanol was updated with our specific inventory data. GREET's soil emissions were replaced with site-and-management specific values modeled with DayCent. Since we only focused on the effects of N fertilizer use, we assumed the same P application rate of 76.21 kg P ha^{-1} for all scenarios. No K fertilizer, manure, or lime applications were simulated in our analysis. The amounts of herbicide (7g per Mg of dry wt. corn grain) and insecticide (0.06g per Mg of dry wt. corn grain) were taken from the GREET's default repository for corn farming and were assumed as application rates for conventional tillage. The herbicide and insecticide application rates were increased by 4% for reduced tillage, and 47% for no-till compared to those of conventional tillage, as reported in Penn State Extension (2017).

The information on fuels used for farm operation was obtained from different sources (Rathke et al., 2007; Luo et al., 2009; Vadas and Digman, 2013) to compute the average energy use for

different tillage types in our analysis (Appendix B2). The energy use for irrigation was 107,242 Btu cm^{-1}, calculated based on the method presented in Martin et al., (2010) assuming an average pumping lift of 80-feet and a pump discharge pressure of 45 pounds per square inch.

Corn grain was assumed to be transported to the stack by medium-heavy duty truck (MHDT) and from the stack to the refinery by heavy-heavy duty truck (HHDT). Corn stover was assumed to be transported by MHDT from field to the refinery in the form of dry bales. The transportation emissions were calculated with GREET's default emission factors and the Euclidean distances from farms to the refinery. The emission factors used in our analysis are summarized in Table 4.2. The GHG emissions were reported in terms of Mg of CO_2e per hectare per year.

Table 4.2. GREET's emission factors for farm operation and feedstock transport

Emission source	Unit	Emission factor
Farm energy	g CO_2e MJ^{-1}	90
N fertilizer	g CO_2e g^{-1}	3.9
P fertilizer	g CO_2e g^{-1}	1.5
Herbicide	g CO_2e g^{-1}	20
Insecticide	g CO_2e g^{-1}	23
Storage and handling of stover	g CO_2e Mg^{-1}	1167
Corn transportation	g CO_2e Mg^{-1} km^{-1}	208
Stover transportation	g CO_2e Mg^{-1} km^{-1}	240
Corn ethanol refinery-to-wheel emissions	g CO_2e MJ^{-1}	35
Stover ethanol refinery-to-wheel emissions	g CO_2e MJ^{-1}	16

4.3.4.3 Nitrogen leaching

Nitrogen (N) leaching due to corn grain and stover production for ethanol was calculated by the nitrogen submodel in DayCent. DayCent calculates N leaching as a function of soil nitrate (inorganic N leaching) and active soil SOM pool decomposition (organic N leaching), soil texture, and the amount of water moving through the soil profile. N leaching was reported in the

unit of gram of nitrate-N per square meter by DayCent and was converted to kg of nitrate-N per hectare per year averaged over the entire simulation period.

4.3.5 Multi-objective optimization of feedstock landscape design

4.3.5.1 Formulation of the optimization problem

To integrate different competing design objectives for linear programming, we used the weighted sum method (WS) (Fishburn, 1967). This method converts a multi-objective optimization problem into a single-objective one using a substitute objective function. Each objective was assigned a weight relative to its importance in decision-making and the weighted sum of all objectives was then be optimized. Since WS method is only applicable when all objectives are expressed in the same unit, we monetized GHG emissions and N leaching by multiplying their actual values with the corresponding social costs, where the social costs of GHG emissions (SC-CO_2) and N leaching (SC-NL) were defined as the estimated damages to the ecosystem in monetary term caused by an incremental increase in CO_2e emissions and N leaching, respectively (Tol, 2011; Brink and Grinsven, 2011). The multi-objective optimization problem was formulated as follows, with detailed nomenclature presented in Table 4.3:

Let $i \in \{1, \ldots, NF\}$ be the corn fields in the landscape and $j \in \{1, \ldots, MP\}$ be the management practice scenarios of each corn field. (NF = 4782, MP = 181).

For all i, let $x_{i,j} \in \{0, \ldots, 1\}$ be the fraction of field i where management practice scenario j is chosen. We assumed that a field can be divided into subfields for multiple management practice scenarios.

Objective:

$$\text{Min} \sum_i^{NF} \sum_j^{MP} x_{i,j} * (w_1 * \text{Cost}_{i,j} + w_2 * (\text{SC-CO}_2 * \text{GHG}_{i,j}) + w_3 * (\text{SC-NL} * \text{NL}_{i,j})) \quad \text{(Eq 4.1)}$$

Subject to (s.t.):

$$x_{i,j} \in \{0, \ldots, 1\} \quad \text{(Eq 4.2)}$$

$$\sum_j^{MP} x_{i,j} \leq 1, \qquad \forall\, i \in NF \quad \text{(Eq 4.3)}$$

$$TC = \sum_i^{NF} \sum_j^{MP} x_{i,j} * CY_{i,j} \quad \text{(Eq 4.4)}$$

$$TS = \sum_i^{NF} \sum_j^{MP} x_{i,j} * SY_{i,j} \quad \text{(Eq 4.5)}$$

$$EtOH = TC * CTE + TS * STE \quad \text{(Eq 4.6)}$$

$$EtOH \geq minCAP \quad \text{(Eq 4.7)}$$

$$GHGcp = \frac{\sum_i^{NF} \sum_j^{MP} x_{i,j} * GHGc_{i,j}}{TC * CTE * LHVe} \quad \text{(Eq 4.8)}$$

$$GHGsp = \frac{\sum_i^{NF} \sum_j^{MP} x_{i,j} * GHGs_{i,j}}{TS * STE * LHVe} \quad \text{(Eq 4.9)}$$

$$\frac{GHGgs - (GHGcp + GHGcep)}{GHGgs} * 100 \geq 20\% \quad \text{(Eq 4.10)}$$

$$\frac{GHGgs - (GHGsp + GHGsep)}{GHGgs} * 100 \geq 60\% \quad \text{(Eq 4.11)}$$

Table 4.3. Nomenclature of the optimization problem

Name	Description
	Indices
i	Index for fields in the landscape
j	Index for management practice scenarios in field i
	Set sizes
NF	Number of corn fields in the landscape (NF = 4782)
MP	Number of considered management practices for each field (MP = 181)
	Parameters
$Cost_{i,j}$	Production costs of field i under management practice scenario j ($ year^{-1})
$GHG_{i,j}$	Farm-to-refinery GHG emissions of field i under management practice scenario j (Mg CO_2e year^{-1})
$NL_{i,j}$	N leaching of field i under management practice scenario j (kg N year^{-1})
SC-CO_2	Social cost of GHG emissions ($ per Mg CO_2e)
SC-NL	Social cost of N leaching ($ per kg N)
w_1	Weighting coefficient for production costs
w_2	Weighting coefficient for GHG emissions, set to 1 for convenience (see section 2.4.2)
w_3	Weighting coefficient for N leaching, set to 1 for convenience (see section 2.4.2)
TC	Total corn grain used for grain-based ethanol production (Mg dry wt. mass year^{-1})
TS	Total harvested stover used for stover-based ethanol production (Mg dry wt. biomass year^{-1})
EtOH	The total amount of ethanol produced from corn grain and harvested stover (L year^{-1})
CTE	Corn grain to ethanol ratio (CTE = 429 L Mg^{-1} (113.4 gallons Mg^{-1}))
STE	Stover to ethanol ratio (STE = 355 L Mg^{-1} (93.7 gallons Mg^{-1}))
minCAP	Ethanol production capacity of FRE (minCAP = 1.51 x 10^8 L year^{-1})
GHGcp	Farm-to-refinery GHG emissions of grain-based ethanol (g CO_2e MJ^{-1})

$GHGsp$	Farm-to-refinery GHG emissions of stover-based ethanol (g CO_2e MJ^{-1})
$GHGc_{i,j}$	GHG emissions attributed to corn grain production for farm i under management practice scenario j (g CO_2e Mg^{-1})
$GHGs_{i,j}$	GHG emissions attributed to stover production for farm i under management practice scenario j (g CO_2e Mg^{-1})
$LHVe$	Lower heating value of ethanol ($LHVe$ = 21.2 MJ L^{-1})
$GHGcep$	Refinery-to-wheel GHG emissions of grain-based ethanol ($GHGcep$ = 34 g CO_2e L^{-1})
$GHGsep$	Refinery-to-wheel GHG emissions of stover-based ethanol ($GHGsep$ = 15 g CO_2e L^{-1})
$GHGgs$	Well-to-wheel GHG emission of conventional gasoline ($GHGgs$ = 94 g CO_2e L^{-1})

Binary variables

$x_{i,j}$	models the fraction of field i and the corresponding management practice scenario j is chosen in the final feedstock landscape for FRE.

The objective (Eq 4.1) represents the minimization of production costs, GHG emissions, and N leaching associated with the feedstock production for both grain- and stover-derived ethanol. The importance of each design objective in the integrated objective function was expressed by their weighting coefficients. The constraints (Eq 4.2) and (Eq 4.3) allow the selection of multiple management scenarios for each corn field as long as the sum of their corresponding areal fractions does not exceed 1. The constraints (Eq 4.4), (Eq 4.5), (Eq 4.6), and (Eq 4.7) ensure that enough corn grain and stover are produced to meet the feedstock demand for 1.51 x 10^8 L of ethanol per year. We assumed that if a field is chosen as a feedstock site, all the corn grain and harvested stover from that field will be used for ethanol production. Corn grain and stover biomass were converted to ethanol using the conversion ratios of 429 L per Mg of dry weight mass of corn grain (~ 113.4 gal Mg^{-1}) and 355 L per Mg of dry weight biomass of corn stover (~93.7 gal Mg^{-1}), respectively (UChicago Argonne, 2015). The constraints (Eq 4.8), (Eq 4.9), (Eq 4.10), and (Eq 4.11) require at least 20% and 60% reductions in life-cycle GHG emissions of grain-based and stover-based ethanol versus conventional gasoline, respectively. This is pursuant to the U.S. Environmental Protection Agency's (EPA) approved fuel pathways under the

Renewable Fuel Standard 2 Program (EPA-RFS2, 2010). This set of constraints is driven by the co-product handling method and might produce different solutions depending on the allocation of GHG emission burdens between corn grain and corn stover.

4.3.4.2 Pareto surface simulation

Any adjustment in the weighting coefficients or social costs will likely result in changes in the integrated objective value (Eq 4.1), thus generating a different Pareto-optimal solution for the landscape design. Varying the weighting coefficients w_2 and w_3 might be appropriate when one wants to examine the trade-offs among the design objectives at a specific valuation of SC-CO_2 and SC-NL. However, for studies that aim to assess the effects of different SC-CO_2 and SC-NL on optimum landscape designs like ours, the variation of both weighting coefficients and social costs might add unnecessary complexities to the analysis. Therefore, we decided to fix the weighting coefficients w_2 and w_3 at 1 while varying SC-CO_2 and SC-NL to generate the Pareto-optimal set. The weighting coefficient w_1 was permitted values of 1 and 0 to allow the inclusion or exclusion of the production costs objective in the optimization. The steps used to create the Pareto-optimal set is shown in Fig. 4.2. Each of the Pareto-optimal solutions can be defined by a set of three design objective values (i.e., the landscape summations of production costs, GHG emissions, and N leaching), and thus lies on a Pareto surface, which is the 3-D graphical presentation of all the Pareto-optimal solutions. The extent of the Pareto surface was identified by iteratively solving bi-objective optimization problems (ignoring one objective). When only one objective was minimized, the algorithm generated a corner solution representing the approximation of global minimum for that objective. We simulated the corner solutions for all design objectives based on the setup in Table 4.4.

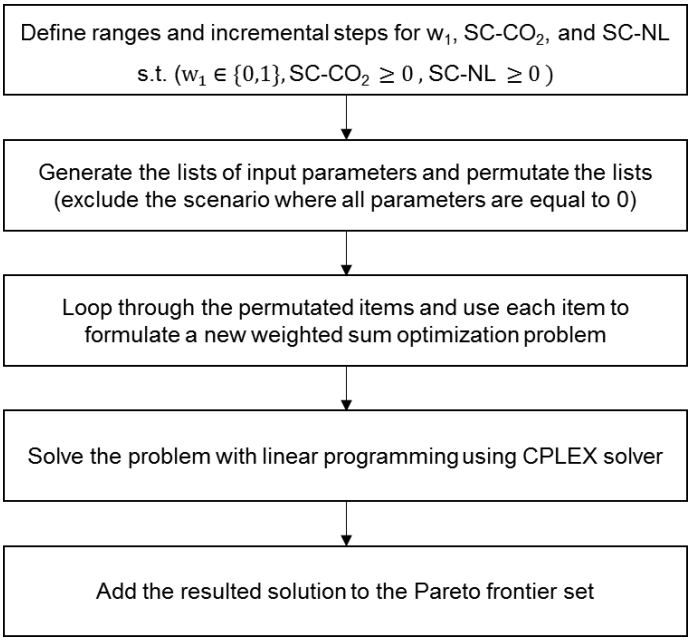


Fig. 4.2. Pareto frontier set simulation where SC-CO_2 and SC-NL are social costs of CO_2 and N leaching.

Since it is impossible to identify all points on a surface, we used interpolation techniques to construct the Pareto surface based on the simulated Pareto-optimal set. The natural neighbor interpolation (Sibson, 1981) was employed in this study through gridata function in MATLAB curve fitting toolbox (MATLAB and Neural Network Toolbox, 2017). This interpolation is a weighted moving average technique that uses geometric relationships between a query point and its closest subset of input samples to interpolate a value. The natural neighbor interpolation was used instead of other conventional methods such as nearest-neighbor, inverse-distance weighted averaging, and kriging because it has been proven to produce good results for unevenly distributed input datasets (Sambridge et al., 1995; Ledoux and Gold, 2005), which is one of the caveats when using WS method to generate the Pareto trade-off frontiers (Marler and Arora, 2010; Motta et al., 2012). Similar Pareto frontier approximation techniques have also been

applied in other studies (Bramanti et al., 2001; Wilson et al., 2001; Ruzika and Wiecek, 2005; Martín et al., 2005).

Table 4.4. Objective coefficients for corner objective scenarios

Objective scenario	Parameters		
	w_1	SC-CO_2	SC-NL
Minimizing Production costs (minCost)	1	0	0
Minimizing farm-to-refinery emissions (minGHG)	0	1	0
Minimizing N leaching (minNL)	0	0	1

Because our scope was to approximate the extent of the Pareto surface, we did not concentrate on the refinement of Pareto optimal solutions on the produced Pareto surface. Therefore, our method might not guarantee the property of non-dominance for many interpolated data points. The fine-tuning of the Pareto front and Pareto surface approximation, especially for nonconvex optimization problems with more than two objectives, can be done with more advanced algorithms such as adaptive weighted sum (Kim and Weck, 2006), modified normal boundary intersection, modified normal constraint (Motta et al., 2012), PAINT (Hartikainen et al., 2012), and Interactive Decision Maps (Lotov et al., 2013).

4.4 RESULTS AND DISCUSSION

4.4.1 Ecosystem responses to changes in management practices

Our results showed various trade-offs among ecosystem function indicators induced by changes in management practices (Fig. 4.3). Increasing N fertilization resulted in a win-lose situation between improving corn grain, harvested stover, and soil organic carbon (SOC) stock and increasing soil N_2O emissions and N leaching where N fertilizer rates were moderate to low ($\leq$ 110 kg N ha^{-1}). However, as the N fertilizer rate exceeded 110 kg N ha^{-1}, this trade-off became a lose-lose situation where the effects of additional N fertilizer on corn grain, harvested stover, and

SOC stock diminished and the negative impacts of N fertilization on elevating soil emissions and N leaching linearly increased. While higher irrigation intensified N leaching, the positive impacts of irrigation on biomass production increased SOC stocks due to higher C inputs from residues. This C credit offset a large portion of soil N_2O emissions. The increases in stover removal rate or tillage intensity reduced C inputs to soil resulting in lower SOC stocks and higher soil emissions. Corn grain yield and N leaching were quite unresponsive to changes in stover removal rate and tillage intensity. The complex interactions between ecosystem parameters and management practices illustrated the necessity of the subsequent optimizations of the landscape design for multi-criteria decision-making.

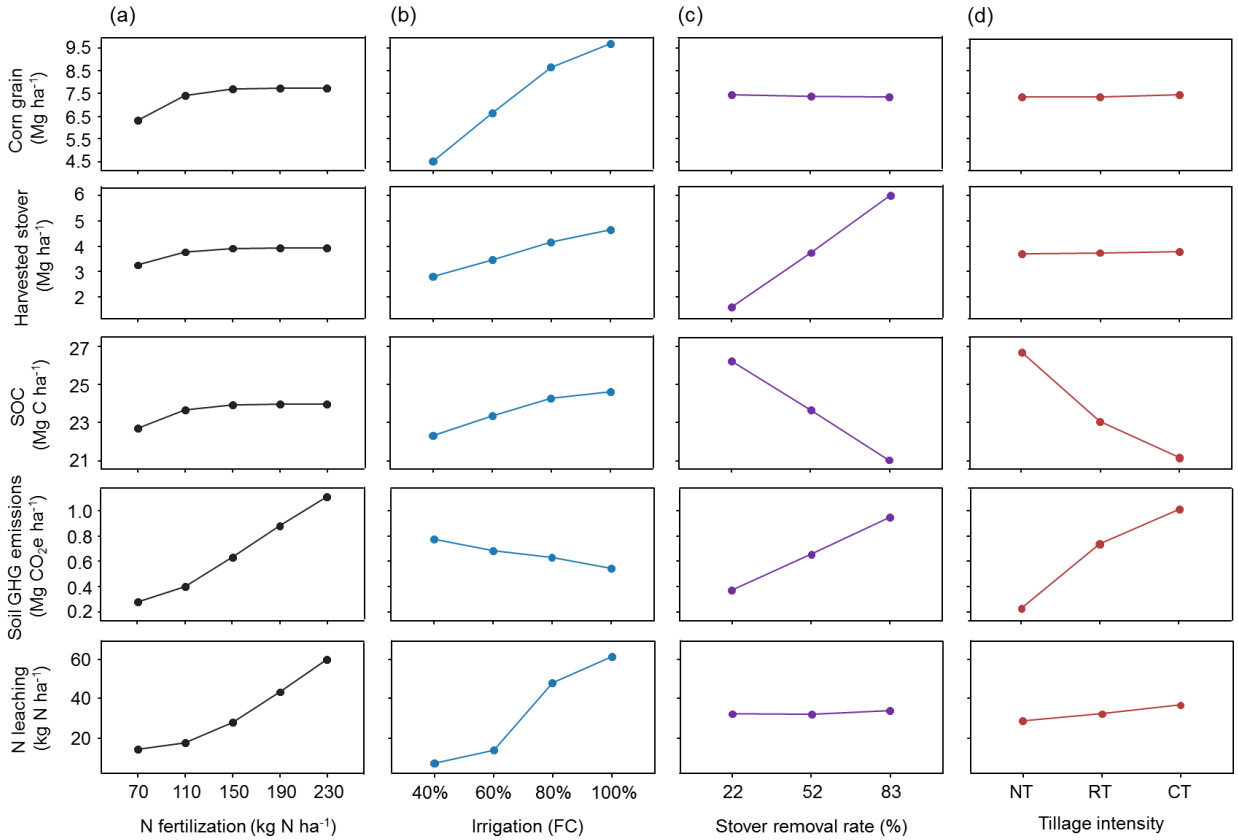


Fig. 4.3. Changes in ecosystem parameters by (a) N fertilizer rates averaged across other management practices; (b) irrigation rates averaged across other management practices; (c) stover removal rates averaged across other management practices; (d) tillage intensity averaged across other management practices. FC is soil-specific field capacity. Soil organic carbon (SOC)

is reported for top soil (0 – 20 cm) at the end of the 30-year simulation period. Corn grain and corn stover are reported as 30-year average of dry weight biomass at harvest. Soil GHG emissions are the 30-year average of all soil-related CO_2e emissions. N leaching is the 30-year average of nitrogen that leached out of the soil profile.

4.4.2 Corner solutions of feedstock landscape designs

We found win-lose trade-offs among different corner design objectives since the minimization of one objective exclusively resulted in increases in others (Table 4.5). These design objectives corresponded to the situations where single objectives were considered in the optimization. The minimum production costs, GHG emissions, and N leaching associated with the feedstock production for 1.51×10^8 L year^{-1} were \$$4.5 \times 10^7$ year^{-1}, 1.8×10^4 Mg CO_2e year^{-1}, and 0.06×10^5 kg N year^{-1}, respectively. The wide variation of each objective value across the objective scenarios suggested a large trade-off space for the simultaneous optimizations of the three design objectives.

Table 4.5. Corner solutions of feedstock landscape design for FRE*

Objective scenario	Objective value		
	Production costs ($\times 10^7$ \$ year^{-1})	GHG emissions ($\times 10^4$ Mg CO_2e year^{-1})	N leaching ($\times 10^5$ kg N year^{-1})
minCost	4.5	3.9	13.3
minGHG	6.3	1.8	3.7
minNL	5.9	4.1	0.06

Note: * The results are a summation of all corn fields in the final landscape based on the feedstock demand for 1.51×10^8 L year^{-1}.

The corner solutions were obtained using different land management input levels (Table 4.6) and spatial configurations (Fig. 4.4). The minimum production costs (minCost) was achieved with higher average management inputs but reduced feedstock area. This was because the fixed costs increased with increasing number of feedstock sites. To attain minimum GHG emissions

(minGHG), lower N fertilizer, irrigation, and stover removal rates were applied to reduce farm operation emissions and increase the C credits from SOC stock change. This led to a substantial expansion of feedstock area and a higher proportion of grain-based ethanol. To minimize N leaching (minNL), the model used low N fertilizer and irrigation rates to reduce N leaching sources while increasing stover removal rates to curtail the required feedstock area. No-till was dominantly applied as the optimum tillage type in all corner scenarios with 100% of landscape under minCost and minGHG scenarios and 94.3% under minNL scenario. The boxplots representing the detailed distributions of field management inputs associated with the corner solutions are provided in Appendix B3.

Table 4.6. Landscape average of management inputs for the corner solutions

Objective scenario	Landscape average* of			Feedstock area (10^3 ha)	Grain-based ethanol (%)
	N fertilizer (kg N ha^{-1})	Irrigation (cm $year^{-1}$)	Stover removal rate (%)		
minCost	150	45	83	19.6	61
minGHG	88	29	22	39.4	85
minNL	80	23	80	34.4	59

Note: * To calculate the landscape average of a scenario, each field was weighted by its area and the weighted sum was divided by the corresponding feedstock area.

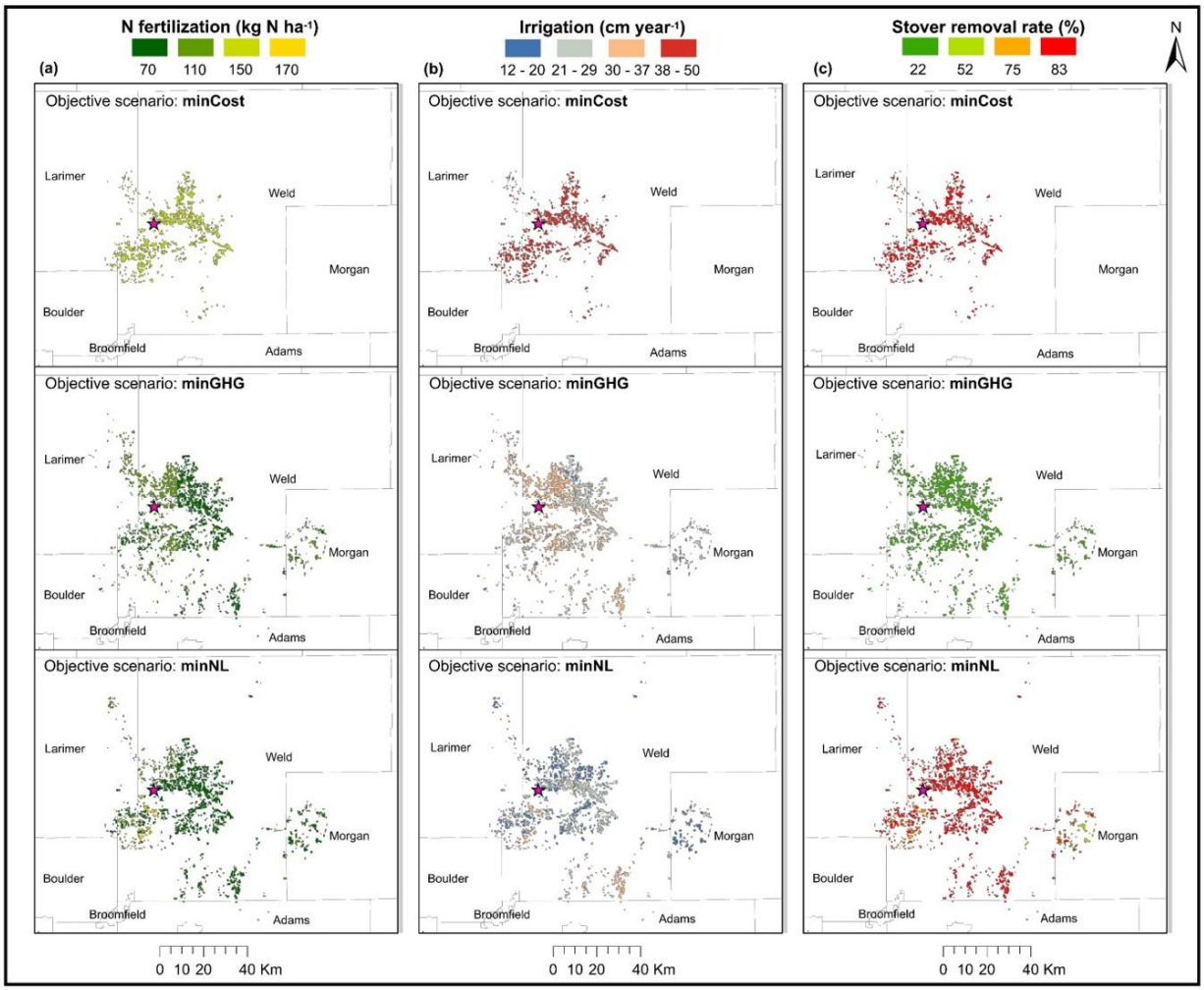


Fig. 4.4. Spatial configurations of (a) N fertilization, (b) irrigation, and (c) stover removal rates corresponding to Pareto-optimal solutions for the corner objective scenarios (minCost, minGHG, and minNL) presented in Table 4.4. The management practice rates are reported on a field-to-field basis.

4.4.3 Pareto trade-off surface for three design objectives

The Pareto surface for simultaneous minimization of production costs, farm-to-refinery GHG emissions, and N leaching exhibits non-smooth patterns (Fig. 4.5a). The surface's borders define the Pareto frontiers representing trade-offs between each pair of the design objectives (i.e., scenarios where one of the weighting coefficients is zero), including Pareto frontiers between

production costs and farm-to-refinery GHG emissions (Cost vs. GHG) (Fig. 4.5b), Pareto frontiers between production costs and N leaching (Cost vs. NL) (Fig. 4.5c), and Pareto frontiers between farm-to-refinery GHG emissions and N leaching (GHG vs. NL) (Fig. 4.5d). The two long tails associated with minCost and minNL scenarios suggest that the marginal improvements of production costs and N leaching objectives at these ends were accompanied with extreme trade-offs from other design objectives. In particular, we found that minimum production costs came with maximum N leaching while minimum N leaching corresponded to maximum GHG emissions. This was due to the decrease in required feedstock area and the intensification of management inputs. The maximum production costs of \$7.04 $\times 10^7$ $year^{-1}$ was identified on the GHG vs. NL frontier at the GHG emission of 2.22 $\times 10^4$ Mg CO_2e $year^{-1}$and N leaching of 0.46 $\times 10^5$ kg N $year^{-1}$. Using different co-product allocation methods did not alter the satisfaction of the emission reduction constraints (Eq. 4.10 and 4.11), thus it did not affect the solutions on Pareto surface. However, allocation methods influenced the attribution of economic and environmental burdens between feedstocks, thus affecting the feedstock costs and environmental footprints of corn grain and corn stover associating with the Pareto surface solutions (see Appendix B4).

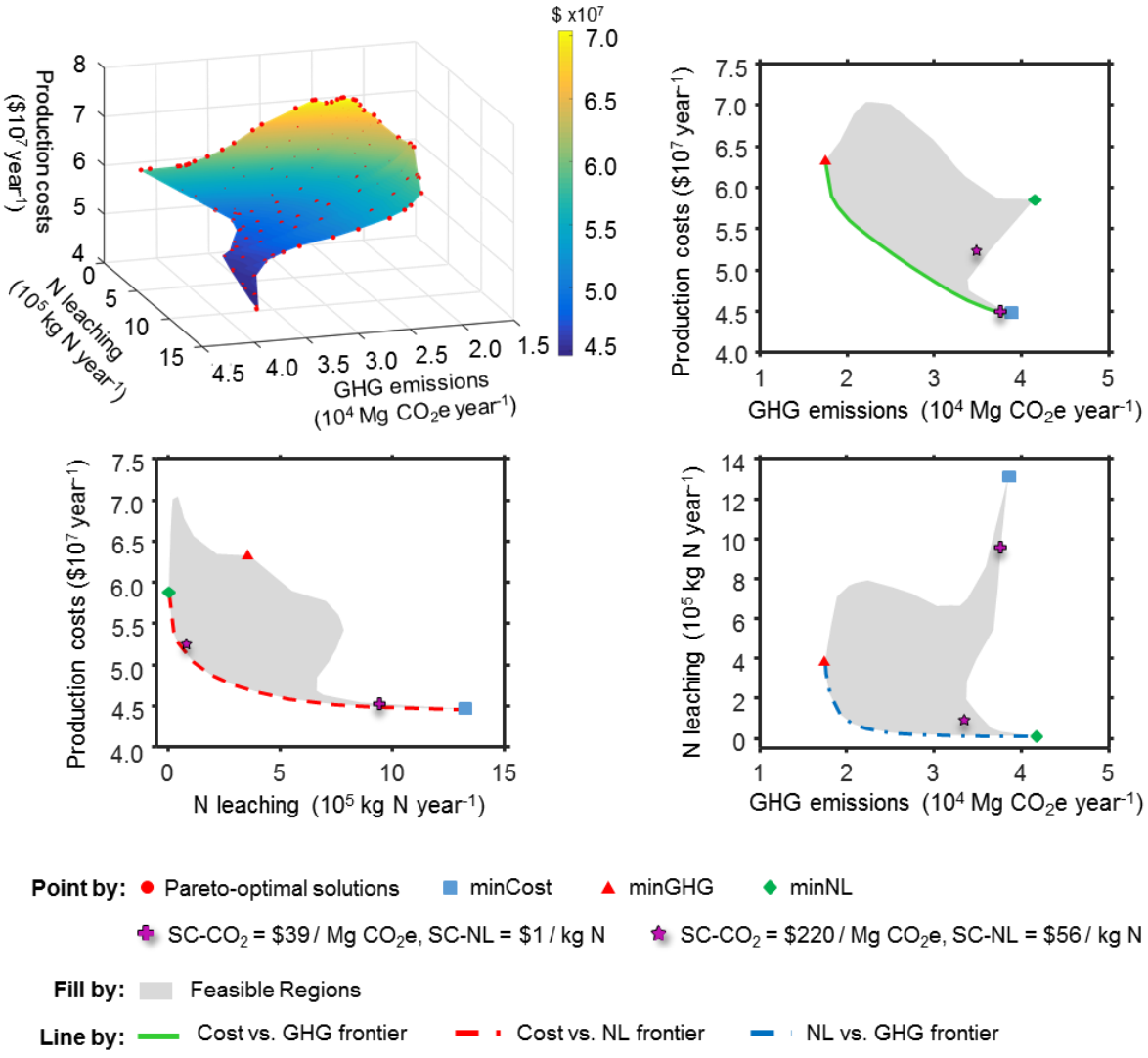


Fig. 4.5. Pareto surface for trade-offs among three design objectives. (a) 3D Pareto surface; (b) 2D projection of the 3D Pareto surface highlighting trade-off between production costs and farm-to-refinery GHG emissions, (c) 2D projection of the 3D Pareto surface highlighting trade-off between production costs and N leaching, (d) 2D projection of the 3D Pareto surface highlighting trade-off between N leaching and farm-to-refinery GHG emissions. The 3D Pareto surface is colored based on production costs. The colored points show Pareto-optimal solutions for scenarios where different social cost of CO_2 (SC-CO_2) and social cost of N leaching (SC-NL) were used as weighting coefficients in the objective functions (Eq. 4.1).

4.4.4 Optimum management practices

The modulation of management practices, feedstock area, and corn grain ethanol percentage to attain optimal solutions are given in Fig. 4.6. In general, our model maintained minimal stover removal rate while lowering N fertilization and irrigation to minimize GHG emissions whereas it retained minimal N fertilization and irrigation while increasing stover removal (thus reducing feedstock area) to minimize N leaching. To minimize production costs, our model attempted to reduce the feedstock area by orchestrating the rates of change among different management practices. For example, we observed constant increases in N fertilization, irrigation, and stover harvest when shifting from minNL to minCost scenarios and periodic adjustments of N fertilization, irrigation, and stover harvest rates moving along the Cost vs. GHG Pareto frontier. The minimum feedstock area required for 1.51×10^8 L year^{-1} was 19,629 ha corresponding to the minCost solution while maximum feedstock area was 47,324 ha obtained at a solution on the NL vs. GHG Pareto frontier. This was the situation where reducing N leaching and GHG emissions were equally weighted and thus the model sustained the lowest production level. The percentage of grain-based ethanol strongly correlated with stover removal (r = -0.98) and ranged from 59% to 85% of the total 1.51×10^8 L year^{-1}.

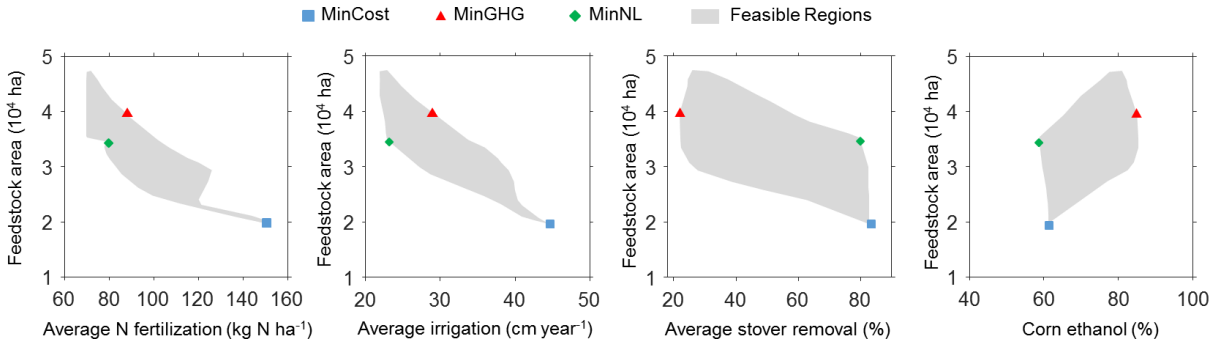


Fig. 4.6. The feedstock area, average management practices, and average corn ethanol percentage corresponding to solutions on the Pareto surface.

4.4.5 Emission credits from ethanol production

Based on the Pareto-optimal solutions, our results estimated an average reduction of 51 – 54% and 72 – 77% in life-cycle GHG emissions as compared to conventional gasoline for the production and utilization of grain-based and stover-based ethanol, respectively (Fig. 4.7). These corresponded to the avoided CO_2e emissions of 0.44 – 0.46 Mg CO_2e Mg^{-1} for corn grain and 0.51 – 0.54 Mg CO_2e Mg^{-1} for corn stover used as feedstock for biofuels (Fig. 4.7b), which outweighed those achieved by SOC increases as compared to BAU scenario (Fig. 4.7c). We did not consider the GHG credits from ethanol production and utilization in the calculation of farm-to refinery GHG since that requires extrapolations beyond the boundary of our analysis. However, if such GHG credits were applied in the analysis, our model would favor the intensification of management inputs and stover removal rates to maximize feedstock yields. This, in turn, would substantially decrease SOC stocks and might result in potential reductions of long term ecosystem health. Therefore, such analysis should be coupled with a SOC objective or apply a critical SOC threshold as an optimization constraint to ensure ecosystem sustainability. This threshold reflects the SOC concentration (%) below which many soil functions are significantly affected. The influences of SOC on ecosystem functioning have been studied for a wide range of soils, cropping types, and geographical regions (e.g., Greenland et al., 1975; Battiston et al., 1987; Bauer and Black, 1992; Körschens et al., 1998; Balesdent et al., 2000, Arshad and Martin, 2002; Loveland and Webb, 2003). These studies suggested that potential SOC threshold might range between 1% to 2% for agricultural soils. However, this threshold varies significantly depending on climate and soil physical properties.

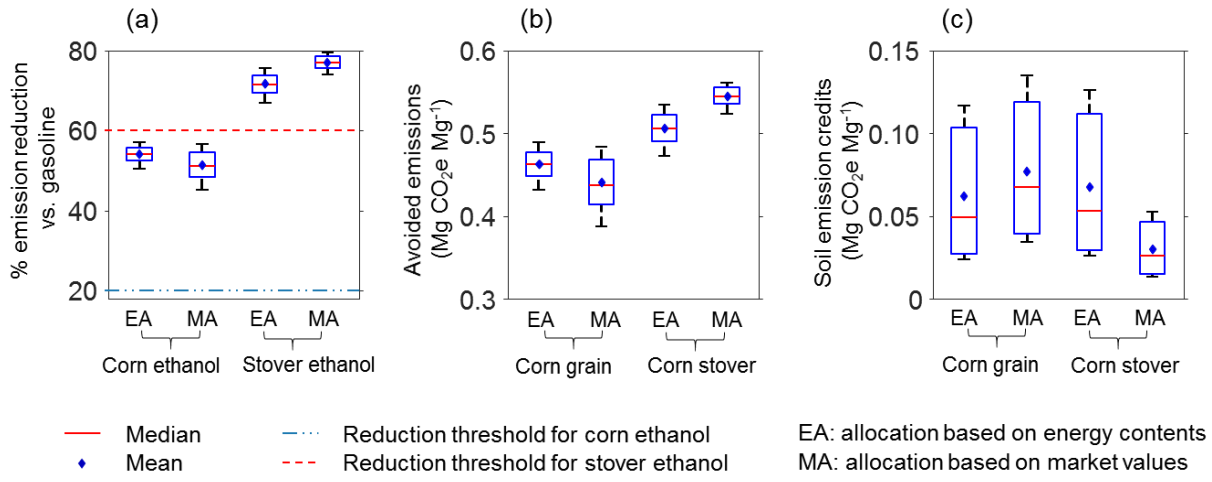


Fig. 4.7. Emission-related results for Pareto-optimal solutions. (a) Emissions reductions of grain- and stover-based ethanol as compared to conventional gasoline. (b) Avoided emissions from the production and utilization of ethanol attributed to corn grain and corn stover. (c) Emission credits due to increases in SOC stock as compared to the business-as-usual scenario attributed to corn grain and corn stover.

4.4.6 Effects of social costs on objective trade-offs

The effects of SC-CO_2 and SC-NL on the trade-offs between each pair of the design objectives along the Pareto frontiers (a.k.a., the borders of the Pareto surface) are illustrated in Fig. 4.8. We did not present such effects for data on the Pareto surface since they required a higher dimensional interpolation of the relationships among the weighting coefficient for production costs (w_1), SC-CO_2, and SC-NL. The generation of GHG vs. NL frontier required varying both SC-CO_2 and SC-NL in the integrated objective function, thus the SC-CO_2/SC-NL ratio was used in the x-axis (Fig. 4.8c). We limited SC-CO_2 between \$0 - \$5000 per Mg CO_2e, SC-NL between \$0 - \$500 per kg N, and the ratio between 0 and 600 to highlight the most dynamic variations in the design objective values. Our result indicated that the feedstock landscape design was more

sensitive to changes of SC-CO_2 between \$400 - \$800 per Mg CO_2e (Fig. 4.8a), changes of SC-NL below \$50 per kg N (Fig. 4.8b), and changes in SC-CO_2/SC-NL ratio below 350 (Fig. 4.8c).

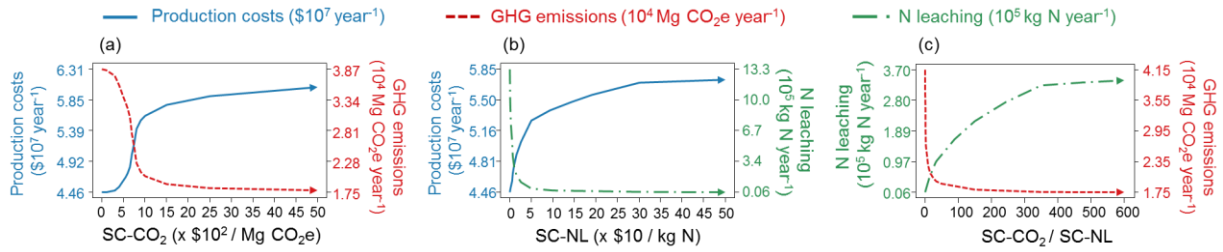


Fig. 4.8. Effects of social costs of carbon and N leaching on design objective trade-offs along Pareto frontiers.

We further examined the internalization of SC-CO_2 and SC-NL as a mean to estimate environmental benefits of GHG emissions and N leaching reduction for biofuel supply chains. The literature estimates of social costs of GHG emissions and N leaching varied greatly with SC-CO_2 ranging from \$39 - \$220 per Mg CO_2e (US EPA, 2016; Moore and Diaz, 2015; Tol, 2011; Rusu, 2012; Nordhaus, 2017) and SC-NL ranging from < \$1 - \$56 per kg N (Brink and Van Grinsven, 2011; Van Grinsven et al., 2013; Keeler et al., 2016; Compton et al., 2011). We compared the solutions for the lower bound (w_1 = 1, SC-CO_2 = \$39 Mg^{-1}, SC-NL = \$1 kg^{-1}) and upper bound (w_1 = 1, SC-CO_2 = \$220 Mg^{-1}, SC-NL = \$56 kg^{-1}) scenarios with that of the control scenario (minCost: w_1 = 1, SC-CO_2 = \$0 Mg^{-1}, SC-NL = \$0 kg^{-1}) (Fig. 4.5). Our result showed that the avoided damages of GHG emission reduction were \$0.42 × 10^5 $year^{-1}$ and \$9.76 × 10^5 $year^{-1}$ and those of N leaching reduction were \$3.89 × 10^5 $year^{-1}$ and \$716.37 × 10^5 $year^{-1}$ for lower and upper bound scenarios, respectively. These came with the corresponding trade-offs of \$3.21 × 10^5 $year^{-1}$ (0.7%) and \$80.32 × 10^5 $year^{-1}$ (18%) increases in production costs.

4.5 CONCLUSION

In this study, we demonstrated a high spatial resolution and multi-objective optimization of feedstock landscape design for a hybrid first- and second-generation biofuel supply chain. A discrete optimization approach was used in junction with the life cycle optimization framework to optimize system outcomes for the cultivation of corn grain and corn stover for ethanol production at Front Range Energy biorefinery (FRE), Windsor, Colorado, USA. The analysis was formulated as a tri-objective linear programming problem, including minimization of production costs, farm-to-refinery greenhouse gas emissions, and nitrogen leaching, subject to meeting feedstock demand for the ethanol production capacity of 1.51×10^8 L $year^{-1}$ (~ 40 million U.S. gallons per year). In particular, we addressed the questions of (1) how much land should be used for feedstock production; (2) where to produce the feedstocks; (3) how much of corn grain and corn stover to produce; and (4) what the optimum management practices (e.g., rates of N fertilization, irrigation, stover removal, and tillage intensity) are to apply for each corn field in the chosen landscape. Our results showed a broad win-lose Pareto surface among the three design objectives and various modulating patterns of the required feedstock area, management input investments, and grain-based and stover-based ethanol ratio associating with optimum solutions on the Pareto surface. Although our model relies heavily on ecosystem modeling expertise and advanced computational infrastructure, it could be a useful tool for solving many complex resource management problems that require detailed trade-off analysis of different objectives at high spatial resolutions. We also believe that the approach and modeling platform employed in this site-specific study would be equally applicable for other problems to support environmentally-conscious decision making in sustainable resource management for biofuel supply chains.

CHAPTER 5. SURROGATE-BASED MULTI-OBJECTIVE OPTIMIZATION OF MANAGEMENT OPTIONS FOR AGRICULTURAL LANDSCAPES USING ARTIFICIAL NEURAL NETWORK

5.1 SUMMARY

Simulation-based optimizations are often employed to inform decision makers about the trade-offs among competing objectives in agricultural production. However, computational limitations remain a major barrier to the implementation of these approaches for large-scale analyses. Surrogate-based optimization approaches offer potential solutions through the use of metamodeling techniques, which substitute the high-fidelity models with its simplified surrogate model to reduce simulation time. This study demonstrates the use of a surrogate-based optimization framework for large-scale and high-resolution landscape optimization, using irrigated corn production systems in Colorado, USA as a case study. An artificial neural network was employed to create a surrogate of the DayCent biogeochemical model. Our optimization considered the trade-offs among seven different objectives at farm and landscape scales, including farm profits, irrigation water use, corn grain, corn stover, soil organic carbon (SOC), greenhouse gas (GHG) emissions, and nitrogen leaching. Our results show that the surrogate captured greater than 99% of the variations in the DayCent's simulated outputs. Farm-level optimization improved farm profits, SOC, grain yield, and GHG emissions as compared to the 'business-as-usual' scenario. Landscape-level optimization suggested that the regional level of irrigation water use could be reduced without affecting both food production and environmental performance through careful land management decisions.

5.2 INTRODUCTION

Modern agricultural ecosystems need to be managed for many competing objectives such as food and fiber production, carbon storage, climate change mitigation, water quality improvement, and biodiversity conservation. Due to the resource constraints (e.g., land, water, nutrients, technology, and labor) in agricultural production, simultaneously achieving these management goals is not possible. Therefore, multi-objective optimization techniques can inform decision makers about possible trade-offs among multiple management objectives and the suite of management decisions they need to make to achieve a specific trade-off level. The multi-objective optimization typically generates a set of Pareto-optimal solutions (also called non-dominated solutions) where any improvement in one objective always impairs some others. Such optimizations are often carried out via simulation-based optimization approaches, which couple a single or a bundle of simulation models to quantify ecosystem responses with mathematical optimization algorithms to identify the trade-off surface (Wu et al., 2016). However, despite the significant improvement of computing power over recent decades, computational limitations remain a major barrier to the effective and systematic implementation of this approach, especially when process-based simulation models are employed for large-scale analyses in a high-resolution and spatially-explicit manner (Lardy et al., 2014).

The use of process-based ecosystem models provides an efficient and non-intrusive means to quantify ecosystem processes and their continuous interactions with decision-making (Nguyen et al., 2018). Simulation-based optimizations are often carried out via scenario analysis, for which a set of scenarios first needs to be formulated and fed to an ecosystem model for exploratory simulations. The results can then be optimized to choose the one that best fits the management goals. This approach helps improve decision-making but does not guarantee actual optima since

solutions are forced to some discrete options. Besides, the formulation of management scenarios and comparison of the simulated results can be time-consuming, laborious, and highly expertise-driven (Seppelt and Voinov, 2002).

Another disadvantage of simulation-based optimizations is that using ecosystem models for large-scale simulations often require high computational power with clusters of superpower machines. This computational burden is magnified when finer spatial resolutions (e.g. field and sub-field scales) and more detailed changes in management practices (e.g. fertilization, irrigation, and tillage) are considered. In addition, the consideration of many decision variables for robust and operational optimization in agricultural resource management can result in very large solution search spaces, which require more time and resources to find solutions. These challenges make simulation-based optimization approaches extremely costly and infeasible in several circumstances. Consequently, the use of process-based simulation models for detailed large-scale landscape optimization is impractical in most institutional contexts that require timely analysis to support rational decision making (MA, 2005).

One of the potential solutions to overcome the inefficiency of process-based ecosystem models in optimization frameworks while maintaining acceptable prediction accuracies is to employ the surrogate-based optimization approaches (Ratto et al., 2012). These approaches use meta-modeling techniques to substitute the high-fidelity models within an optimization framework with its cost-effective simplified surrogate model to reduce simulation time, expertise requirement, and data storage (Wu et al., 2016). A surrogate is a model of the detailed model that highlights properties of the detailed model itself. Meta-modeling is a data-driven approach, which involves fitting a response surface to approximate the relationship between multiple input variables and a target output. The replacement of complex models with their surrogates brings

many practical benefits. First, the operation of surrogate models is highly cost-effective due to the reduction in the simulation time and data storage requirement. Second, they can easily be applied across different spatial and/or temporal scales with an acceptable degree of accuracy, providing the data corresponding to the new system parameterization are available. Third, with higher flexibility and adaptability, surrogate models can be integrated into other processes and simulation platforms to achieve a variety of simulation goals such as prediction, optimization, validation, and sensitivity analysis (Simpson et al., 2001).

Although surrogate-based optimization has been widely used in mechanical engineering for a long time (Meckesheimer et al., 2002), its application in ecosystem research is still very recent (Castelletti et al. 2012). Surrogate models have been employed in various studies such as land use and nitrogen fertilizer use distribution (Audsley et al., 2008), pesticide leaching (Stenemo et al., 2007), air quality (Carnevale et al., 2012), N_2O and nitrogen leaching (Villa-Vialaneix et al., 2012), and groundwater (Roy et al., 2016). However, incorporating these models into optimization frameworks to support decision making is less common, with most applications to date dealing with water resource management such as conjunctive surface- and ground-water use (Wu et al., 2016) and coastal aquifer management (Kourakos and Mantoglou, 2013; Song et al., 2018). Despite the on-going development of surrogate-based optimization approaches, to our knowledge, there have been no studies that have applied these approaches for large-scale agricultural landscape optimization, considering multiple stakeholders and ecosystem services at high spatial and temporal resolutions.

This study aims to demonstrate the use of a surrogate-based optimization framework for spatially-explicit trade-off analyses in agricultural landscapes to support resource management decision-making. We present our framework with a case study on optimization of the landscape

design of management practices for irrigated corn production systems in Colorado, USA. The optimization considered the trade-offs among farm profits, irrigation water use, and several ecosystem services at a high spatial resolution (field scale) and detailed consideration of historical land use (dating back to the 1880s).

5.3 CASE STUDY AND METHOD

5.3.1 Study site and problem statement

Our study focused on the irrigated corn growing area in eastern Colorado, USA (Fig. 5.1). The whole area is 298,897 hectare (ha) with a majority of fine-loamy soils. Like many other agricultural states, agriculture in Colorado is subjected to the trade-offs among food and biomass production, greenhouse gas emissions, water pollution from excessive fertilizer run-off (Strange et al., 1999). The allocation of limited water resources to meet the needs of agricultural production and municipal and industrial development is also a challenge to the state (Water Conservation Board, 2010). In addition, the potential use of crop residues for ligno-cellulosic biofuels production in the region introduces another trade-off between the residue harvest and the long-term soil health (Wilhelm et al., 2007). The complexity and the multidimensional nature of these trade-offs necessitate multi-objective optimizations of land management decisions at different levels (farm and landscape) to improve the efficiency of resource allocation and minimize negative impacts on the environment while meeting future demands.

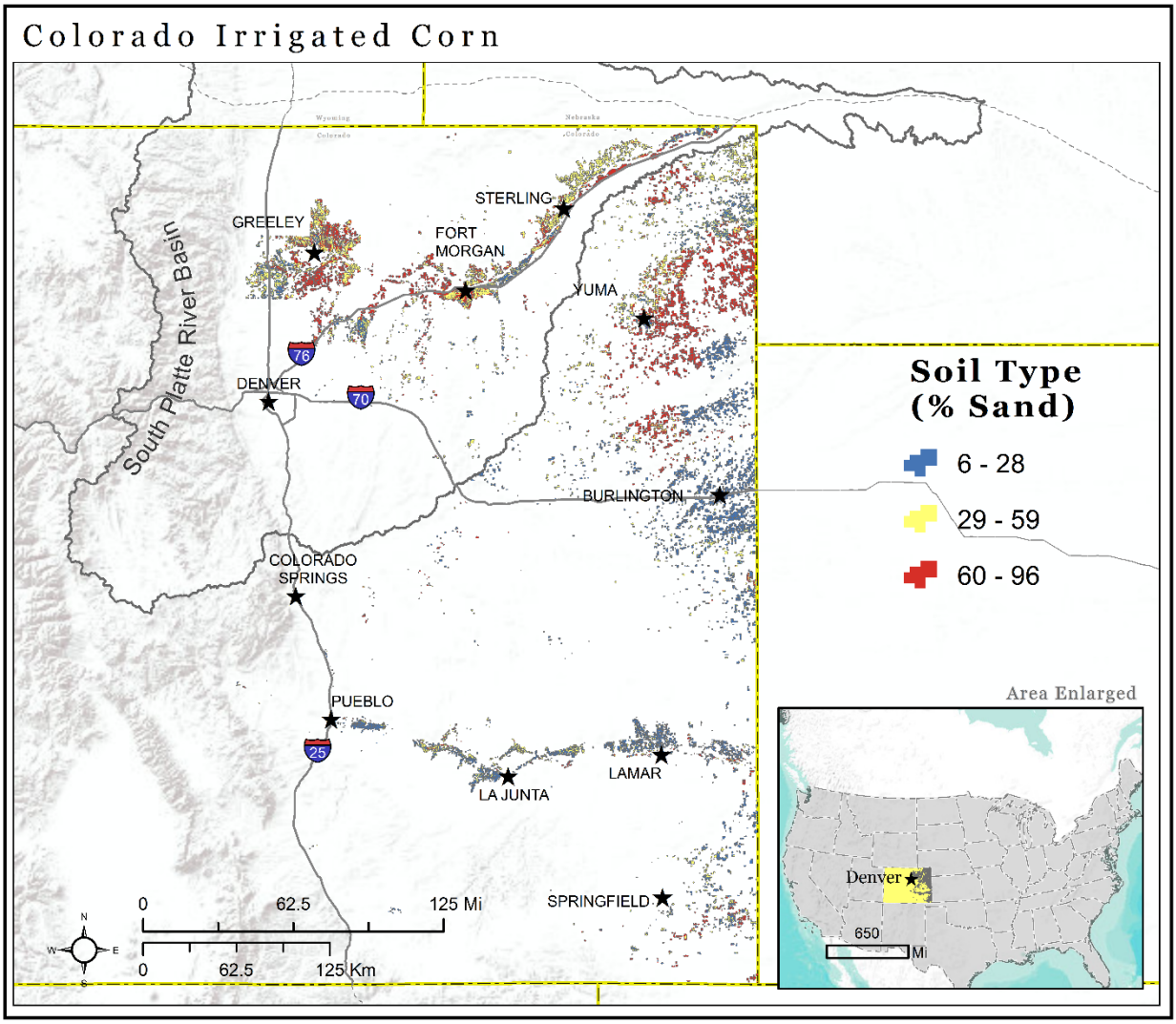


Fig. 5.1. Irrigated corn land units with the percent sand content of the soils within Colorado state.

The main objective of this case study was to quantify the trade-offs among seven management objectives for irrigated corn production systems in eastern Colorado. These objectives consisted of farm profits, irrigation water use, and five ecosystem services including food and biomass production, soil carbon storage, water quality, and climate regulation. We also aimed to identify the optimal spatial configuration of management practices for the study region corresponding to different trade-off levels. The optimization considered farmers' objective of maximizing farm profits and a regional planner's multi-objectives for maximization of ecosystem services. To

quantify the ecosystem services, we used corn grain, corn stover, soil organic carbon (SOC), greenhouse gas (GHG) emissions, and nitrogen (N) leaching as indicators representing food and biomass production, ecosystem carbon storage, climate regulation, and water quality, respectively. For solvability purpose, we structured our analysis as a two-level optimization problem.

At the farm level, we optimized management practices to maximize farm's profits, subject to multiple relative SOC constraints. This constraint defines the percentage increase in SOC due to changes in management practices relative to a 'business-as-usual' scenario and was used as a proxy for long-term soil health. The management practice decision variables considered in our analysis were N fertilization (kg N ha^{-1}), irrigation (cm $year^{-1}$), and residue removal (%) rates. Since our previous study (Nguyen et al., in review) of irrigated corn in the region indicated that higher tillage intensity negatively impacted SOC stock while insignificantly affecting farm's profit and the yields of corn grain and stover, no-till was used for all corn fields in this optimization. Artificial neural networks were employed via a metamodeling technique to rapidly predict farm's SOC, corn grain and stover yields as functions of farm management practices and other site-specific variables. The predicted yields and SOC were then used to compute the objective (farm profits) and constraint (percentage change in SOC level) functions, respectively. Farm-level optimization was formulated as a continuous, bounded, and non-linear programming problem, and was iteratively solved for different levels of relative SOC constraint using Fmincon solver in the Optimization Toolbox of MATLAB software (MATLAB and Optimization Toolbox, 2017).

At the landscape level, we carried out a combinatorial multi-objective optimization based on the outcomes from farm-level optimizations. For this, the optimized farm-level management

practices were simulated with the Agricultural Ecosystem Service Optimization system to derive the dataset for landscape-level optimization. We then optimized the farm-level relative SOC constraint to maximize the production of corn stover while minimizing GHG emissions and N leaching. The landscape-level optimization was subjected to various constraints in corn grain production, SOC accumulation, and irrigation water use. Landscape-level optimization was carried out using the ε-constraint methods (Haimes et al., 1971) for multi-objective problems. An overview of our surrogate-optimization framework is presented in Fig. 5.2.

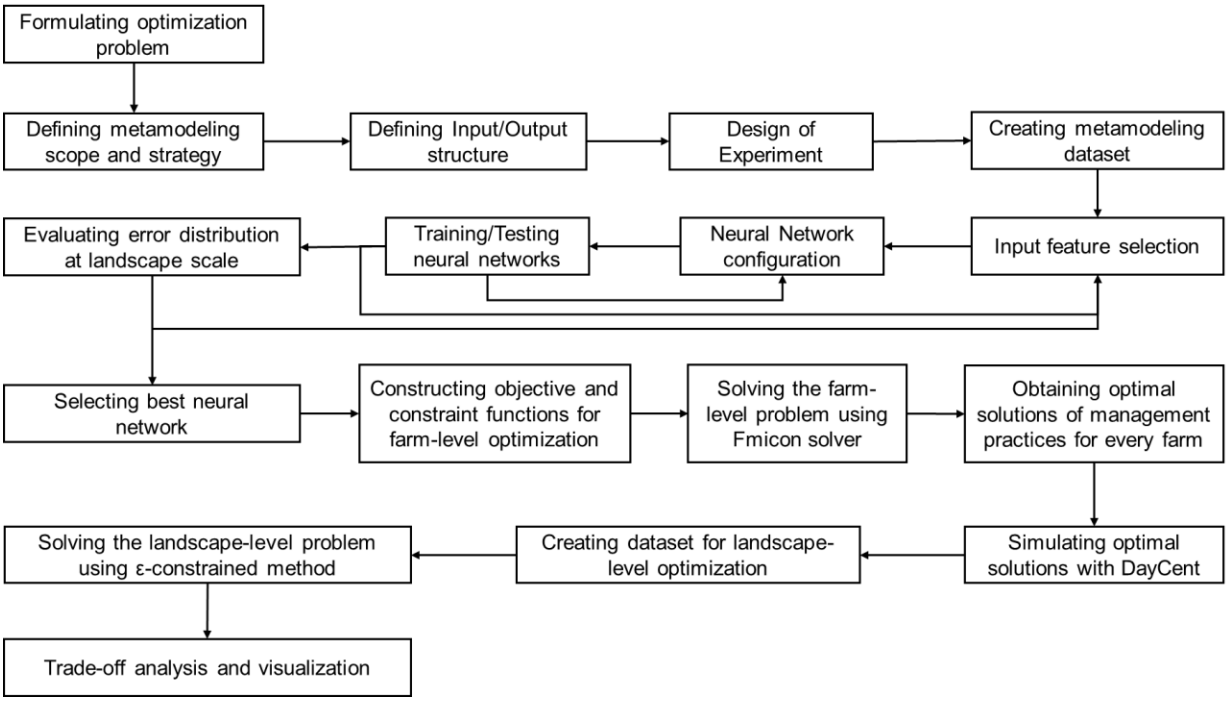


Fig. 5.2. Overview of the surrogate-based optimization framework employed by our study.

5.3.2 Metamodeling for farm-level optimization

5.3.2.1 Agricultural Ecosystem Service Optimization tool

The Agricultural Ecosystem Service Optimization (Ag-EcoSOpt) tool (Nguyen et al., 2018) was used to simulate corn grain and stover yields and farm-level SOC to generate the input dataset

for metamodeling. The Ag-EcoSOpt tool is comprised of the DayCent biogeochemical model (Del Grosso et al., 2002), a database of historical land use for U.S. cropland (Ogle et al., 2010), a geospatial component for spatially-explicit processing of inputs and outputs, and an optimization component. Several optimization algorithms such as min-max (Hwang et al., 1980), weighted sum (Fishburn, 1967), and ε-constraint (Haimes et al., 1971) were included in Ag-EcoSOpt to optimize DayCent's simulation outputs for land use and land management decision making.

The DayCent model is a daily time-step process-based model that simulates biogeochemical flows of carbon, nitrogen, and water among the atmosphere, vegetation, and soil in many different ecosystems. DayCent simulates plant growth as a function of nutrient availability, soil water and temperature, shading, vegetation type, and plant phenology (Metherell et al., 1993). Grain yield and harvested residue biomass are calculated from the aboveground carbon based on a crop-specific harvest index and a user-defined residue removal rate, respectively. Soil organic matter (SOM) dynamics are simulated for surface litter pools and the top soil layer (0-20 cm), and are represented as functions of substrate availability, substrate quality (lignin content, C/N ratio), soil texture, temperature, water availability, and tillage intensity. The model has been previously used to predict crop yields, soil carbon dynamics, and greenhouse gas emissions in a wide variety of US agricultural systems (e.g., Del Grosso et al., 2008a; Del Grosso et al., 2008b; Kim et al., 2009; David et al., 2008; Ogle et al., 2010; US EPA, 2015; Zhang, 2016).

5.3.2.2 Artificial neural networks

There are many widely used metamodeling methods such as kriging, polynomial functions, artificial neural networks (ANNs), radial basis functions, k-nearest neighbor, and support vector machines. Of these methods, ANNs have been shown to be more suitable for CPU-intensive applications like optimization because of their high accuracy and low prediction time (i.e., the

time needed to obtain new predictions) (Villa-Vialaneix et al., 2012; Roy et al., 2016). Therefore, in this analysis we employed ANNs to metamodel DayCent using the input dataset generated by Ag-EcoSOpt.

ANNs (Haykin, 1998) is a machine learning technique that mimics biological nervous systems, such as the brain, to identify patterns (classification) or to fit functions (regression). An elementary artificial neuron consists of five basic elements, including inputs, input weights, an adder, an activation (a.k.a., transfer) function, and an output. The input values are multiplied by their weights and the weighted sum computed by the adder is fed to the activation function to generate the output, which is propagated to other units in the ANN. The activation function limits the range of the output to some finite value and typically falls into one of three categories including linear, threshold, and sigmoid. ANNs learn by adapting their input weights to identify the underlying input-output relationships from a collection of training examples within a specific task domain (Haykin, 1998). The multilayer feedforward neural networks, a.k.a., multilayer perceptrons (MLP), are the most widely studied and used neural networks in practice, especially for supervised learning problems. A typical MLP consists of at least three layers of neurons, including an input layer, one or more hidden layers of sigmoid neurons, and an output layer of linear neurons (Fig. 5.3) (Kaastra and Boyd, 1995). MLP training, i.e., the network's weight optimization, is commonly carried out using a gradient-descent based algorithm known as backpropagation (Rumelhart et al., 1986). Many studies have shown that MLP of one hidden layer with adequate number of neurons can be trained to approximate any measurable function to any desired degree of accuracy (Hornik et al., 1989; Funahashi, 1989). As a result, our analysis used one-hidden-layer MLP and varied the number of neurons in the hidden layer, via a sensitivity analysis, to identify the best network.

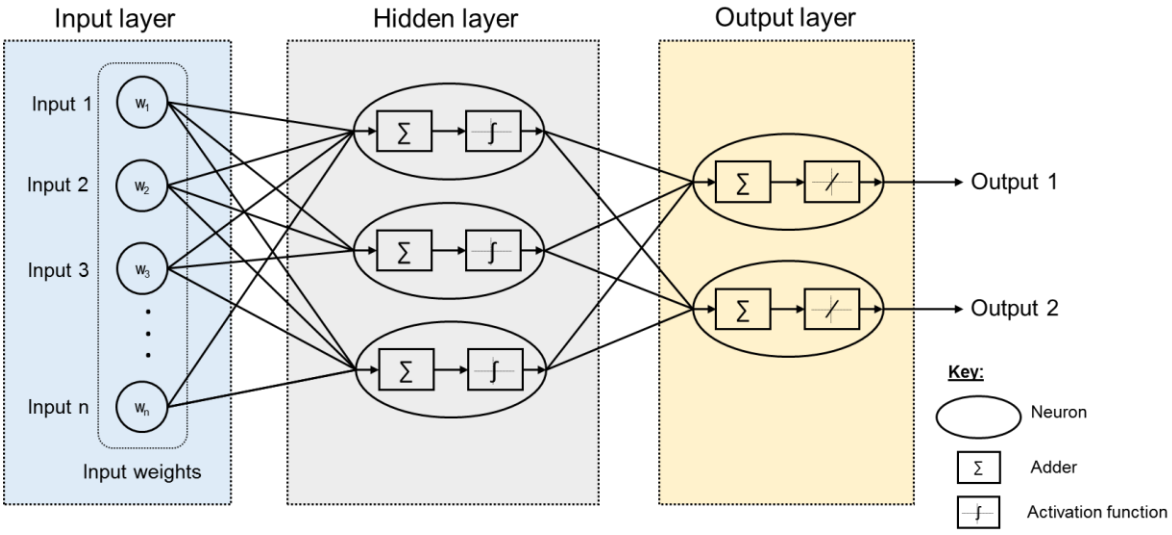


Fig. 5.3. Schematic presentation of a multilayer perceptron

5.3.2.3 Scope of metamodeling

The high computation time of Ag-EcoSOpt's landscape simulations could mainly be attributed to DayCent model initialization and forward simulations. Prior to any forward simulation, DayCent needs to be initialized for all farms in the landscape. The model initialization consists of three simulation stages including spin-up, base, and 'business as usual' (BAU). The spin-up simulation (5,000 – 10,000 years) defines the natural steady-state of the land prior to any disturbance from human activities. The base simulation (hundreds of years), extended from this equilibrium state, reflects the land condition in the distant past. This historical simulation is followed by a BAU simulation (<100 years), which represents the land condition under the current management regime, and is used to compare with forward simulations into the future. The forward simulations, extended from the BAU condition, model different predefined scenarios of changes in land use and/or land management to support decision making. This sequential model simulation is important for capturing the long-term and cumulative effect of management practices, soil properties, and weather data on carbon and nitrogen dynamics.

Although the model initialization is imperative, it is repetitive for an area under different analyses. Therefore, we carried out initialization simulations for the study region and stored the initialized results in our database as the initial conditions for the forward simulations. Metamodeling technique was employed to further reduce the computational burden of the forward simulations and thus increased the efficacy of large-scale and complex landscape optimizations. The scope of our metamodeling is to train an MLP to quickly and accurately predict the DayCent's modeled SOC, corn grain and stover yields at farm level under the forward simulations.

5.3.2.4 Defining inputs and outputs structure

In this analysis, we used a time frame of 30 years for the forward simulations. The time frame was selected to account for the effects of the variation in one cycle of weather data (1980 – 2010) from the North American Regional Reanalysis (NARR) climate database used for our model simulations (Mesinger et al., 2006). It was also sufficient for soil organic carbon (SOC) to approach a new equilibrium after switching management practices from the BAU scenario for a fair comparison of SOC and a thorough accounting of annual CO_2 emissions. The effects of climate change and increases in atmospheric CO_2 concentration were not simulated in our study. The SOC stock at the end of the forward simulation period and the 30-year-average corn grain and stover yields were used as the output variables for metamodeling. These variables are efficient to capture both the cumulative effects of changing management practices on SOC dynamics and the seasonal variations of yields. SOC was reported in Mg of carbon per hectare while corn grain and stover were reported in Mg of dry weight biomass per hectare per assuming 43.5% carbon content and moisture content of 15.5% for corn grain and 20% for harvested stover (Gesch et al., 2010).

The DayCent model's inputs were used as the input variables for metamodeling. The primary inputs of DayCent are daily maximum and minimum air temperature and precipitation, soil properties for each horizon in the soil profile, soil (rooting) depth, land cover/use data (e.g., vegetation types, land use history), and management practices (e.g., irrigation, tillage, fertilization). To reduce the number of input variables for faster training of MLP, soil property-related input variables such as bulk density and textures were averaged across all soil layers. Daily min and max temperatures and daily precipitation were also aggregated into growing-season values and then averaged over the forward simulation timeframe (30 years) before using as input for MLP training. The soil organic carbon and soil nitrogen at the end of the BAU simulations were used as the input variables that describe the historical land management. The list of the input variables used for our MLP metamodel building is given in the Appendix C1.

5.3.2.5 Design of experiment

The metamodel input variables were divided into two groups for sampling purposes including site-specific and management-specific input variables. Site-specific input variables describe the intrinsic properties of a corn field such as temperature and precipitation, soil texture, soil depth, and historic conditions. These variables, related to a specific field, do not change during the forward simulations of DayCent and thus were sampled based on their spatial locations. Management-specific input variables refer to N fertilizer, irrigation, and stover removal rates, which were modulated during DayCent forward simulations to measure the response in SOC, corn grain and stover yields. They were sampled based on experimental design in each corn field.

The site-specific input data were acquired from several spatial databases including Natural Resource Conservation Service Soil Survey Geographic database (SSURGO, NRCS-USDA,

2014), the NARR climate database (Mesinger et al., 2006), U.S irrigated land database (Pervez and Brown, 2010), and the National Agricultural Statistics Service Crop Data Layer (NASS-CDL, 2016). The shapefiles of these databases were then intersected to create input polygons of various sizes (>1 ha) representing 35,365 virtual irrigated corn fields in the study site (Fig. 5.1), including 2707 unique combinations of soil and weather data (DayCent input strata). Only these input strata were simulated with DayCent to create dataset for metamodeling.

The DayCent input strata were initialized with the sequential simulation procedure used in the Inventory of U.S. Greenhouse Gas Emissions and Sinks (US EPA, 2015) and the COMET-Farm^{TM} system (http://cometfarm.nrel.colostate.edu/). Each stratum was first spun up with native grassland condition for 5000 years and then extended with the base simulations using the historical agricultural land use assumptions for Major Land Resource Area (MLRA) up to 1979 (Ogle et al., 2010). Finally, they were extended to 2017 using the state-average BAU management practices reported by the USDA Agricultural Resource Management Survey for Colorado's irrigated corn crop, which were 170 kg of nitrogen per hectare, 34 cm of irrigated water per growing season, 75% residue removal, and conventional tillage (ERS-ARMS, 2010).

To insure balanced sampling, N fertilizer, irrigation, and residue removal rates were first stratified into lower and upper ranges based on the BAU management practice levels and their corresponding lower and upper bounds. These bounds were chosen to reflect the possible ranges for irrigated, continuous corn in the region (Halvorson et al., 2006), including 0 – 300 kg ha^{-1} for N fertilization, 0 – 70 cm year^{-1} for irrigation, and 0 – 90% for stover removal. We then sampled from these ranges assuming uniform distributions of the management-specific input variables. A 3 × 2 factorial design (i.e., 3 management practices each at 2 levels) was then employed for each DayCent input strata in the landscape. The sampling yielded 21,656 samples (pairs of

inputs/output) for the metamodeling. All DayCent simulations were conducted in parallel on an 18-node, 216-processor cluster computing system at the Natural Resource Ecology Laboratory at Colorado State University.

5.3.2.6 Developing the metamodel

There are four main steps in developing an ANN metamodel including preprocessing data, configuring, training, and testing the network (Maier et al., 2010). First, we normalized the input and output variables between [-1, 1]. The normalization of input data increases the network training efficacy and prevents the saturation effect of large input values on the activation functions (Li et al., 2000). The ranges for data normalization depend on the activation functions used to configure neural networks. For this analysis, we used hyperbolic tangent sigmoid and linear activation functions for the hidden and output layers, respectively. The preprocessed data were then randomly partitioned (without replacement) into three subsets including training, validation, and test sets. The training set was used to optimize network's weights and biases; the validation set was used to prevent overfitting during network training ("early stopping") (Prechelt, 2012); and the test set was used to evaluate the performance of the trained network. To reduce the training and prediction time of the metamodel, we developed a nested MLP to predict three output variables simultaneously.

There is no apparent rule for MLP configuration for a specific problem, and thus "trial and error" approach is often employed in practice (Maier and Dandy, 2000). For our study, we conducted a sensitivity analysis for the number of input variables, training ratio, and number of neurons in the hidden layer to identify the best configuration of the MLP. First, we ran the relative importance analysis (Johnson and Lebreton, 2004) on the input variables using a simple configuration of MLP with ten hidden neurons. The input variables were then added to the MLP training

sequence in the descending order of their relative importance. The relative importance refers to the proportional contribution each input variable makes in explaining the variation in the output variables, considering both its direct effect (i.e., its correlation with the outputs) and its effect when combined with the other input variables (Johnson and Lebreton, 2004). For each input variable, importance was calculated as the difference between the accuracy of the full model (i.e., including all input variables) and the mean accuracy of all submodels that excluded the input variable of interest. The values were normalized against each other so that they sum up to 100%. For our study, the relative importance analysis was carried out via a sensitivity analysis, which was explained in detail in IBM Corp (2015). The training ratio defines the percentage of the input dataset that is used to train the MLP. We varied the training ratio from 10% to 70% incrementing by a step of 20%. The test set was kept at 20% of the total sample size in all cases. The number of neurons was varied between 5 and 60 with an incremental step of 5. The sensitivity analysis for MLP configuration was automated and incorporated into Ag-EcoSOpt for the sake of future studies.

The configured MLPs were initialized with random weights sampled from a Gaussian distribution with a mean of 0 and a standard deviation of 0.01 and were trained using Levenberg-Marquardt backpropagation (Marquardt, 1963) for 500 epochs (training iterations) using Neural Network Toolbox in MATLAB software (MATLAB and Neural Network Toolbox, 2017). Other parameters of the training algorithm were kept at their default values. Our MLP training goal was to minimize the mean squared errors (MSE) of the training set ($MSE_{train} = 0$). The training stopped if one of the following conditions was met: (1) the number of training iteration exceeds 500; (2) the MSE of training set equates zero; (3) the magnitude of the error gradient is less than 10^{-7}; and (4) MSE of the validation set has increased more than 6 training iterations

since the last time it decreased (early stopping) (Prechelt, 2012). The performance of the trained MLP was evaluated using the test set based on the coefficient of determination (R^2) and MSE. One of the drawbacks of the backpropagation training algorithm is that the network might be trapped in some local minima especially for non-smooth error surfaces (Maier et al., 2010). To rectify this problem, each MLP configuration was trained 10 times using different random weight initialization and the performance means were compared. The final surrogate model was selected based on a balance among metamodel accuracy, training and prediction time.

5.3.2.7 Testing the spatial distribution of prediction errors

It is important that surrogate models are tested based on their domains of use (Marie and Simioni, 2014). Since our objective was to use the selected surrogate model to construct the objective and constraint functions for the farm-level optimization of management practices for irrigated corn in Colorado, we calculated the root mean squared errors (RMSE) between the surrogate model predictions and the DayCent's outputs for each corn farm. The RMSE was computed based on 216 random samples of management practices per farm. The farm-level RMSE were then visualized for the whole landscape via maps to identify the spatial pattern of metamodeling uncertainty as well as locations of extreme errors.

5.3.3 Formulation of the optimization problems

For clarity, hereafter we refer to the farm-related metrics with a prefix "farm" (e.g., farm SOC and farm GHG) while the landscape sum of a farm-related metric (i.e., sum of all farms in the landscape) will be referred to only by the metric's name (e.g., SOC and GHG). In the case of corn stover and corn grain, the farm-level metrics will be denoted with the suffix "yield" (e.g., corn stover yield and corn grain yield) while their landscape summation will be denoted as corn stover and corn grain (with no suffix), respectively. For constraint function, we differentiate

between the constraint metric (left side of constraint function) and constraint value (right side of constraint function). For example, in equation (11), corn grain (i.e., $\sum_i^N \sum_j^M \text{Grain}_{i,j} * x_{i,j}$) is the constraint metric and A is the constraint value. Farm profits were calculated as the difference between revenues from corn grain and stover and the corresponding farm production costs. Farm GHG emissions were computed as the sum of all soil-related and on-farm machinery emissions. The calculation of all the metrics and their corresponding data are reported in Appendix C2 using the method published in our previous papers (Nguyen et al., 2018; Nguyen et al., in review).

Our optimization problem can be formulated as follows, with detailed nomenclature presented in Table 5.1:

Farm-level optimization: This is a bounded and non-linear programming problem

$\forall i \in N$:

$\forall j \in M$: maximizing $f_p(F_i, I_i, R_i)$ (Eq. 5.1)

Subject to (s.t):

$F_{LB} \leq F_i \leq F_{UB}$ (Eq. 5.2)

$I_{LB} \leq I_i \leq I_{UB}$ (Eq. 5.3)

$R_{LB} \leq R_i \leq R_{UB}$ (Eq. 5.4)

$\frac{f_s(F_i,I_i,R_i) - \text{bauSOC}_i}{\text{bauSOC}_i} \geq \text{dSOC}_j$ (Eq. 5.5)

Landscape-level optimization:

Maximize $\sum_i^N \sum_j^M \text{Stover}_{i,j} * x_{i,j}$ (Eq. 5.6)

Minimize $\sum_i^N \sum_j^M \text{GHG}_{i,j} * x_{i,j}$ (Eq. 5.7)

Minimize $\sum_i^N \sum_j^M \text{NL}_{i,j} * x_{i,j}$ (Eq. 5.8)

Subject to:

$x_{i,j} \in \{0, \ldots, 1\}$ (Eq. 5.9)

$\sum_j^M x_{i,j} \leq 1, \quad \forall i \in N$ (Eq. 5.10)

$\sum_i^N \sum_j^M \text{Grain}_{i,j} * x_{i,j} \geq A$ (Eq. 5.11)

$$\sum_{i}^{N}\sum_{j}^{M} SOC_{i,j} * x_{i,j} \geq B \quad \text{(Eq. 5.12)}$$

$$\sum_{i}^{N}\sum_{j}^{M} Irrigation_{i,j} * x_{i,j} \leq C \quad \text{(Eq. 5.13)}$$

Table 5.1. The nomenclature of the optimization problem

Name	Description
	Indices
i	Index for corn fields in the landscape
j	Index for farm's relative SOC constraint
	Set sizes
N	Number of corn fields in the landscape (N = 38,299)
M	Number of considered relative SOC constraints (M = 6)
	Parameters
Flb	Lower bound of N fertilizer (Flb = 0 kg ha^{-1})
Fub	Upper bound of N fertilizer (Fub = 300 kg ha^{-1})
Ilb	Lower bound of irrigation (Ilb = 0 cm $year^{-1}$)
Iub	Upper bound of irrigation (Iub = 70 cm $year^{-1}$)
Rlb	Lower bound of residue removal rate (Rlb = 0 %)
Rub	Upper bound of residue removal rate (Rub = 90%)
$bauSOC_i$	SOC level of the 'business-as-usual' scenario in farm i (Mg C ha^{-1})
$dSOC_j$	Relative SOC constraint $dSOC_j$ for farm i, $dSOC_j \in \{0\%, 10\%, 20\%, 30\%, 40\%, 50\%\}$
$Stover_{i,j}$	Corn stover produced in farm i under relative SOC constraint $dSOC_j$ (Mg biomass $year^{-1}$)
$GHG_{i,j}$	GHG emissions of farm i under relative SOC constraint $dSOC_j$ (Mg CO_2e $year^{-1}$)
$NL_{i,j}$	N leaching of farm i under relative SOC constraint $dSOC_j$ (kg N $year^{-1}$)
$Grain_{i,j}$	Corn grain produced in farm i under relative SOC constraint $dSOC_j$ (Mg $year^{-1}$)
$SOC_{i,j}$	SOC of farm i under relative SOC constraint $dSOC_j$ (Mg C)
$Irrigation_{i,j}$	Irrigation water used by farm i under relative SOC constraint $dSOC_j$ (m^3 $year^{-1}$)
A	Landscape constraint of corn grain production based on NASS's 2013 census for irrigated corn in Colorado (A = 3.34 x 10^6 Mg $year^{-1}$)
B	Landscape constraints for SOC. B is the total SOC under the 'business-as-usual' scenario with tillage changed from conventional to no-till (BAU-NT) (B = 8.19 x 10^6 Mg C $year^{-1}$)
C	Landscape constraints for irrigation water use based on NASS's 2013 census for irrigated corn in Colorado (C = 1.50 x 10^9 m^3 $year^{-1}$)
	Non-linear functions

$f_p(F_i, I_i, R_i)$	Function measures profits of farm i corresponded to the management practices F_i, I_i, and H_i ($ ha^{-1} year^{-1}).
$f_s(F_i, I_i, R_i)$	Function measures SOC of farm i corresponded to the management practices F_i, I_i, and H_i (Mg C ha^{-1})
	Continuous variables
F_i	N fertilizer rate of farm i (kg ha^{-1})
I_i	Irrigation rate of farm i (cm year^{-1})
R_i	Residue removal rate of farm i (%)
	Binary variables
$x_{i,j}$	Fraction of field i under relative SOC constraint $dSOC_j$ is chosen in the final landscape.

The objective (Eq. 5.1) represents the maximization of farm profits, which is a non-linear function based on farm management practices. The constraints (Eq. 5.2), (Eq. 5.3), and (Eq. 5.4) depict the upper and lower bounds for farm N fertilizer, irrigation, and residue removal rates, respectively. The constraint (Eq. 5.5) imposed a certain level of relative SOC constraint on the farm optimization. The relative SOC constraint for farm i is defined as the percentage increase in SOC due to changes in management practices, relative to the 'business-as-usual' SOC level of that particular farm. To examine the trade-off between farm profits and farm SOC as well as to generate fine-scaled scenarios for the landscape optimization, we varied the farm's relative SOC constraint value between 0 - 50%, incrementing by a step of 10%. In other words, each farm in the landscape was optimized six times corresponding to six levels of the relative SOC constraint.

The objectives (Eq. 5.6), (Eq. 5.7), and (Eq. 5.8) define the simultaneous maximization of corn stover and minimizations of GHG emissions and N leaching at the landscape level, respectively. The constraints (Eq. 5.9) and (Eq. 5.10) allow the selection of multiple relative SOC constraints for each corn farm as long as the sum of their corresponding areal fractions does not exceed 1. The constraints (Eq. 5.11), (Eq. 5.12), and (Eq. 5.13) require the optimization to at least maintain

the empirical production levels of corn grain and SOC while keeping the irrigation water use below an empirical amount. The most recent census of U.S agriculture by the National Agricultural Statistics Service (NASS) in 2012 reported an average of 11.17 Mg ha^{-1} and 5029 m^3 ha^{-1} $year^{-1}$ for corn grain yield and farm irrigation for irrigated corn in Colorado, respectively (NASS, 2012). These corresponded to about 3.34 x 10^6 Mg $year^{-1}$ of corn grain and 1.50 x 10^9 m^3 $year^{-1}$ of irrigation water use on the landscape level. For SOC, since empirical data were not available at state level, we used the simulated landscape sum of SOC under the 'business-as-usual' scenario with tillage adjusted from conventional to no-till. Switching the tillage under the 'business-as-usual' scenario was consistent with the tillage assumption of the optimization while placing a more conservative constraint on SOC for sustainability. These empirical/simulated values were set as the base constraint values for the landscape multi-objective optimization.

We also varied the landscape constraint values to examine how improvements in corn grain, SOC, and irrigation water saving affected the pareto trade-offs among the landscape objectives (i.e., corn stover, GHG emissions, and N leaching). For this, we first carried out different single-objective optimizations (both maximization and minimization) of the landscape constraint metrics (i.e., corn grain, SOC, irrigation water use). These are scenarios where one of the constraint metrics was turned into an objective function while the other metrics remained as constraints using their base values (Table 5.2). The results from the minimization and maximization of a certain constraint metric revealed its feasible value range (FVR) (i.e., [min, max]). These ranges were first used to verify the base constraint values (A, B, C). If a base constraint value lied beyond the FVR of its corresponding constraint metric (i.e., no feasible solution), it would be reset to the mean of the feasible value range before carrying out the landscape multi-objective optimization. The ranges between the base constraint values and either

the maximum values of corn grain and SOC or the minimum value of irrigation water use were defined as the 'Constraint Improvement Region' (CIR) (see details in Appendix C6). This was because setting the constraint values along these ranges resulted in increases in corn grain and SOC and decreases in irrigation water use. To examine the effects of improvements in constraint metrics on the landscape multi-objective optimization, we separately varied each constraint value along its CIR and observed the changes in the resulting pareto frontiers.

Table 5.2. Objective and constraint functions of the single-objective optimizations of corn grain, SOC, and irrigation water use

Objective function		Constraint functions
$\sum_i^N \sum_j^M \text{Grain}_{i,j} * x_{i,j}$	s.t.	$x_{i,j} \in \{0, \ldots, 1\}$ $\sum_j^M x_{i,j} \leq 1, \quad \forall i \in N$ $\sum_i^N \sum_j^M \text{SOC}_{i,j} * x_{i,j} \geq B$ $\sum_i^N \sum_j^M \text{Irrigation}_{i,j} * x_{i,j} \leq C$
$\sum_i^N \sum_j^M \text{SOC}_{i,j} * x_{i,j}$	s.t.	$x_{i,j} \in \{0, \ldots, 1\}$ $\sum_j^M x_{i,j} \leq 1, \quad \forall i \in N$ $\sum_i^N \sum_j^M \text{Grain}_{i,j} * x_{i,j} \geq A$ $\sum_i^N \sum_j^M \text{Irrigation}_{i,j} * x_{i,j} \leq C$
$\sum_i^N \sum_j^M \text{Irrigation}_{i,j} * x_{i,j}$	s.t.	$x_{i,j} \in \{0, \ldots, 1\}$ $\sum_j^M x_{i,j} \leq 1, \quad \forall i \in N$ $\sum_i^N \sum_j^M \text{Grain}_{i,j} * x_{i,j} \geq A$ $\sum_i^N \sum_j^M \text{SOC}_{i,j} * x_{i,j} \geq B$

Note: Each objective function was used for both maximization and minimization

5.4 RESULTS

5.4.1 Model accuracy

Our results indicated high accuracy of the surrogate models with appropriate configuration and training (Fig. 5.4). The minimum Root Mean Squared Error (RMSE) of test sets for different MLP configurations varied from 0.4 – 4.1 Mg C ha^{-1} for SOC, 0.3 – 3.3 Mg ha^{-1} for corn grain

yield, 0.2 – 2.4 Mg ha^{-1} for corn stover yield (Fig. 5.4a). These corresponded to the test-set R^2 of 73.6% – 99.5%, 1.7% – 99.4%, and 0.8% – 99.3%, respectively (Fig. 5.4b). The performance of the surrogate models was more sensitive to variations in the numbers of input variables and hidden neurons as compared to that of training ratio. Improvement in the surrogate performance when configuring the MLPs with more than four input variables or 25 hidden neurons came with significant increases in training and prediction time. These metrics depict the number of seconds taken to train an MLP and to conduct one million predictions using the trained surrogate model, respectively. Under the stop rules for MLP training defined in section 5.3.2.6, an MLP took from 2 seconds to over 15 minutes to train and required less than 1.3 seconds for one million predictions.

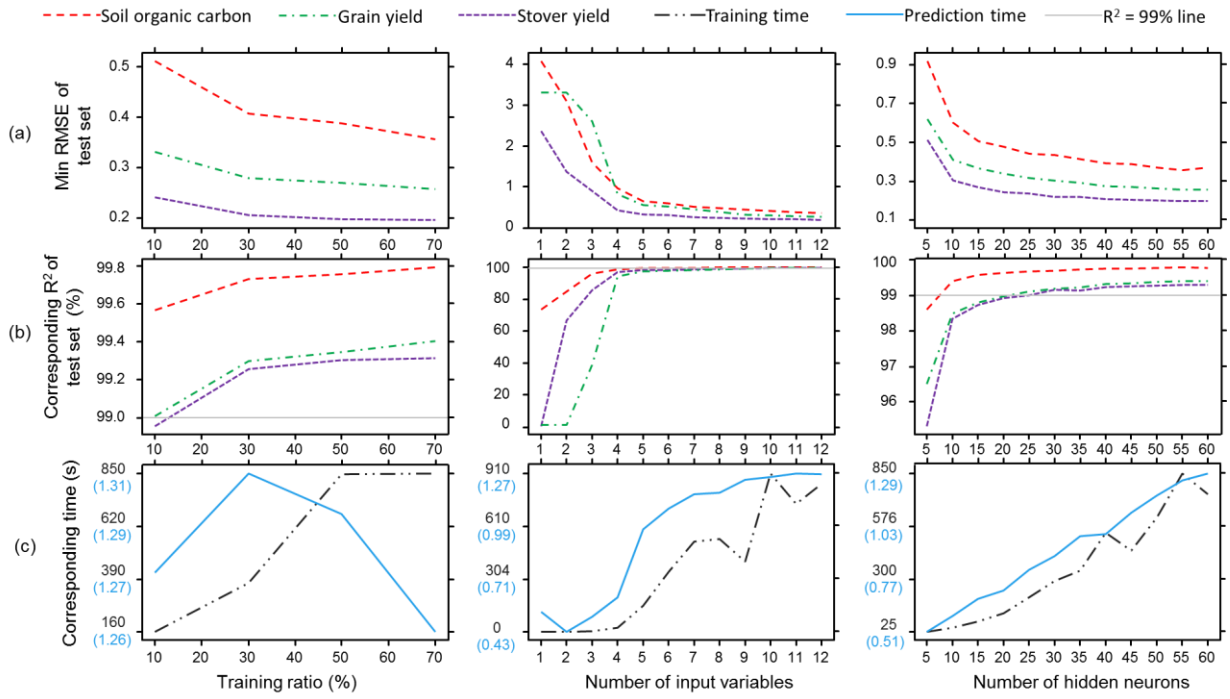


Fig. 5.4. (a) Minimum Root Mean Squared Error (RMSE) of test sets averaged across different train ratios, number of inputs variables, and number of hidden neurons; and their corresponding (b) test-set R^2, and (c) training and prediction time. The grey horizontal lines in the middle plots are the test-set R^2= 99% lines. Prediction time is labeled in blue color with parenthesis.

The final surrogate model that balances between metamodeling accuracy (i.e., test-set R^2 for all targets are greater than 99%) and prediction time (i.e., fastest) is presented in Table 5.3. Other details including regression plots and MLP's weight and bias configuration are available in the Appendix C3. The final surrogate model exhibited high accuracy with the RMSE less than 0.46 Mg C ha^{-1}, 0.25 Mg ha^{-1}, and 0.33 Mg ha^{-1} for SOC, corn stover, and corn grain, respectively. The gain in simulation time was substantial as compared to the original DayCent model. While the DayCent took about 59.6 ± 0.33 days (extrapolated from 5.15 ± 0.33 second per simulation; n=10) for one million simulations on a quad core i7 CPU (2.20GHz and 12Gb RAM), its surrogate model only required 0.83 ± 0.08 seconds on the same setup. This indicated that the surrogate model was about 6.2 million times faster for our analysis.

Table 5.3. The configuration and statistics of the chosen surrogate model

	SOC (Mg C ha^{-1})	Stover yield (Mg ha^{-1})	Grain yield (Mg ha^{-1})
Train set ratio (%)	---------- 70 ----------		
Number of input features	---------- 12 ----------		
Number of hidden neurons	---------- 25 ----------		
Training time (s)	---------- 179 ----------		
Prediction time (s)	---------- 0.82 ± 0.08* ----------		
RMSE of training set	0.42	0.23	0.30
RMSE of validation set	0.45	0.24	0.32
RMSE of test set	0.45	0.22	0.31
R^2 of train set (%)	99.7	99.1	99.2
R^2 of validation set (%)	99.7	99.0	99.1
R^2 of test set (%)	99.7	99.1	99.1

Note: * The standard deviation was calculated from ten independent batches of one million predictions.

5.4.2 Input variable importance and dependence

The BAU's SOC and management practices appeared to be the most important variables in the surrogate model (Table 5.4). Importance was defined as the reduction in prediction accuracy (i.e., increase in MSE) when comparing the MSE of the complete model (with all input variables) to the average of MSE of all submodels that excluded the input variable of interest. This result confirmed the diminishing effect of increasing input variables on accuracy improvement (Fig. 5.4). The dependence plot of the first four most important input variables indicated that the output variables responded to variation in input variables in expected trends but with different magnitudes (Fig. 5.5). The effects of varying BAU's SOC and residue removal rate on the output variables were quite linear while those of N fertilizer and irrigation followed convex curves indicating diminishing management effects at higher investment of management practices. Farm SOC seemed to be more influenced by changes in its initial value, whereas corn stover and grain yields responded with higher magnitudes under variation in residue removal, and N fertilization and irrigation, respectively.

Table 5.4. Relative importance of the input variable to the surrogate model

Input variables	Unit	Relative importance (%)
BAU's soil organic carbon	Mg C ha^{-1}	30.43
Residue removal rate	%	16.53
Fertilizer	kg N ha^{-1}	15.67
Irrigation	cm $year^{-1}$	13.30
BAU's soil nitrogen	Mg N ha^{-1}	3.68
Silt	%	3.57
30-year-average maximum temperature	^{o}C	3.03
Number of soil layer		2.93
30-year-average precipitation	cm $year^{-1}$	2.92
Sand	%	2.91

Bulk density	g cm^{-3}	2.69
30-year-average minimum temperature	°C	2.34

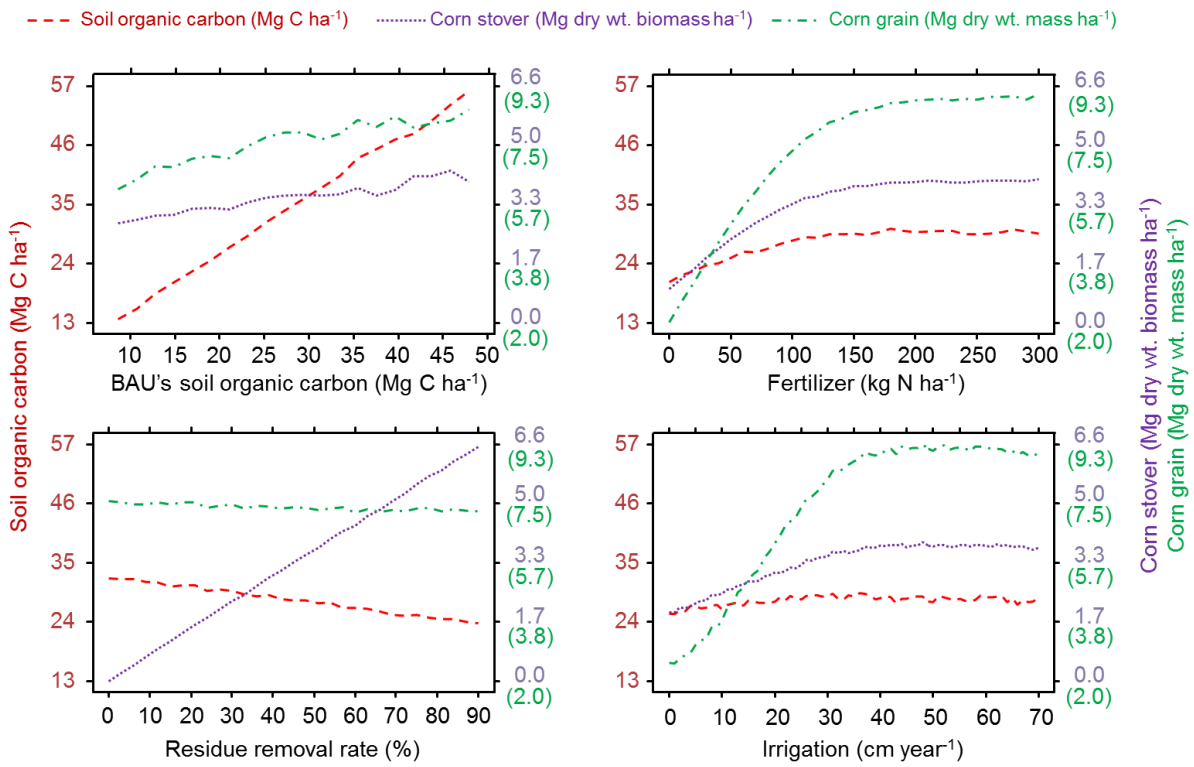


Fig. 5.5. Dependence plot of the first four most important input variables. The dependence plot quantified the ranges of MLP predicted outputs averaged across the entire landscape. The left y-axes were labeled with corn stover yield in Mg dry wt. biomass ha^{-1} (top) and corn grain yield in Mg dry wt. mass ha^{-1} (bottom with parenthesis). The dependence plots for other input variables are available in Appendix C4.

5.4.3 Spatial variation of the predictive errors

The spatial distributions of the RMSE between DayCent simulated outputs and surrogate model predictions indicated a high accuracy of the surrogate model throughout the landscape (Fig. 5.6). The 95th percentiles of the RMSE for the prediction of SOC, corn stover and corn grain yields were 0.77 Mg C ha^{-1}, 0.37 Mg ha^{-1}, and 0.53 Mg ha^{-1}, respectively. The maps of RMSE spatial

distributions revealed higher concentration of sites with lower accuracy (i.e., RMSE above 95th percentiles) around areas west of the city of Greeley and east of the town of Sterling. Although, our exploratory analysis did not reveal any specific correlation between the input variables and the errors (Appendix C5), it might suggest that the aggregation of input variables for more efficient metamodeling was not efficient in explaining potential seasonal variations of output variables for certain regions in the landscape.

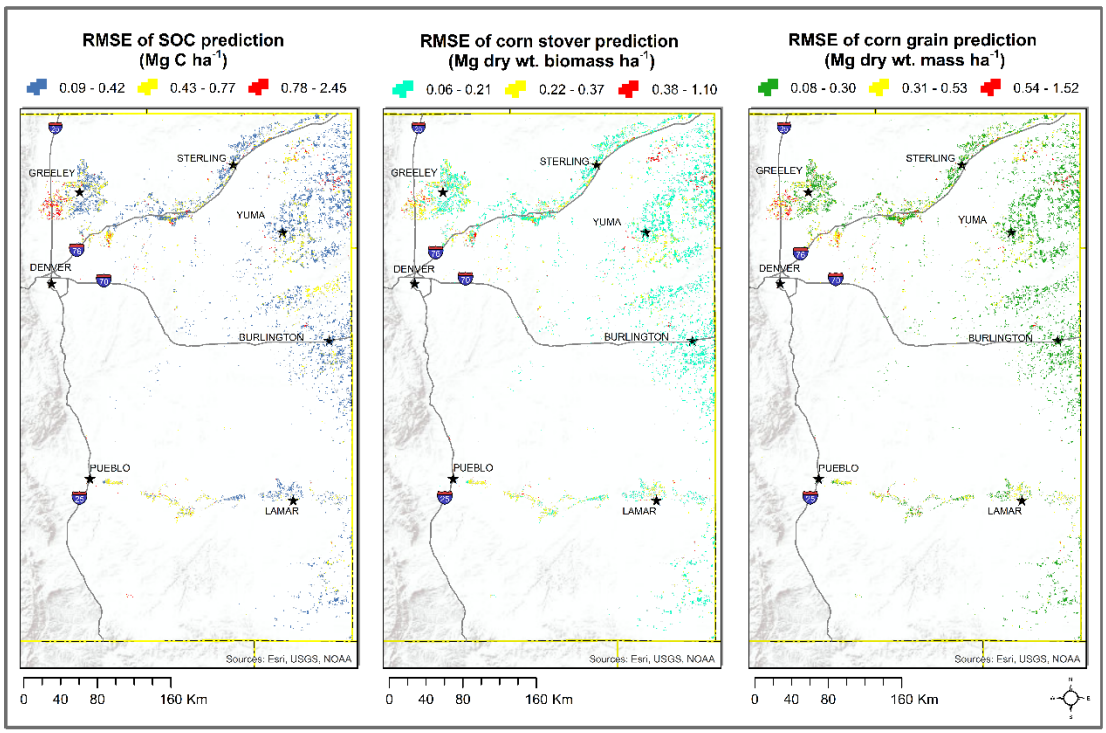


Fig. 5.6. Spatial variation of Root Mean Squared Error (RMSE) between surrogate model and DayCent for (a) soil organic carbon; (b) corn stover yield; and (c) corn grain yield. The maps were colored with three distinct value ranges based on the minimum values, 75th percentiles, 95th percentiles, and maximum values to better identify the spatial patterns of the area with high prediction error. RMSE was calculated from 216 random samples per corn field.

Table 5.5. Average results of farm-level optimization*

Relative SOC constraint (%)	Profit ($ ha^{-1}year^{-1})	Stover yield (Mg ha^{-1}year^{-1})	Grain yield (Mg ha^{-1}year^{-1})	SOC (Mg ha^{-1})	Fertilizer (kg ha^{-1})	Irrigation (cm year^{-1})	Residue removal (%)	GHG (Mg ha^{-1}year^{-1})	N leaching (kg ha^{-1}year^{-1})
BAU	152	7.2	10.4	22.1	170	34.0	75.0	2.66	45
BAU-NT**	233	7.2	10.3	27.4	170	34.0	75.0	1.83	31
0	379	9.1	11.4	25.8	176	48.0	89.2	2.14	51
10	382	9.1	11.4	25.7	176	48.1	89.0	2.15	51
20	369	8.4	11.4	27.0	178	48.0	80.1	2.01	51
30	340	6.7	11.5	29.3	179	47.3	63.6	1.74	48
40	309	4.9	11.5	31.5	181	46.8	46.6	1.47	47
50	277	3.1	11.5	33.8	184	46.5	29.5	1.21	46

Notes:
*The average results were calculated by dividing the area-weighted sum of farm values over total area.
** "Business-as-usual" scenario where conventional tillage was converted to no-till.
Color scales were added to aid the visualization. White-to-green and white-to-red color scales depict metrics whose higher values (green) and lower values (white) were assumed to be preferred, respectively. Each column is colored based on its value range with white color being the min value. Black and white numbers are values below and above the average, respectively

5.4.4 Farm-level optimization of profits

The results associated with farm profits maximization are presented in Table 5.5. The optimization improved farm's profits by 83% to 150% as compared to the BAU scenario. The improvements of farm's profits were due to the lowered production fixed costs of no-till and the higher biomass production attained by the intensified fertilization and irrigation. As we imposed higher relative SOC constraints, our model reduced irrigation and residue removal rates to satisfy SOC constraints while increasing N fertilization rate to improve biomass production for farm profits maximization. Residue removal rate was decreased to below the BAU's level at the relative SOC constraint greater than 20%. While both SOC accumulation and GHG emissions reduction were significantly improved under all relative-SOC-constrained scenarios as compared to the BAU scenario, N leaching was slightly worse due to higher fertilization and irrigation rates. When switching from conventional tillage to no-till under the BAU scenario (BAU-NT), we observed 54% and 24% increases in farm profits and SOC, and 31% and 32% decreases in GHG emissions and N leaching, respectively.

5.4.5 Feasible value ranges and improvement regions of landscape constraints

The feasible value ranges of SOC, irrigation water, and corn grain ranged from 7.42 – 10.15 ($\times 10^6$ Mg C year^{-1}), 1.29 – 1.44 ($\times 10^9$ m^3 year^{-1}), and 2.48 – 3.45 ($\times 10^6$ Mg year^{-1}), respectively (Table 5.6). The base landscape constraints of SOC and corn grain fell between their feasible ranges while the base constraint of irrigation water use laid above its feasible range. This indicated that the empirical irrigation constraint had been satisfied through the farm's profits optimizations. Therefore, for the landscape multi-objective optimization, we re-set irrigation constraint to the mean of its feasible value range (1.36 $\times 10^9$ m^3 year^{-1}) while maintaining SOC

and corn grain constraints at their empirical values of 8.19 $\times 10^6$ Mg C $year^{-1}$ and 3.34 $\times 10^6$ Mg $year^{-1}$, respectively.

Table 5.6. Feasible value ranges and constraint improvement regions of landscape constraints

Constraint variables	Feasible value range	Constrain improvement region (CIR)	Base values*
SOC ($\times 10^6$ Mg C $year^{-1}$)	7.42 - 10.15	8.19 – 10.15	8.19
Irrigation water ($\times 10^9$ m^3 $year^{-1}$)	1.29 – 1.44	1.29 – 1.36	1.50
Corn grain ($\times 10^6$ Mg $year^{-1}$)	2.48 – 3.45	3.34 – 3.45	3.34

Note: * Base values were derived from literature or simulation of BAU scenario under no-till.

5.4.6 Pareto trade-offs

The pareto surface of trade-offs among corn stover, GHG, and N leaching of irrigated corn production in eastern Colorado had a bowed shape revealing three different trade-off patterns between different pairs of landscape objectives (Fig. 5.7). While corn stover and GHG displayed a straight-forward linear relation (Fig. 5.7d), N leaching-related trade-offs showed more complicated relationships with a concave corn stover-N leaching pareto frontier (Fig. 5.7b) and a convex GHG-N leaching pareto frontier (Fig. 5.7c). Under the base landscape constraints, the maximum amount of corn stover that could be produced was 2.35 $\times 10^6$ Mg biomass $year^{-1}$ bearing the highest impacts on ecosystem at 14.37 $\times 10^3$ Mg N $year^{-1}$ and 57 $\times 10^4$ Mg CO_2e $year^{-1}$. Improving the environmental performance of the ecosystem was achieved at the trade-off on corn stover. The minimum GHG and N leaching was 33.50 $\times 10^4$ Mg CO_2e $year^{-1}$ and 11.95 $\times 10^3$ Mg N $year^{-1}$ corresponding to the corn stover production of 0.83 (minimum) and 0.96 ($\times 10^6$ Mg biomass $year^{-1}$), respectively. The landscape sum of profits linearly correlated with corn stover reaching the maximum of \$106.89 $\times 10^6$ $year^{-1}$ at the maximum stover production level and the

minimum of \$80.63 $\times 10^6$ $year^{-1}$ at 0.87 $\times 10^6$ Mg biomass $year^{-1}$ of corn stover, 33.90 $\times 10^4$ Mg CO_2e $year^{-1}$ of GHG, and 12.02 $\times 10^3$ Mg N $year^{-1}$ of N leaching.

Raising the landscape constraints (i.e., higher SOC and corn grain, and lower irrigation water use) worsened corn stover, N leaching, and GHG objectives (Fig. 5.7). Maximum SOC and corn grain production and irrigation water saving were obtained at the stover production levels below 1.1 $\times 10^6$ Mg biomass $year^{-1}$ and fell far in the sub-optimal regions of the stover, N leaching and GHG trade-off frontiers. The mean-CIR improvement scenarios for SOC and irrigation water saving were achieved mainly by decreasing either stover removal rate or irrigation amount (Table 5.7), and thus resulted in decreases in corn stover production. The trade-off showed that we could improve SOC and reduce irrigation water use to a certain extent without affecting either food production or environmental performance of the ecosystem by producing less corn stover. Contrarily, although improving food production (mean CIR_{grain}) did not greatly reduce stover or increase GHG emissions, it significantly increased N leaching as showed by the shifting of the green dotted lines to the right of the base constraint pareto frontiers (Fig. 5.7b). The increase in N leaching under this scenario was due to higher application of both N fertilizer and irrigation (Table 5.7). The summary of landscape trade-offs among production constraints and optimization objectives was given in the Appendix C6.

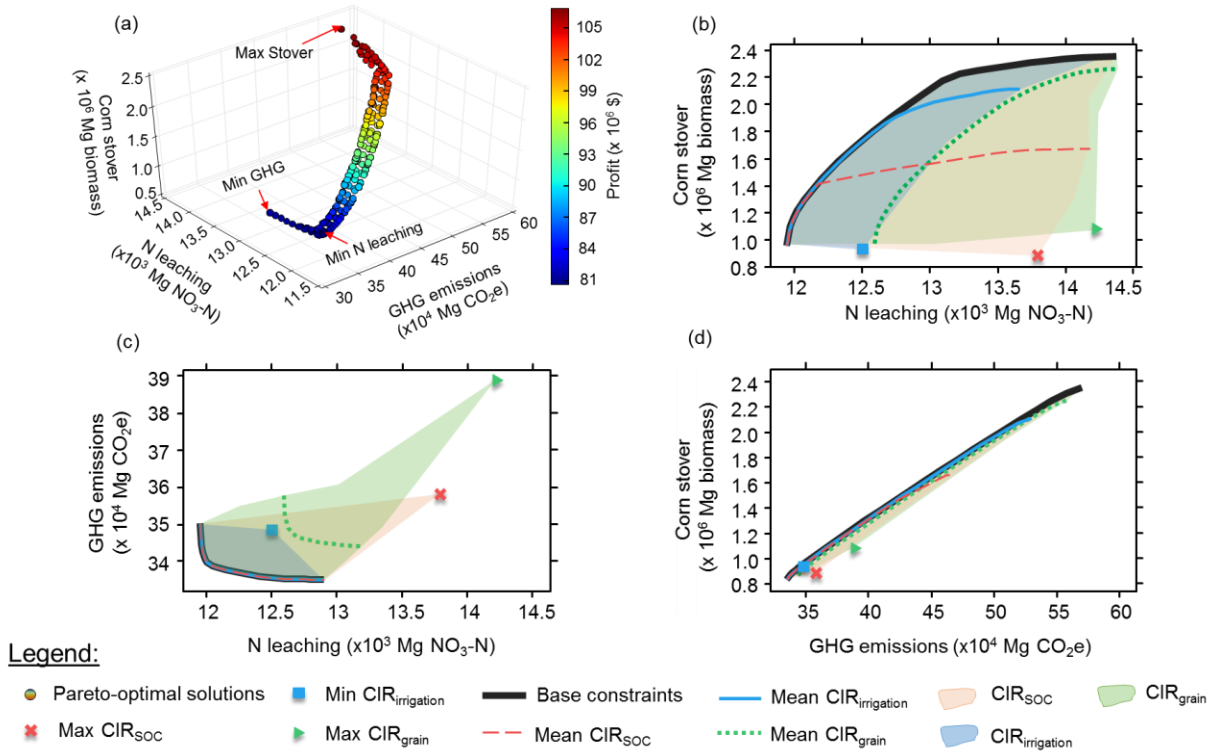


Fig. 5.7. Pareto trade-off among landscape optimization objectives and constraints. (a) 3D pareto surface showing trade-off among corn stover, GHG, and N leaching. The pareto solutions are colored based on the sum of farm profits; (b) Pareto trade-off between corn stover and N leaching; (c) Pareto trade-off between GHG emissions and N; (d) Pareto trade-off between corn stover and GHG emissions. The line shapes are Pareto frontiers. The filled areas defined the feasible regions of Pareto trade-offs when varying a constraint value along its Constrain Improvement Region (CIR). The min, mean, and max CIR are scenarios where the subscripted constraint metric was improved to the min, mean, and max of its CIR, respectively.

Table 5.7. Average management practices of landscape trade-off pareto frontiers.

Pareto trade-off	Constraint scenario	N fertilizer (kg N ha^{-1})	Irrigation (cm $year^{-1}$)	Residue removal (%)
	Max CIR_{grain}	184	47	37
	Min $CIR_{irrigation}$	176	43	32
	Max CIR_{soc}	184	46	30
Stover vs. N leaching	Base	174	44	60
	Mean CIR_{grain}	177	45	59
	Mean $CIR_{irrigation}$	174	44	58
	Mean CIR_{soc}	177	44	50
GHG vs. N leaching	Base	176	43	29
	meanCIR_{grain}	179	44	31
	mean$CIR_{irrigation}$	176	43	29
	meanCIR_{soc}	176	43	29
Stover vs. GHG	Base	174	44	51
	meanCIR_{grain}	177	45	51
	mean$CIR_{irrigation}$	174	44	48
	meanCIR_{soc}	176	44	42

Note: The average of a Pareto frontier was computed from all solutions in that Pareto frontier. The Pareto trade-offs and constraint scenarios are corresponded to those in Fig. 5.7. Each column is colored based on its value range with white color being the min value. The min, mean, and max CIR are scenarios where the subscripted constraint metric was improved to the min, mean, and max of its Constrain Improvement Region (CIR), respectively.

5.4.7 Optimal stover harvest

Fig. 5.8 shows an example of optimized landscape configuration of residue removal rate at different pareto trade-off levels among stover production, GHG emissions, and N leaching, under the base landscape constraints. To increase stover production, residue removal rate was first increased in the areas around the towns of Yuma, Sterling and Greeley and then the areas around Burlington and Lamar. This revealed that more stover could be produced with less environmental impacts (i.e., lower environmental footprints) in the upper part as compared to lower part of the study region.

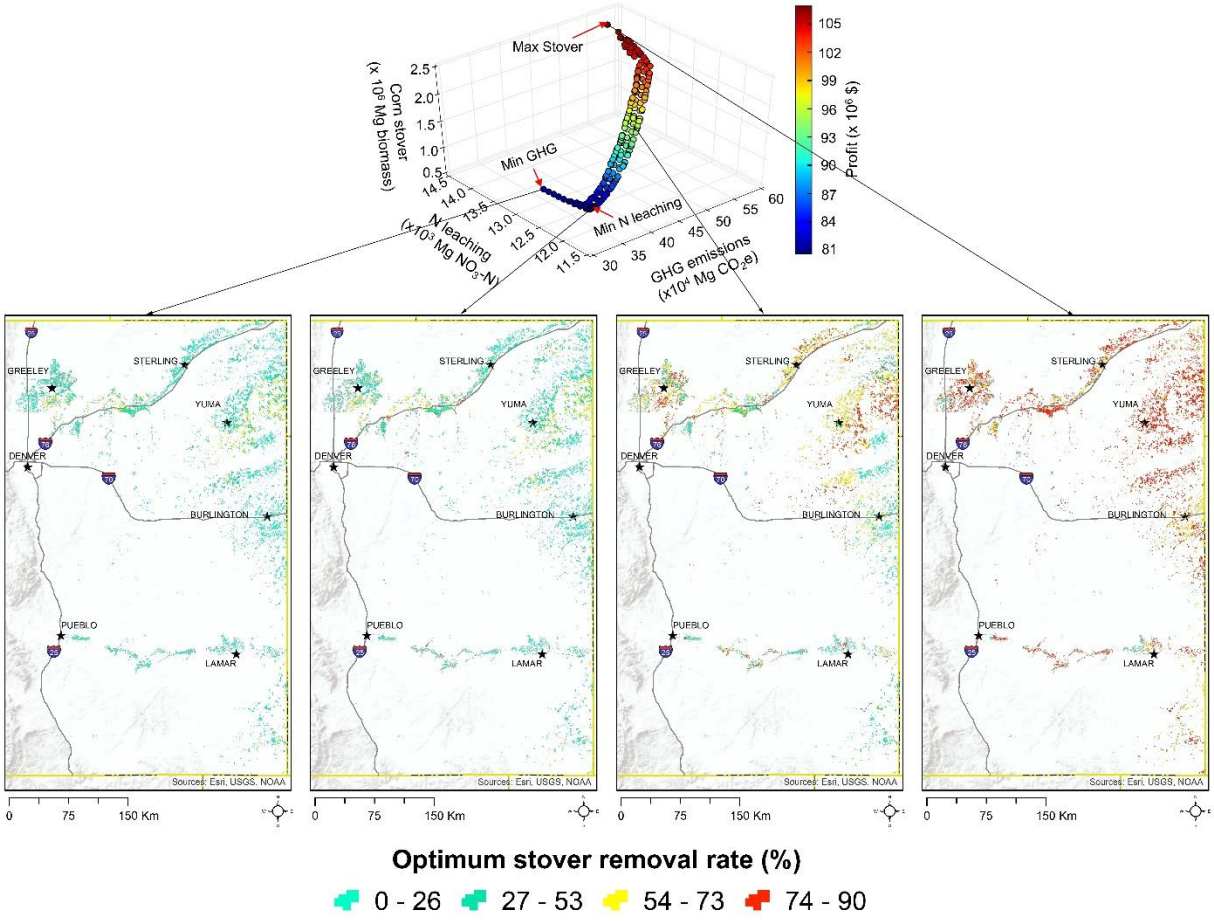


Fig. 5.8. Landscape configuration of stover removal rates at different trade-off level among stover production, GHG emissions, and N leaching. The maps were colored with four distinct value ranges based on the minimum values, 25th percentile, 50th percentile, 75th percentile and maximum values of all the pareto-optimum solutions.

5.5 DISCUSSION

The results of our analysis could be used to support decision-making at multiple levels. At farm level, farmers (the direct decision makers of agricultural ecosystems) could be informed with the optimum levels of resource investment to maximize farm profits and improve soil's carbon storage. Farmers could also adjust their management decisions based on the quantitative trade-offs among different ecosystem services (e.g., carbon storage and food and biomass production)

and disservices (e.g., GHG emissions and N leaching). This increases farm-level efficiency and sustainability. Furthermore, as farm-level environmental performance was made available, farmers could be encouraged to incorporate these metrics in their farm planning and budgeting, allowing more thorough state-wide agricultural statistic reports. At regional level, our results could be used to support several applications such as life cycle assessment and supply chain optimization of first- and/or second-generation biofuel production, biorefinery sitting, and feedstock landscape optimization. They could also be aggregated into county-level or state-level assessments to support policy-making and regional planning. For example, different stover removal rates could be recommended for different counties to approximate an optimal trade-off level among food and biomass production, carbon storage, and environmental impacts (an example is given in Appendix C7).

As argued by Levin (1992), the problem of pattern and scale in ecology and ecosystem science is to understand the simultaneous changes that are taking place on many scales. With significant gain in simulation time, as exampled by our study, metamodeling allows a model to be used at larger scales to explore the spatial and temporal dynamics of interactions among ecosystem processes. This advantage also permits simulation-expensive applications like sensitivity analysis and optimization, which were conventionally impractical for the original complex models, in many cases. While our study demonstrated the metamodeling of the DayCent model for an agricultural landscape optimization, our framework would be equally applicable for other ecosystem models as well as optimization problems at different scales. Realizing that the complexity of such a modeling framework might hinder its practical use, we automated the workflow and incorporated it into our Ag-EcoSOpt system to facilitate future analyses. Besides,

efforts are being made to integrate other metamodeling techniques and optimization algorithms for more dynamic applications of our modeling platform.

Despite the potential advantages of the framework, we do realize some drawbacks in the methodology. First, the computational burden of the metamodeling process was substantial. Generating the dataset for metamodeling required the use of a cluster computing system. While we could argue that the surrogate model could be trained once for different analyses, this computational burden might still challenge the adoption of our metamodeling framework especially when the computing infrastructure is not available for a large number of original model simulations. However, our results showed that training ratio did not significantly affect surrogate model performance (Fig. 5.4 a, b). This indicated that we could have decreased the sample size to reduce the computational burden of dataset generation as well as the MLP training time. For this, space-filling sampling techniques like Latin Hypercube design (Alam et al., 2004) could be used in place of the full factorial design employed by our study.

Although the training time of each MLP configuration ranged from a couple seconds to 15 minutes, optimization of MLP configuration via sensitivity analysis required thousands of iterative MLP trainings, which was time-consuming. Since our results suggested that a slight increase in accuracy could come with a considerable increase in training and prediction time (Fig. 5.4), we could set the MSE_{train} (training goal) at a certain level above zero (instead of zero) to stop MLP training earlier. In addition, the maximum number of neurons in the hidden layer could be reduced in future studies. By training a surrogate model for a single crop in a small region, we had reduced the crop-specific input variables and many other site- and management-specific input variables such as planting dates and fertilization timing. This improved the efficiency of the surrogate model. Fig. 5.4a also suggests that training MLP with more than four

input variables did not significantly improve the surrogate model accuracy. This implied that other input reduction techniques such as principle component analysis could be used to decrease number of input variables, thus lower the training time.

The MLP prediction time could also be improved if farm profits (i.e., the farm-level objective function) were directly metamodeled. In our study, simulated corn grain and stover yields were first metamodeled and then used to calculate farm profits. While this approach allowed a more flexible use of the surrogate model to address changes in economic factors such as commodity prices, it slightly increased the prediction time by 8%. The optimizations required more than 24 hours on a quad core i7 CPU for both farm and landscape levels. For applications where time saving is more crucial, reducing intermediate steps is recommended to improve optimization efficiency. To further reduce optimization time, we equipped the Ag-EcoSOpt with the ability to use the Parallel Computing Toolbox in MATLAB for simultaneous simulation of multiple farm-level optimization instances when resources are available.

Our analysis was subject to several sources of uncertainty. By nature, surrogate models inherit the uncertainties of the original models, which are those from model input, parameters, and structure. In addition, surrogate models suffer from their own structural uncertainty, which was measured as the errors between predicted outputs and the modeled outputs. The latter was quantified in our study for surrogate model selection. There were also other uncertainties associated with the optimization process such as those of economic data. These uncertainties would propagate through the system and should be reflected in the output to inform decision makers. While the quantification of uncertainty is valuable to decision making, such analysis is difficult to include in regional assessments such as this due to a) limited amounts of empirical data against which to validate model function, and b) the highly computational-intensive nature

of such operations. These challenges are significant enough to fall outside the scope of our study, and merit separate thorough studies. However, we note that meaningful sensitivity and uncertainty analysis is only possible after a system has been initially characterized, and the important potential feedbacks and dependencies identified.

It is also important to note that the surrogate should not be used outside of the domain for which it was designed. In our analysis, the MLP was trained for irrigated corn in eastern Colorado with certain bounds of management-specific input variables. The use of the surrogate model for another corn growing regions or beyond the management practice bounds is not recommended. In these cases, one should follow the framework described in this study to construct and thoroughly test surrogate models before employing them for their specific study. Constructing a robust surrogate model to be used over larger areas and to predict several outputs for many crop species would improve the generalization of the metamodeling. This could also encourage studies of broader scopes on larger scales to support strategic decision-making process. However, such effort requires laborious data acquisition and processing as well as comprehensive experimental designs that encompass all possible input distributions. Furthermore, a generic surrogate is likely to be bulky and slow, and thus is inefficient for simulation-intensive applications like optimization. Alternatively, simplified surrogate models developed for specific problems like ours are more efficient for local and operational decision making. Our future efforts will focus on building more generic surrogate models for different crops in the U.S. as well as extending the Ag-EcoSOpt's database to facilitate the adoption of our framework for similar resource optimization problems.

Although multi-level and multi-objective optimization at high spatial resolutions is supposed to support decision-making process, we found that communicating the results of such optimization

to stakeholders and decision makers is a big challenge. Complicated optimization problems like ours could result in thousands of graphs, tables, and maps, which would be difficult to present through the traditional communication means such as reports and scientific papers. The use of web mapping applications allows integrated and customized presentations of information at different scales, serving the needs of multiple stakeholders. Such applications will also permit the involvement and feedback from stakeholders for more practical and operational analyses. The future development of Ag-EcoSOpt platform will focus on building a web mapping application to facilitate the visualization of analyses that employed our platform.

5.6 CONCLUSION

In this study, we demonstrated the use of metamodeling techniques in a surrogate-based optimization framework for decision support in agricultural resource management. For this, we conducted a case study to optimize the management practices of irrigated corn production systems in Colorado. The optimization explored the trade-offs among seven objectives at farm and landscape scales, including farm profits, irrigation water use, corn grain, corn stover, soil organic carbon (SOC), greenhouse gas (GHG) emissions, and nitrogen (N) leaching. The Multilayer perceptron (MLP), a feedforward artificial neural network, was employed to create a surrogate of the DayCent biogeochemical model. The surrogate model was used to create the objective and constraint functions for farm-level optimization. Outputs from farm-level optimization were then used to generate scenarios for landscape-level optimization. The optimizations were subjected to certain constraints in soil carbon storage, food production and irrigation water use. Our results showed that the surrogate model was able to capture above 99% variations in the simulated SOC, corn grain and stover yields. In addition, it was 6.2 million times faster than the DayCent model and was suitable for use in computationally-intensive

applications like sensitivity analysis and optimization. Optimizing management practices at the farm-level improved farm profit, SOC, grain yield, and reduced GHG emissions while slightly increasing N leaching. Landscape-level optimization resulted in different unique trade-off patterns among corn stover, GHG, and N leaching. The optimum stover removal rates were also visualized for each corn farm to illustrate the landscape configuration of management practices to obtain certain level of trade-offs. We discussed the advantages and disadvantages of our surrogate-based optimization framework as well as the potential to couple our framework with web mapping applications to facilitate better communication of high spatial-resolution optimization results to decision makers. We believe that this combination would encourage the participation and feedback from stakeholders for more practical and operational optimization analyses to support decision-making in sustainable resource management.

BIBLIOGRAPHY

Adelekan, B.A., 2011. Cassava as a Potent Energy Crop for the Production of Ethanol and Methane in Tropical Countries. Int. J. Therm. Environ. Eng. 4, 25–32. https://doi.org/10.5383/ijtee.04.01.004

Akgul, O., Zamboni, A., Bezzo, F., Shah, N., Papageorgiou, L.G., 2011. Optimization-Based Approaches for Bioethanol Supply Chains. Ind. Eng. Chem. Res. 50, 4927–4938. https://doi.org/10.1021/ie101392y

Alam, F.M., McNaught, K.R., Ringrose, T.J., 2004. A comparison of experimental designs in the development of a neural network simulation metamodel. Simul. Model. Pract. Theory, Simulation in Operational Research 12, 559–578. https://doi.org/10.1016/j.simpat.2003.10.006

Aldea, J., Martínez-Peña, F., Romero, C., Diaz-Balteiro, L., 2014. Participatory Goal Programming in Forest Management: An Application Integrating Several Ecosystem Services. Forests 5, 3352–3371. https://doi.org/10.3390/f5123352

Arshad, M.A., Martin, S., 2002. Identifying critical limits for soil quality indicators in agro-ecosystems. Agric. Ecosyst. Environ., Soil Health as an Indicator of Sustainable Management 88, 153–160. https://doi.org/10.1016/S0167-8809(01)00252-3

Audsley, E., Pearn, K.R., Harrison, P.A., Berry, P.M., 2008. The impact of future socio-economic and climate changes on agricultural land use and the wider environment in East Anglia and North West England using a metamodel system. Clim. Change 90, 57–88. https://doi.org/10.1007/s10584-008-9450-9

Awudu, I., Zhang, J., 2012. Uncertainties and sustainability concepts in biofuel supply chain management: A review. Renew. Sustain. Energy Rev. 16, 1359–1368. https://doi.org/10.1016/j.rser.2011.10.016

Bagstad, K.J., Semmens, D.J., Waage, S., Winthrop, R., 2013. A comparative assessment of decision-support tools for ecosystem services quantification and valuation. Ecosyst. Serv. 5, 27–39. https://doi.org/10.1016/j.ecoser.2013.07.004

Bakker, H., 2012. Sugar Cane Cultivation and Management. Springer Science & Business Media.

Balbi, S., Prado, A. del, Gallejones, P., Geevan, C.P., Pardo, G., Pérez-Miñana, E., Manrique, R., Hernandez-Santiago, C., Villa, F., 2015. Modeling trade-offs among ecosystem services in agricultural production systems. Environ. Model. Softw. 72, 314–326. https://doi.org/10.1016/j.envsoft.2014.12.017

Balesdent, J., Chenu, C., Balabane, M., 2000. Relationship of soil organic matter dynamics to physical protection and tillage. Soil Tillage Res. 53, 215–230.

Bandaranayake, W., Qian, Y.L., Parton, W.J., Ojima, D.S., Follett, R.F., 2003. Estimation of soil organic carbon changes in turfgrass systems using the CENTURY model. Agron. J. 95, 558–563.

Battiston, L.A., Miller, M.H., Shelton, I.J., 1987. Soil Erosion and Corn Yield in Ontario. I. Field Evaluation. Can. J. Soil Sci. 67, 731–745. https://doi.org/10.4141/cjss87-072

Bauer, A., Black, A.L., 1992. Organic Carbon Effects on Available Water Capacity of Three Soil Textural Groups. Soil Sci. Soc. Am. J. 56, 248–254. https://doi.org/10.2136/sssaj1992.03615995005600010038x

Berhe, A.A., Harte, J., Harden, J.W., Torn, M.S., 2007. The Significance of the Erosion-induced Terrestrial Carbon Sink. BioScience 57, 337–346. https://doi.org/10.1641/B570408

Bernardi, A., Giarola, S., Bezzo, F., 2013. Spatially Explicit Multiobjective Optimization for the Strategic Design of First and Second Generation Biorefineries Including Carbon and Water Footprints. Ind. Eng. Chem. Res. 52, 7170–7180. https://doi.org/10.1021/ie302442j

Bernardi, A., Giarola, S., Bezzo, F., 2012. Optimizing the economics and the carbon and water footprints of bioethanol supply chains. Biofuels Bioprod. Biorefining 6, 656–672. https://doi.org/10.1002/bbb.1358

BioGrace, 2015. The BioGrace GHG calculation tool: a recognised voluntary scheme [WWW Document]. Harmon. Calc. Bioenergy Greenh. Gas Emiss. Eur. URL http://www.biograce.net/content/ghgcalculationtools/recognisedtool/ (accessed 3.20.15).

Bramanti, A., Di Barba, P., Farina, M., Savini, A., 2001. Combining response surfaces and evolutionary strategies for multiobjective pareto-optimization in electromagnetics. Int. J. Appl. Electromagn. Mech. 15, 231–236.

Brink, C., Van Grinsven, H., 2011. The European Nitrogen Assessment: Costs and benefits of nitrogen in the environment - Chapter 22. Camb. Univ. Press.

Brown, L.R., 2012. Full planet, empty plates: the new geopolitics of food scarcity, First Edition. ed. W.W. Norton & Company, New York.

Bruinsma, J., 2009. The resource outlook to 2050. Economic and Social Development Department,Food and Agriculture Organization of the United Nations.

Buck-Sorlin, P.D.G., 2013. Process-based Model, in: Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H. (Eds.), Encyclopedia of Systems Biology. Springer New York, pp. 1755–1755. https://doi.org/10.1007/978-1-4419-9863-7_1545

Carnevale, C., Finzi, G., Guariso, G., Pisoni, E., Volta, M., 2012. Surrogate models to compute optimal air quality planning policies at a regional scale. Environ. Model. Softw., Emulation techniques for the reduction and sensitivity analysis of complex environmental models 34, 44–50. https://doi.org/10.1016/j.envsoft.2011.04.007

Carré, F., Intergovernmental Panel on Climate Change (Eds.), 2010. Background guide for the calculation of land carbon stocks in the biofuels sustainability scheme: drawing on the 2006 IPCC guidelines for national greenhouse gas inventories, EUR Scientific and technical research series. Publ. Off. of the European Union, Luxembourg.

CDSS, 2010. Geographic Information System Data, Colorado's Decision Support Systems [WWW Document]. URL http://cdss.state.co.us/GIS/Pages/GISDataHome.aspx (accessed 11.28.16).

Cerri, C.E.P., Easter, M., Paustian, K., Killian, K., Coleman, K., Bernoux, M., Falloon, P., Powlson, D.S., Batjes, N., Milne, E., Cerri, C.C., 2007. Simulating SOC changes in 11 land use change chronosequences from the Brazilian Amazon with RothC and Century models. Agric. Ecosyst. Environ. 122, 46–57. https://doi.org/10.1016/j.agee.2007.01.007

Chan, K.M.A., Shaw, M.R., Cameron, D.R., Underwood, E.C., Daily, G.C., 2006. Conservation Planning for Ecosystem Services. PLOS Biol. 4, e379. https://doi.org/10.1371/journal.pbio.0040379

Cheng, K., Ogle, S.M., Parton, W.J., Pan, G., 2014. Simulating greenhouse gas mitigation potentials for Chinese Croplands using the DAYCENT ecosystem model. Glob. Change Biol. 20, 948–962. https://doi.org/10.1111/gcb.12368

Cheng, K., Ogle, S.M., Parton, W.J., Pan, G., 2013. Predicting methanogenesis from rice paddies using the DAYCENT ecosystem model. Ecol. Model. 261–262, 19–31. https://doi.org/10.1016/j.ecolmodel.2013.04.003

Clark, M., Tilman, D., 2017. Comparative analysis of environmental impacts of agricultural production systems, agricultural input efficiency, and food choice. Environ. Res. Lett. 12, 064016. https://doi.org/10.1088/1748-9326/aa6cd5

Coleman, K., Jenkinson, D.S., 1996. RothC-26.3 - A Model for the turnover of carbon in soil, in: Powlson, D.S., Smith, P., Smith, J.U. (Eds.), Evaluation of Soil Organic Matter Models, NATO ASI Series. Springer Berlin Heidelberg, pp. 237–246.

Compton, J.E., Harrison, J.A., Dennis, R.L., Greaver, T.L., Hill, B.H., Jordan, S.J., Walker, H., Campbell, H.V., 2011. Ecosystem services altered by human changes in the nitrogen cycle: a new perspective for US decision making. Ecol. Lett. 14, 804–815. https://doi.org/10.1111/j.1461-0248.2011.01631.x

Corsano, G., Vecchietti, A.R., Montagna, J.M., 2011. Optimal design for sustainable bioethanol supply chain considering detailed plant performance model. Comput. Chem. Eng., Energy & Sustainability 35, 1384–1398. https://doi.org/10.1016/j.compchemeng.2011.01.008

Čuček, L., Klemeš, J.J., Kravanja, Z., 2012. A Review of Footprint analysis tools for monitoring impacts on sustainability. J. Clean. Prod., Recent Cleaner Production Advances in

Process Monitoring and Optimisation 34, 9–20. https://doi.org/10.1016/j.jclepro.2012.02.036

Čuček, L., Lam, H.L., Klemeš, J.J., Varbanov, P.S., Kravanja, Z., 2010. Synthesis of regional networks for the supply of energy and bioproducts. Clean Technol. Environ. Policy 12, 635–645. https://doi.org/10.1007/s10098-010-0312-6

Dal-Mas, M., Giarola, S., Zamboni, A., Bezzo, F., 2011. Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty. Biomass Bioenergy 35, 2059–2071. https://doi.org/10.1016/j.biombioe.2011.01.060

David, M.B., Grosso, S.J.D., Hu, X., Marshall, E.P., McIsaac, G.F., Parton, W.J., Tonitto, C., Youssef, M.A., 2008. Modeling denitrification in a tile-drained, corn and soybean agroecosystem of Illinois, USA. Biogeochemistry 93, 7–30. https://doi.org/10.1007/s10533-008-9273-9

Davis, S.C., Boddey, R.M., Alves, B.J.R., Cowie, A.L., George, B.H., Ogle, S.M., Smith, P., van Noordwijk, M., van Wijk, M.T., 2013. Management swing potential for bioenergy crops. GCB Bioenergy 5, 623–638. https://doi.org/10.1111/gcbb.12042

Davis, S.C., Parton, W.J., Grosso, S.J.D., Keough, C., Marx, E., Adler, P.R., DeLucia, E.H., 2011. Impact of second-generation biofuel agriculture on greenhouse-gas emissions in the corn-growing regions of the US. Front. Ecol. Environ. 10, 69–74. https://doi.org/10.1890/110003

de Groot, R.S., Alkemade, R., Braat, L., Hein, L., Willemen, L., 2010. Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. Ecol. Complex. 7, 260–272. https://doi.org/10.1016/j.ecocom.2009.10.006

DeAngelis, D.L., Yurek, S., 2017. Spatially Explicit Modeling in Ecology: A Review. Ecosystems 20, 284–300. https://doi.org/10.1007/s10021-016-0066-z

Deb, K., 2011. Multi-objective optimisation using evolutionary algorithms: an introduction, in: Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing. Springer, pp. 3–34.

Decision no. 177/2007/QD-TTg of Government of Vietnam, 2007. Approval of the scheme on development of biofuel up to 2015 and a vision to 2025.

Deininger, K., Byerlee, D., 2011. Rising Global Interest in Farmland: Can It Yield Sustainable and Equitable Benefits? The World Bank. https://doi.org/10.1596/978-0-8213-8591-3

Del Grosso, S., Ojima, D., Parton, W., Mosier, A., Peterson, G., Schimel, D., 2002. Simulated effects of dryland cropping intensification on soil organic matter and greenhouse gas exchanges using the DAYCENT ecosystem model. Environ. Pollut. Barking Essex 1987 116 Suppl 1, S75-83.

Del Grosso, S.J., Halvorson, A.D., Parton, W.J., 2008a. Testing DAYCENT Model Simulations of Corn Yields and Nitrous Oxide Emissions in Irrigated Tillage Systems in Colorado. J. Environ. Qual. 37, 1383. https://doi.org/10.2134/jeq2007.0292

Del Grosso, S.J., Ogle, S.M., Parton, W.J., Breidt, F.J., 2010. Estimating uncertainty in N2O emissions from U.S. cropland soils. Glob. Biogeochem. Cycles 24, GB1009. https://doi.org/10.1029/2009GB003544

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Ojima, D.S., Kulmala, A.E., Phongpan, S., 2000. General model for N2O and N2 gas emissions from soils due to dentrification. Glob. Biogeochem. Cycles 14, 1045–1060. https://doi.org/10.1029/1999GB001225

Del Grosso, S.J., Parton, W.J., Ojima, D.S., Keough, C.A., Riley, T.H., Mosier, A.R., 2008b. DAYCENT Simulated Effects of Land Use and Climate on County Level N Loss Vectors in the USA. Publ. USDA-ARSUNL Fac. 255.

Delgrosso, S., Mosier, A., Parton, W., Ojima, D., 2005. DAYCENT model analysis of past and contemporary soil NO and net greenhouse gas flux for major crops in the USA. Soil Tillage Res. 83, 9–24. https://doi.org/10.1016/j.still.2005.02.007

D'haeze, D., Deckers, J., Raes, D., Phong, T.A., Minh Chanh, N.D., 2003. Over-irrigation of Coffea canephora in the Central Highlands of Vietnam revisited: Simulation of soil moisture dynamics in Rhodic Ferralsols. Agric. Water Manag. 63, 185–202. https://doi.org/10.1016/S0378-3774(03)00181-1

Easter, M., Paustian, K., Killian, K., Boyack, T., Williams, S., Feng, T., Coleman, K., Swan, A., Al-Adamat, R., Bhattacharrya, T., Cerri, C.E.P., Kamoni, P., Batjes, N., Milne, E., 2005. User Manual: GEFSOC Soil Carbon Modeling System. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, USA.

Efroymson, R.A., Dale, V.H., Kline, K.L., McBride, A.C., Bielicki, J.M., Smith, R.L., Parish, E.S., Schweizer, P.E., Shaw, D.M., 2013. Environmental Indicators of Biofuel Sustainability: What About Context? Environ. Manage. 51, 291–306. https://doi.org/10.1007/s00267-012-9907-5

EPA-RFS2, 2010. Renewable Fuel Standard (RFS2): Final Rule (No. Legal Authority: 42 U.S.C. §7545). United States Environmental Protection Agency.

ERS-ARMS, 2010. Farm Financial and Crop Production Practices. Economic Research Service, United States Department of Agriculture.

ESRI, 2014. ArcGIS for Desktop. Environmental Systems Research Institute, Redlands, CA.

Eve, M., Pape, D., Flugge, M., Steele, R., Man, D., Gilbert, M., Biggar, S., 2014. Quantifying Greenhouse Gas Fluxes in Agriculture and Forestry: Methods for Entity-Scale Inventory (No. Technical Bulletin Number 1939). Washington, DC: U.S. Department of Agriculture, Office of the Chief Economist.

Ewing, P.M., Runck, B.C., 2015. Optimizing nitrogen rates in the midwestern United States for maximum ecosystem value. Ecol. Soc. 20. https://doi.org/10.5751/ES-06767-200118

FAO, 2014. FAOSTAT commodity definitions and correspondences: Classifications and standards [WWW Document]. URL http://www.fao.org/economic/ess/ess-standards/en/#.WE8Mc1MrKpo (accessed 12.12.16).

FAO, 2013. Food and Agriculture Organization of The United Nations - Statistics Division [WWW Document]. URL http://faostat3.fao.org/download/Q/QC/E (accessed 3.15.15).

Farrell, A.E., Plevin, R.J., Turner, B.T., Jones, A.D., O'Hare, M., Kammen, D.M., 2006. Ethanol Can Contribute to Energy and Environmental Goals. Science 311, 506–508. https://doi.org/10.1126/science.1121416

Finkbeiner, M., Inaba, A., Tan, R., Christiansen, K., Klüppel, H.-J., 2006. The New International Standards for Life Cycle Assessment: ISO 14040 and ISO 14044. Int. J. Life Cycle Assess. 11, 80–85. https://doi.org/10.1065/lca2006.02.002

Fischer, G., Heilig, G.K., 1997. Population momentum and the demand on land and water resources. Philos. Trans. R. Soc. B Biol. Sci. 352, 869–889. https://doi.org/10.1098/rstb.1997.0067

Fishburn, P.C., 1967. Additive Utilities with Incomplete Product Sets: Application to Priorities and Assignments. Oper. Res. 15, 537–542.

Funahashi, K.-I., 1989. On the approximate realization of continuous mappings by neural networks. Neural Netw. 2, 183–192. https://doi.org/10.1016/0893-6080(89)90003-8

Gasparatos, A., Stromberg, P., Takeuchi, K., 2011. Biofuels, ecosystem services and human wellbeing: Putting biofuels in the ecosystem services narrative. Agric. Ecosyst. Environ. 142, 111–128. https://doi.org/10.1016/j.agee.2011.04.020

Gesch, R.W., Archer, D.W., Forcella, F., 2010. Rotational Effects of Cuphea on Corn, Spring Wheat, and Soybean. Agron. J. 102, 145. https://doi.org/10.2134/agronj2009.0215

Giarola, S., Zamboni, A., Bezzo, F., 2011. Spatially explicit multi-objective optimisation for design and planning of hybrid first and second generation biorefineries. Comput. Chem. Eng., Energy Systems Engineering 35, 1782–1797. https://doi.org/10.1016/j.compchemeng.2011.01.020

Gibbs, H.K., Johnston, M., Foley, J.A., Holloway, T., Monfreda, C., Ramankutty, N., Zaks, D., 2008. Carbon payback times for crop-based biofuel expansion in the tropics: the effects of changing yield and technology. Environ. Res. Lett. 3, 034001. https://doi.org/10.1088/1748-9326/3/3/034001

Glendining, M.J., Dailey, A.G., Williams, A.G., Evert, F.K. van, Goulding, K.W.T., Whitmore, A.P., 2009. Is it possible to increase the sustainability of arable and ruminant agriculture by reducing inputs? Agric. Syst. 99, 117–125. https://doi.org/10.1016/j.agsy.2008.11.001

Gnansounou, E., Dauriat, A., Villegas, J., Panichelli, L., 2009a. Life cycle assessment of biofuels: Energy and greenhouse gas balances. Bioresour. Technol. 100, 4919–4930. https://doi.org/10.1016/j.biortech.2009.05.067

Gnansounou, E., Dauriat, A., Villegas, J., Panichelli, L., 2009b. Life cycle assessment of biofuels: Energy and greenhouse gas balances. Bioresour. Technol. 100, 4919–4930. https://doi.org/10.1016/j.biortech.2009.05.067

Greenland, D.J., Rimmer, D., Payne, D., 1975. Determination of the Structural Stability Class of English and Welsh Soils, Using a Water Coherence Test. J. Soil Sci. 26, 294–303. https://doi.org/10.1111/j.1365-2389.1975.tb01953.x

GSOV, 2014. General Statistics Office of Vietnam [WWW Document]. URL http://gso.gov.vn/Default_en.aspx?tabid=766 (accessed 11.1.15).

Ha, P.Q., 2010. Carbon in Vietnamese Soils and Experiences to Improve Carbon Stock in Soil, in: Proceeding. Presented at the Int. Workshop on Evaluation and Sustainable Management of Soil Carbon Sequestration in Asian Countries, Bogor, Indonesia.

Haimes, Y.Y., Lasdon, L.S., Wismer, D.A., 1971. On a Bicriterion Formulation of the Problems of Integrated System Identification and System Optimization. IEEE Trans. Syst. Man Cybern. SMC-1, 296–297. https://doi.org/10.1109/TSMC.1971.4308298

Haines-Young, R., Potschin, M., 2010. The links between biodiversity, ecosystem services and human well-being. Ecosyst. Ecol. New Synth. 110–139.

Halvorson, A.D., Mosier, A.R., Reule, C.A., Bausch, W.C., 2006. Nitrogen and Tillage Effects on Irrigated Continuous Corn Yields. Agron. J. 98, 63. https://doi.org/10.2134/agronj2005.0174

Harden, J.W., Berhe, A.A., Torn, M., Harte, J., Liu, S., Stallard, R.F., 2008. Soil Erosion: Data Say C Sink. Science 320, 178–179. https://doi.org/10.1126/science.320.5873.178

Hartikainen, M., Miettinen, K., Wiecek, M.M., 2012. PAINT: Pareto front interpolation for nonlinear multiobjective optimization. Comput. Optim. Appl. 52, 845–867. https://doi.org/10.1007/s10589-011-9441-z

Haykin, S., 1998. Neural Networks: A Comprehensive Foundation, 2nd ed. Prentice Hall PTR, Upper Saddle River, NJ, USA.

Hill, R.D., Peart, M.R., 1998. Land use, runoff, erosion and their control: a review for southern China. Hydrol. Process. 12, 2029–2042. https://doi.org/10.1002/(SICI)1099-1085(19981030)12:13/14<2029::AID-HYP717>3.0.CO;2-O

Hoang, H., Phan, C., Hoang, T., Chen, W., Bell, R., 2010. Sandy soils in South Central Coastal Vietnam: Their origin, constraints and management, in: Proceedings of the 19th World Congress of Soil Science: Soil Solutions for a Changing World, Brisbane, Australia, 1-6 August 2010. pp. 251–254.

Hoang, K., Nguyen, V., Hoang, L., Nguyen, T., Ceballos, H., Howeler, R., 2010. Recent progress in cassava breeding and the selection of improved cultivars in Vietnam, in: In: Howeler RH, Editor. A New Future for Cassava in Asia: Its Use as Food, Feed and Fuel to Benefit the Poor. Proceedings of the 8th Region Al Workshop, 2008 Oct 20-24. Bangkok: Centro Internacional de Agriciltura Tropical, Vientiane, Lao PDR.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Howeler, R.H., 2000. Cassava production practices in Asia-Can they maintain soil productivity. Presented at the International Symposium on Cassava, Starch and Starch Derivatives, Guangxi (China), 11-15 Nov 1996.

Howeler, R.H., 1991. Long-term effect of cassava cultivation on soil productivity. Field Crops Res. 26, 1–18.

Hu, Z., Fang, F., Ben, D., Pu, G., Wang, C., 2004. Net energy, CO2 emission, and life-cycle cost assessment of cassava-based ethanol as an alternative automotive fuel in China. Appl. Energy 78, 247–256. https://doi.org/10.1016/j.apenergy.2003.09.003

Huang, Y., Chen, C.-W., Fan, Y., 2010. Multistage optimization of the supply chains of biofuels. Transp. Res. Part E Logist. Transp. Rev. 46, 820–830. https://doi.org/10.1016/j.tre.2010.03.002

Husson, O., Castella, J.-C., Ha Dinh Tuan, N.K., 2001. Agronomic diagnosis and identification of factors limiting upland rice yield in mountainous areas of northern Vietnam. SAM Pap. Ser. 2.

Hwang, C.L., Paidy, S.R., Yoon, K., Masud, A.S.M., 1980. Mathematical programming with multiple objectives: A tutorial. Comput. Oper. Res. 7, 5–31. https://doi.org/10.1016/0305-0548(80)90011-8

Iback, D., Adams, J., 1967. Fertilizer Use in the United States, by Crops and Areas- 1964 Estimates. Statistical Reporting Service, Economic Research Service, Department of Agriculture, United States.

Ibendahl, G., O'Brien, D.M., Duncan, 2015. Center-Pivot-Irrigated Corn Cost-Return Budget in North Central Kansas (No. MF2601). Department of Agricultural Economics, Kansas State University.

IBM®, 2017. IBM ILOG CPLEX Optimization Studio. IBM®.

IBM Corp, 2015. IBM SPSS Modeler 17 Algorithms Guide.

IPCC, 2014. Climate Change 2014: Synthesis Report, Fifth Assessment Report (AR5).

IPCC, 2006. 2006 IPCC guidelines for national greenhouse gas inventories. Institute for Global Environmental Strategies, Hayama, Japan.

ISO 14044, 2006. Environmental management -Life cycle assessment - Requirements and guidelines. International Organisation for Standardisation (ISO), Geneve.

Jenks, G.F., 1967. The Data Model Concept in Statistical Mapping. Int. Yearb. Cartogr. 7 186–190.

Johnson, J.W., Lebreton, J.M., 2004. History and Use of Relative Importance Indices in Organizational Research. Organ. Res. Methods 7, 238–257. https://doi.org/10.1177/1094428104266510

Kaastra, I., Boyd, M.S., 1995. Forecasting futures trading volume using neural networks. J. Futur. Mark. 15, 953–970. https://doi.org/10.1002/fut.3990150806

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron., Modelling Cropping Systems: Science, Software and Applications 18, 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9

Keeler, B.L., Gourevitch, J.D., Polasky, S., Isbell, F., Tessum, C.W., Hill, J.D., Marshall, J.D., 2016. The social costs of nitrogen. Sci. Adv. 2, e1600219. https://doi.org/10.1126/sciadv.1600219

Kennedy, C.M., Hawthorne, P.L., Miteva, D.A., Baumgarten, L., Sochi, K., Matsumoto, M., Evans, J.S., Polasky, S., Hamel, P., Vieira, E.M., Develey, P.F., Sekercioglu, C.H.,

Davidson, A.D., Uhlhorn, E.M., Kiesecker, J., 2016. Optimizing land use decision-making to sustain Brazilian agricultural profits, biodiversity and ecosystem services. Biol. Conserv. 204, Part B, 221–230. https://doi.org/10.1016/j.biocon.2016.10.039

Kershner, J., Samhouri, J.F., James, C.A., Levin, P.S., 2011. Selecting Indicator Portfolios for Marine Species and Food Webs: A Puget Sound Case Study. PLoS ONE 6, e25248. https://doi.org/10.1371/journal.pone.0025248

Kim, I.Y., Weck, O.L. de, 2006. Adaptive weighted sum method for multiobjective optimization: a new method for Pareto front generation. Struct. Multidiscip. Optim. 31, 105–116. https://doi.org/10.1007/s00158-005-0557-6

Kim, S., Dale, B.E., Jenkins, R., 2009. Life cycle assessment of corn grain and corn stover in the United States. Int. J. Life Cycle Assess. 14, 160–174. https://doi.org/10.1007/s11367-008-0054-4

King, E., Cavender-Bares, J., Balvanera, P., Mwampamba, T.H., Polasky, S., 2015. Trade-offs in ecosystem services and varying stakeholder preferences: evaluating conflicts, obstacles, and opportunities. Ecol Soc 20, 25.

Klinnert, C., others, 2001. Problems with and local solutions for organic matter management in Vietnam. Nutr. Cycl. Agroecosystems 61, 89–97.

Körschens, M., Weigel, A., Schulz, E., 1998. Turnover of soil organic matter (SOM) and long-term balances — tools for evaluating sustainable productivity of soils. Z. Für Pflanzenernähr. Bodenkd. 161, 409–424. https://doi.org/10.1002/jpln.1998.3581610409

Kourakos, G., Mantoglou, A., 2013. Development of a multi-objective optimization algorithm using surrogate models for coastal aquifer management. J. Hydrol. 479, 13–23. https://doi.org/10.1016/j.jhydrol.2012.10.050

Kragt, M.E., Robertson, M.J., 2014. Quantifying ecosystem services trade-offs from agricultural practices. Ecol. Econ. 102, 147–157. https://doi.org/10.1016/j.ecolecon.2014.04.001

Lal, R., Griffin, M., Apt, J., Lave, L., Morgan, M.G., 2004. Ecology. Managing soil carbon. Science 304, 393. https://doi.org/10.1126/science.1093079

Lardy, R., Bachelet, B., Bellocchi, G., Hill, D.R.C., 2014. Towards vulnerability minimization of grassland soil organic matter using metamodels. Environ. Model. Softw. 52, 38–50. https://doi.org/10.1016/j.envsoft.2013.10.015

Lautenbach, S., Volk, M., Strauch, M., Whittaker, G., Seppelt, R., 2013. Optimization-based trade-off analysis of biodiesel crop production for managing an agricultural catchment. Environ. Model. Softw. 48, 98–112. https://doi.org/10.1016/j.envsoft.2013.06.006

Le, L.T., van Ierland, E.C., Zhu, X., Wesseler, J., 2013. Energy and greenhouse gas balances of cassava-based ethanol. Biomass Bioenergy 51, 125–135. https://doi.org/10.1016/j.biombioe.2013.01.011

Ledoux, H., Gold, C., 2005. An efficient natural neighbour interpolation algorithm for geoscientific modelling, in: Developments in Spatial Data Handling. Springer, pp. 97–108.

Leduc, S., Lundgren, J., Franklin, O., Dotzauer, E., 2010. Location of a biomass based methanol production plant: A dynamic problem in northern Sweden. Appl. Energy 87, 68–75. https://doi.org/10.1016/j.apenergy.2009.02.009

Leng, R., Wang, C., Zhang, C., Dai, D., Pu, G., 2008. Life cycle inventory and energy analysis of cassava-based Fuel ethanol in China. J. Clean. Prod. 16, 374–384. https://doi.org/10.1016/j.jclepro.2006.12.003

Lester, S.E., Costello, C., Halpern, B.S., Gaines, S.D., White, C., Barth, J.A., 2013. Evaluating tradeoffs among ecosystem services to inform marine spatial planning. Mar. Policy 38, 80–89. https://doi.org/10.1016/j.marpol.2012.05.022

Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. Q. J. Appl. Mathmatics II, 164–168.

Levin, S.A., 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology 73, 1943–1967. https://doi.org/10.2307/1941447

Li, C., Frolking, S., Frolking, T.A., 1992. A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. J. Geophys. Res. Atmospheres 97, 9759–9776. https://doi.org/10.1029/92JD00509

Li, H., Chen, C.L.P., Huang, H.-P., 2000. Fuzzy Neural Intelligent Systems: Mathematical Foundation and the Applications in Engineering, 1 edition. ed. CRC Press, Boca Raton, FL.

Liska, A.J., Yang, H.S., Bremer, V.R., Klopfenstein, T.J., Walters, D.T., Erickson, G.E., Cassman, K.G., 2009. Improvements in Life Cycle Energy Efficiency and Greenhouse Gas Emissions of Corn-Ethanol. J. Ind. Ecol. 13, 58–74. https://doi.org/10.1111/j.1530-9290.2008.00105.x

Liu, B., Wang, F., Zhang, B., Bi, J., 2013. Energy balance and GHG emissions of cassava-based fuel ethanol using different planting modes in China. Energy Policy 56, 210–220. https://doi.org/10.1016/j.enpol.2012.12.052

Loomis, J., Kent, P., Strange, L., Fausch, K., Covich, A., 2000. Measuring the total economic value of restoring ecosystem services in an impaired river basin: results from a contingent valuation survey. Ecol. Econ. 33, 103–117.

Lotov, A., Bushenkov, V.A., Kamenev, G.K., 2013. Interactive Decision Maps: Approximation and Visualization of Pareto Frontier. Springer Science & Business Media.

Loveland, P., Webb, J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil Tillage Res. 70, 1–18.

Luo, L., van der Voet, E., Huppes, G., 2009. An energy analysis of ethanol from cellulosic feedstock–Corn stover. Renew. Sustain. Energy Rev. 13, 2003–2011. https://doi.org/10.1016/j.rser.2009.01.016

MA, 2005a. Ecosystems and human well-being: synthesis, Millennium Ecosystem Assessment. ed. Island Press, Washington, DC.

MA, 2005b. Ecosystems and Human Well-being: A Framework for Assessment, Millennium Ecosystem Assessment. Island Press, Washington, DC.

Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model. Softw. 15, 101–124. https://doi.org/10.1016/S1364-8152(99)00007-9

Maier, H.R., Jain, A., Dandy, G.C., Sudheer, K.P., 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environ. Model. Softw. 25, 891–909. https://doi.org/10.1016/j.envsoft.2010.02.003

MARD, 2015a. Statistics and Food Security Database - Weather - Vietnam Ministry of Agriculture and Rural Development [WWW Document]. URL http://www.mard.gov.vn/Pages/statistic_csdl.aspx?TabId=thongke

MARD, 2015b. National Center for Agricultural Extension - Data - Vietnam Ministry of Agriculture and Rural Development [WWW Document]. URL

http://www.khuyennongvn.gov.vn/en/h%C6%B0%E1%BB%9Bng-d%E1%BA%ABn-s%E1%BA%A3n-xu%E1%BA%A5t-nong-nghi%E1%BB%87p.aspx (accessed 3.10.15).

Marie, G., Simioni, G., 2014. Extending the use of ecological models without sacrificing details: a generic and parsimonious meta-modelling approach. Methods Ecol. Evol. 5, 934–943. https://doi.org/10.1111/2041-210X.12250

Marler, R.T., Arora, J.S., 2010. The weighted sum method for multi-objective optimization: new insights. Struct. Multidiscip. Optim. 41, 853–862. https://doi.org/10.1007/s00158-009-0460-7

Marquardt, D., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. Soc. Ind. Appl. Math. 11, 431–441. https://doi.org/10.1137/0111030

Martin, D.L., Dorn, T.W., Kranz, W.L., Melvin, S.R., Corr, A.J., 2010. Reducing the cost of pumping irrigation water, in: 2010 CENTRAL PLAINS IRRIGATION CONFERENCE. p. 41.

Martín, J., Bielza, C., Insua, D.R., 2005. Approximating nondominated sets in continuous multiobjective optimization problems. Nav. Res. Logist. NRL 52, 469–480. https://doi.org/10.1002/nav.20090

MATLAB and Neural Network Toolbox, 2017. Neural Network Toolbox Documentation. The MathWorks, Inc., Natick, Massachusetts, United States.

MATLAB and Optimization Toolbox, 2017. Optimization Toolbox Documentation - Fmincon solver. The MathWorks, Inc., Natick, Massachusetts, United States.

Maynard, S., James, D., Davidson, A., 2010. The Development of an Ecosystem Services Framework for South East Queensland. Environ. Manage. 45, 881–895. https://doi.org/10.1007/s00267-010-9428-z

McIntosh, B.S., Ascough, J.C., Twery, M., Chew, J., Elmahdi, A., Haase, D., Harou, J.J., Hepting, D., Cuddy, S., Jakeman, A.J., Chen, S., Kassahun, A., Lautenbach, S., Matthews, K., Merritt, W., Quinn, N.W.T., Rodriguez-Roda, I., Sieber, S., Stavenga, M., Sulis, A., Ticehurst, J., Volk, M., Wrobel, M., van Delden, H., El-Sawah, S., Rizzoli, A., Voinov, A., 2011. Environmental decision support systems (EDSS) development – Challenges and best practices. Environ. Model. Softw. 26, 1389–1402. https://doi.org/10.1016/j.envsoft.2011.09.009

Meckesheimer, M., Booker, A.J., Barton, R.R., Simpson, T.W., 2002. Computationally Inexpensive Metamodel Assessment Strategies. AIAA J. 40, 2053–2060. https://doi.org/10.2514/2.1538

Mele, F.D., Guillén-Gosálbez, G., Jiménez, L., 2009. Optimal Planning of Supply Chains for Bioethanol and Sugar Production with Economic and Environmental Concerns. Comput. Aided Chem. Eng., 19th European Symposium on Computer Aided Process Engineering 26, 997–1002. https://doi.org/10.1016/S1570-7946(09)70166-X

Mele, F.D., Kostin, A.M., Guillén-Gosálbez, G., Jiménez, L., 2011. Multiobjective Model for More Sustainable Fuel Supply Chains. A Case Study of the Sugar Cane Industry in Argentina. Ind. Eng. Chem. Res. 50, 4939–4958. https://doi.org/10.1021/ie101400g

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American Regional Reanalysis. Bull. Am. Meteorol. Soc. 87, 343–360. https://doi.org/10.1175/BAMS-87-3-343

Metherell, A.K., Harding, L.A., Cole, C.V., Parton, W.J., 1993. CENTURY Soil Organic Matter Model Environment, Technical Documentation Agroecosystem Version 4.0 (Technical Documentation Agroecosystem Version 4.0 No. Technical Report No. 4), Great Plains System Research Unit. USDA-ARS, Fort Collins, Colorado.

Miettinen, K., Ruiz, F., Wierzbicki, A.P., 2008. Introduction to Multiobjective Optimization: Interactive Approaches, in: Branke, J., Deb, K., Miettinen, K., Słowiński, R. (Eds.), Multiobjective Optimization, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 27–57. https://doi.org/10.1007/978-3-540-88908-3_2

Mitchell, S., Consulting, S.M., Dunning, I., 2011. PuLP: A Linear Programming Toolkit for Python. The University of Auckland, Auckland, New Zealand.

Moore, F.C., Diaz, D.B., 2015. Temperature impacts on economic growth warrant stringent mitigation policy. Nat. Clim. Change 5, 127–131. https://doi.org/10.1038/nclimate2481

Motta, R. de S., Afonso, S.M.B., Lyra, P.R.M., 2012. A modified NBI and NC method for the solution of N-multiobjective optimization problems. Struct. Multidiscip. Optim. 46, 239–259. https://doi.org/10.1007/s00158-011-0729-5

Mueller, D.K., Hamilton, P.A., Helsel, D.R., Hitt, K.J., Ruddy, B.C., 1995. Nutrients in ground water and surface water of the United States—an analysis of data through 1992. US Geol. Surv. Water-Resour. Investig. Rep. 95, 74.

Muth Jr., D.J., Bryden, K.M., Nelson, R.G., 2013. Sustainable agricultural residue removal for bioenergy: A spatially comprehensive US national assessment. Appl. Energy, Special Issue on Advances in sustainable biofuel production and use - XIX International Symposium on Alcohol Fuels - ISAF 102, 403–417. https://doi.org/10.1016/j.apenergy.2012.07.028

Nahlik, A.M., Kentula, M.E., Fennessy, M.S., Landers, D.H., 2012. Where is the consensus? A proposed foundation for moving ecosystem service concepts into practice. Ecol. Econ. 77, 27–35. https://doi.org/10.1016/j.ecolecon.2012.01.001

NASS, 2012. USDA - NASS, Census of Agriculture - 2012 Census Volume 1, Chapter 1: State Level Data [WWW Document]. URL https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_1_State_Level/Colorado/ (accessed 2.26.18).

NASS, 2010. Usual Planting and Harvesting Dates for U.S. Field Crops, Agricultural Handbook Number 628. National Agriculture Statistics Services, United States Department of Agriculture.

NASS-CDL, 2016. CropScape - National Agricultural Statistics Services -Crop Data Layer Program [WWW Document]. URL https://nassgeodata.gmu.edu/CropScape/ (accessed 8.1.16).

Newson, M., 1997. Land, Water and Development: Sustainable Management of River Basin Systems, Second Edition, second edition. ed. Routledge.

Nguyen, Q.H., Egashira, K., 2008. Clay Mineralogy of Grey Degraded Soils and Sandy Soils as Problem in Vietnam.

Nguyen, T.H., Cook, M., Field, J.L., Khuc, Q.V., Paustian, K., 2018. High-resolution trade-off analysis and optimization of ecosystem services and disservices in agricultural landscapes. Environ. Model. Softw. 107, 105–118. https://doi.org/10.1016/j.envsoft.2018.06.006

Nguyen, T.H., Field, J., Paustian, K., in review. High-resolution multi-objective optimization of feedstock landscape design for hybrid first and second generation biorefineries. Environ. Model. Softw.

Nguyen, T.H., Williams, S., Paustian, K., 2017. Impact of ecosystem carbon stock change on greenhouse gas emissions and carbon payback periods of cassava-based ethanol in Vietnam. Biomass Bioenergy 100, 126–137. https://doi.org/10.1016/j.biombioe.2017.02.009

Nguyen, T.L.T., Gheewala, S.H., 2008. Life cycle assessment of fuel ethanol from cassava in Thailand. Int. J. Life Cycle Assess. 13, 147–154. https://doi.org/10.1065/lca2007.06.343

Nguyen, T.L.T., Gheewala, S.H., Garivait, S., 2007. Energy balance and GHG-abatement cost of cassava utilization for fuel ethanol in Thailand. Energy Policy 35, 4585–4596. https://doi.org/10.1016/j.enpol.2007.03.012

Nordhaus, W.D., 2017. Revisiting the social cost of carbon. Proc. Natl. Acad. Sci. 114, 1518–1523. https://doi.org/10.1073/pnas.1609244114

NRCS-USDA, 2014. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database [WWW Document]. URL https://sdmdataaccess.sc.egov.usda.gov (accessed 6.20.14).

Ogle, S.M., Breidt, F.J., Easter, M., Williams, S., Killian, K., Paustian, K., 2010. Scale and uncertainty in modeled soil organic carbon stock changes for US croplands using a process-based model. Glob. Change Biol. 16, 810–822. https://doi.org/10.1111/j.1365-2486.2009.01951.x

Oost, K.V., Quine, T.A., Govers, G., Gryze, S.D., Six, J., Harden, J.W., Ritchie, J.C., McCarty, G.W., Heckrath, G., Kosmas, C., Giraldez, J.V., Silva, J.R.M. da, Merckx, R., 2007. The

Impact of Agricultural Soil Erosion on the Global Carbon Cycle. Science 318, 626–629. https://doi.org/10.1126/science.1145724

Othoniel, B., Rugani, B., Heijungs, R., Benetto, E., Withagen, C., 2016. Assessment of Life Cycle Impacts on Ecosystem Services: Promise, Problems, and Prospects. Environ. Sci. Technol. 50, 1077–1092. https://doi.org/10.1021/acs.est.5b03706

Ou, X., Zhang, X., Chang, S., Guo, Q., 2009. Energy consumption and GHG emissions of six biofuel pathways by LCA in (the) People's Republic of China. Appl. Energy 86, S197–S208. https://doi.org/10.1016/j.apenergy.2009.04.045

Parish, E.S., Hilliard, M.R., Baskaran, L.M., Dale, V.H., Griffiths, N.A., Mulholland, P.J., Sorokine, A., Thomas, N.A., Downing, M.E., Middleton, R.S., 2012. Multimetric spatial optimization of switchgrass plantings across a watershed. Biofuels Bioprod. Biorefining 6, 58–72. https://doi.org/10.1002/bbb.342

Parker, N., Tittmann, P., Hart, Q., Nelson, R., Skog, K., Schmidt, A., Gray, E., Jenkins, B., 2010. Development of a biorefinery optimized biofuel supply curve for the Western United States. Biomass Bioenergy 34, 1597–1607. https://doi.org/10.1016/j.biombioe.2010.06.007

Parton, W.J., Hartman, M., Ojima, D., Schimel, D., 1998. DAYCENT and its land surface submodel: description and testing. Glob. Planet. Change 19, 35–48. https://doi.org/10.1016/S0921-8181(98)00040-X

Parton, W.J., Schimel, D.S., Ojima, D.S., Cole, C.V., 1994. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. Soil Science Society of America Inc., pp. 147–167.

Parton, W.J., Stewart, J.W.B., Cole, C.V., 1988. Dynamics of C, N, P and S in grassland soils: a model. Biogeochemistry 5, 109–131. https://doi.org/10.1007/BF02180320

Paustian, K., Brenner, J., Easter, M., Killian, K., Ogle, S., Olson, C., Schuler, J., Vining, R., Williams, S., 2009. Counting carbon on the farm: Reaping the benefits of carbon offset programs. J. Soil Water Conserv. 64, 36A–40A.

Paustian, K., Levine, E., Post, W.M., Ryzhova, I.M., 1997. The use of models to integrate information and understanding of soil C at the regional scale. Geoderma 79, 227–260. https://doi.org/10.1016/S0016-7061(97)00043-8

Paustian, K., Ogle, S.M., Conant, R.T., 2010. Quantification and Decision Support Tools for US Agricultural Soil Carbon Sequestration, in: Handbook of Climate Change and Agroecosystems, ICP Series on Climate Change Impacts, Adaptation, and Mitigation. IMPERIAL COLLEGE PRESS, pp. 307–341.

Penn State Extension, 2017. The Agronomy Guide 2017 - 2018. College of Agricultural Sciences, The Pennsylvania State University.

Pervez, M.S., Brown, J.F., 2010. Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. Remote Sens. 2, 2388–2412. https://doi.org/10.3390/rs2102388

Plevin, R.J., Jones, A.D., Torn, M.S., Gibbs, H.K., 2010. Greenhouse Gas Emissions from Biofuels' Indirect Land Use Change Are Uncertain but May Be Much Greater than Previously Estimated. Environ. Sci. Technol. 44, 8015–8021. https://doi.org/10.1021/es101946t

Power, A.G., 2010. Ecosystem services and agriculture: tradeoffs and synergies. Philos. Trans. R. Soc. B Biol. Sci. 365, 2959–2971. https://doi.org/10.1098/rstb.2010.0143

Prechelt, L., 2012. Early Stopping — But When?, in: Neural Networks: Tricks of the Trade, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 53–67. https://doi.org/10.1007/978-3-642-35289-8_5

Rathke, G.-W., Wienhold, B.J., Wilhelm, W.W., Diepenbrock, W., 2007. Tillage and rotation effect on corn–soybean energy balances in eastern Nebraska. Soil Tillage Res. 97, 60–70. https://doi.org/10.1016/j.still.2007.08.008

Ratto, M., Castelletti, A., Pagano, A., 2012. Emulation techniques for the reduction and sensitivity analysis of complex environmental models. Environ. Model. Softw., Emulation techniques for the reduction and sensitivity analysis of complex environmental models 34, 1–4. https://doi.org/10.1016/j.envsoft.2011.11.003

Reap, J., Roman, F., Duncan, S., Bras, B., 2008a. A survey of unresolved problems in life cycle assessment. Int. J. Life Cycle Assess. 13, 374–388. https://doi.org/10.1007/s11367-008-0009-9

Reap, J., Roman, F., Duncan, S., Bras, B., 2008b. A survey of unresolved problems in life cycle assessment. Part 1: goal and scope and inventory analysis. Int. J. Life Cycle Assess. 13, 290–300. https://doi.org/10.1007/s11367-008-0008-x

Reijnders, L., Huijbregts, M.A.J., 2007. Life cycle greenhouse gas emissions, fossil fuel demand and solar energy conversion efficiency in European bioethanol production for automotive purposes. J. Clean. Prod. 15, 1806–1812. https://doi.org/10.1016/j.jclepro.2006.05.007

Roy, P., Nei, D., Orikasa, T., Xu, Q., Okadome, H., Nakamura, N., Shiina, T., 2009. A review of life cycle assessment (LCA) on some food products. J. Food Eng. 90, 1–10. https://doi.org/10.1016/j.jfoodeng.2008.06.016

Roy, T., Schütze, N., Grundmann, J., Brettschneider, M., Jain, A., 2016. Optimal groundwater management using state-space surrogate models: a case study for an arid coastal region. J. Hydroinformatics 18, 666–686. https://doi.org/10.2166/hydro.2016.086

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536. https://doi.org/10.1038/323533a0

Russell, J., Dalsted, N., Tranel, J., Young, R.B., Seyler, J., 2016. Custom Ratesfor Colorado Farms & Ranches in 2015. Agriculture & Business Management Notes, Colorado State University Extension.

Rusu, M., 2012. Social Cost of Carbon: Opportunities and Environmental Solutions. Procedia Econ. Finance, International Conference Emerging Markets Queries in Finance and Business, Petru Maior University of Tîrgu-Mures, ROMANIA, October 24th - 27th, 2012 3, 690–697. https://doi.org/10.1016/S2212-5671(12)00215-8

Ruzika, S., Wiecek, M.M., 2005. Approximation Methods in Multiobjective Programming. J. Optim. Theory Appl. 126, 473–501. https://doi.org/10.1007/s10957-005-5494-4

Salami, B.T., Sangoyomi, T.E., 2013. Soil fertility status of cassava fields in south western Nigeria. Am. J. Exp. Agric. 3, 152–164.

Sambridge, M., Braun, J., McQueen, H., 1995. Geophysical parametrization and interpolation of irregular data using natural neighbours. Geophys. J. Int. 122, 837–857. https://doi.org/10.1111/j.1365-246X.1995.tb06841.x

Santibañez-Aguilar, J.E., González-Campos, J.B., Ponce-Ortega, J.M., Serna-González, M., El-Halwagi, M.M., 2011. Optimal Planning of a Biomass Conversion System Considering Economic and Environmental Aspects. Ind. Eng. Chem. Res. 50, 8558–8570. https://doi.org/10.1021/ie102195g

Saxton, K.E., Rawls, W.J., Romberger, J.S., Papendick, R.I., 1986. Estimating Generalized Soil-water Characteristics from Texture. Soil Sci. Soc. Am. J. 50, 1031–1036. https://doi.org/10.2136/sssaj1986.03615995005000040039x

Schneiders, A., Van Daele, T., Van Landuyt, W., Van Reeth, W., 2012. Biodiversity and ecosystem services: Complementary approaches for ecosystem management? Ecol. Indic. 21, 123–133. https://doi.org/10.1016/j.ecolind.2011.06.021

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., Yu, T.-H., 2008. Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change. Science 319, 1238–1240. https://doi.org/10.1126/science.1151861

Seppelt, R., Voinov, A., 2002. Optimization methodology for land use patterns using spatially explicit landscape models. Ecol. Model. 151, 125–142. https://doi.org/10.1016/S0304-3800(01)00455-0

Sibson, R., 1981. A Brief Description of Natural Neighbor Interpolation-Chapter 2. John Wiley Sons N. Y., Interpolating multivariate data 21–36.

Silalertruksa, T., Gheewala, S.H., 2011. Long-Term Bioethanol System and Its Implications on GHG Emissions: A Case Study of Thailand. Environ. Sci. Technol. 45, 4920–4928. https://doi.org/10.1021/es1040915

Simpson, T.W., Poplinski, J.D., Koch, P.N., Allen, J.K., 2001. Metamodels for computer-based engineering design: survey and recommendations. Eng. Comput. 17, 129–150.

Smalley, H.R., Engle, R.H., 1942. Fertilizer use and its effect on crops in 1942. Fertil. Rev. XVIII, 2–3 and 8–11.

Song, J., Yang, Y., Wu, Jianfeng, Wu, Jichun, Sun, X., Lin, J., 2018. Adaptive surrogate model based multiobjective optimization for coastal aquifer management. J. Hydrol. 561, 98–111. https://doi.org/10.1016/j.jhydrol.2018.03.063

Springmann, M., Godfray, H.C.J., Rayner, M., Scarborough, P., 2016. Analysis and valuation of the health and climate change cobenefits of dietary change. Proc. Natl. Acad. Sci. 113, 4146–4151. https://doi.org/10.1073/pnas.1523119113

Sriroth, K., Wanlapatit, S., Piyachomkwan, K., 2012. Cassava Bioethanol. InTech Bioethanol, Marco Aurelio Pinheiro Lima (Ed). https://doi.org/10.5772/23112

Stenemo, F., Lindahl, A.M.L., Gärdenäs, A., Jarvis, N., 2007. Meta-modeling of the pesticide fate model MACRO for groundwater exposure assessments using artificial neural networks. J. Contam. Hydrol. 93, 270–283. https://doi.org/10.1016/j.jconhyd.2007.03.003

Strange, E.M., Fausch, K.D., Covich, A.P., 1999. Sustaining ecosystem services in human-dominated watersheds: biohydrology and ecosystem processes in the South Platte River Basin. Environ. Manage. 24, 39–54.

Sutton, M.A., Bleeker, A., Howard, C.M., Bekunda, M., Grizzetti, B., de Vries, W., van Grinsven, H.J.M., Abrol, Y.P., Adhya, T.K., Billen, G., Davidson, E.A., Datta, A., Diaz, R., Erisman, J.W., Liu, X.J., Oenema, O., Palm, C., Raghuram, N., Reis, S., Scholz, R.W., Sims, T., Westhoek, H., Zhang, F.S., 2013. Our nutrient world: the challenge to produce more food and energy with less pollution. NERC/Centre for Ecology & Hydrology, Edinburgh.

Swinton, S.M., Lupi, F., Robertson, G.P., Hamilton, S.K., 2007. Ecosystem services and agriculture: Cultivating agricultural ecosystems for diverse benefits. Ecol. Econ. 64, 245–252. https://doi.org/10.1016/j.ecolecon.2007.09.020

Thompson, J., Tyner, W.E., 2011. Corn stover for Bioenergy Production: Cost estimates and Farmer supply response, Purdue Extension Renewable Energy series. Purdue Extension, Purdue University.

Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. 108, 20260–20264. https://doi.org/10.1073/pnas.1116437108

Tittmann, P.W., Parker, N.C., Hart, Q.J., Jenkins, B.M., 2010. A spatially explicit techno-economic model of bioenergy and biofuels production in California. J. Transp. Geogr. 18, 715–728. https://doi.org/10.1016/j.jtrangeo.2010.06.005

Tittmann, P. W., Parker, N.C., Hart, Q.J., Jenkins, B.M., 2010. A spatially explicit techno-economic model of bioenergy and biofuels production in California. J. Transp. Geogr., Special Section on Alternative Fuels and Vehicles 18, 715–728. https://doi.org/10.1016/j.jtrangeo.2010.06.005

Tol, R., 2011. The Social Cost of Carbon. Annu. Rev. Resour. Econ. 3, 419–443. https://doi.org/10.1146/annurev-resource-083110-120028

UChicago Argonne, 2015. The Greenhouse gases, Regulated Emissions, and Energy use in Transportation Model. Argonne National Laboratory.

UN, 2017. World Population Prospects The 2017 Revision (Key Findings and Advance Tables No. ESA/P/WP/248). Department of Economic and Social Affairs, Population Division, New York.

U.S. Congress, 1990. Food, Agriculture, Conservation, and Trade Act of 1990. Public Law 101–624: U.S. Farm Bill.

US EPA, 2014. US EPA ANNEX 3 Methodological Descriptions for Additional Source or Sink Categories, DRAFT Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2012. U.S. Environmental Protection Agency, Washington DC.

US EPA, C.C.D., 2015. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013 (Collections & List, No. 430- R-14– 003).

US EPA, O., 2016. Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866 (May 2013, Revised August 2016) [WWW Document]. URL https://www.epa.gov/climatechange/social-cost-carbon (accessed 3.8.17).

USDA-ERS, 1997. Cropping Practices Survey Data—1995.

USDA-NASS, 2004. Published Estimates Database.

Vadas, P.A., Digman, M.F., 2013. Production costs of potential corn stover harvest and storage systems. Biomass Bioenergy 54, 133–139. https://doi.org/10.1016/j.biombioe.2013.03.028

Van Grinsven, H.J.M., Holland, M., Jacobsen, B.H., Klimont, Z., Sutton, M. a., Jaap Willems, W., 2013. Costs and Benefits of Nitrogen for Europe and Implications for Mitigation. Environ. Sci. Technol. 47, 3571–3579. https://doi.org/10.1021/es303804g

Villa, F., Bagstad, K.J., Voigt, B., Johnson, G.W., Portela, R., Honzák, M., Batker, D., 2014. A Methodology for Adaptable and Robust Ecosystem Services Assessment. PLoS ONE 9, e91001. https://doi.org/10.1371/journal.pone.0091001

Villa-Vialaneix, N., Follador, M., Ratto, M., Leip, A., 2012. A comparison of eight metamodeling techniques for the simulation of N2O fluxes and N leaching from corn crops. Environ. Model. Softw., Emulation techniques for the reduction and sensitivity

analysis of complex environmental models 34, 51–66. https://doi.org/10.1016/j.envsoft.2011.05.003

Volk, M., 2013. Modelling ecosystem services – Challenges and promising future directions. Sustain. Water Qual. Ecol. 1–2, 3–9. https://doi.org/10.1016/j.swaqe.2014.05.003

Wang, M., Han, J., Dunn, J.B., Cai, H., Elgowainy, A., 2012. Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for US use. Environ. Res. Lett. 7, 045905. https://doi.org/10.1088/1748-9326/7/4/045905

Wang, Z., Dunn, J.B., Han, J., Wang, M.Q., 2015. Influence of corn oil recovery on life-cycle greenhouse gas emissions of corn ethanol and corn oil biodiesel. Biotechnol. Biofuels 8, 178. https://doi.org/10.1186/s13068-015-0350-8

Water Conservation Board, 2010. South Platte SWSI 2010 Basin Report. Colo. Dep. Nat. Resour. 2.

Wierzbicki, A.P., Makowski, M., Wessels, J., 2000. Model-Based Decision Support Methodology with Environmental Applications. Springer Netherlands.

Wilhelm, W.W., Johnson, J.M.F., Karlen, D.L., Lightle, D.T., 2007. Corn Stover to Sustain Soil Organic Carbon Further Constrains Biomass Supply. Agron. J. 99, 1665–1667. https://doi.org/10.2134/agronj2007.0150

Wilson, B., Cappelleri, D., Simpson, T.W., Frecker, M., 2001. Efficient Pareto Frontier Exploration using Surrogate Approximations. Optim. Eng. 2, 31–50. https://doi.org/10.1023/A:1011818803494

Wu, X., Zheng, Y., Wu, B., Tian, Y., Han, F., Zheng, C., 2016. Optimizing conjunctive use of surface water and groundwater for irrigation to address human-nature water conflicts: A

surrogate modeling approach. Agric. Water Manag. 163, 380–392. https://doi.org/10.1016/j.agwat.2015.08.022

You, F., Tao, L., Graziano, D.J., Snyder, S.W., 2012. Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input–output analysis. AIChE J. 58, 1157–1180. https://doi.org/10.1002/aic.12637

Yu, T.E., Wang, Z., English, B.C., Larson, J.A., 2014. Designing a dedicated energy crop supply system in Tennessee: a multiobjective optimization analysis. J. Agric. Appl. Econ. 46, 357.

Yue, D., Kim, M.A., You, F., 2013. Design of Sustainable Product Systems and Supply Chains with Life Cycle Optimization Based on Functional Unit: General Modeling Framework, Mixed-Integer Nonlinear Programming Algorithms and Case Study on Hydrocarbon Biofuels. ACS Sustain. Chem. Eng. 1, 1003–1014. https://doi.org/10.1021/sc400080x

Yue, D., You, F., Snyder, S.W., 2014. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. Comput. Chem. Eng., Selected papers from ESCAPE-23 (European Symposium on Computer Aided Process Engineering - 23), 9-12 June 2013, Lappeenranta, Finland 66, 36–56. https://doi.org/10.1016/j.compchemeng.2013.11.016

Zamboni, A., Bezzo, F., Shah, N., 2009a. Spatially Explicit Static Model for the Strategic Design of Future Bioethanol Production Systems. 2. Multi-Objective Environmental Optimization. Energy Fuels 23, 5134–5143. https://doi.org/10.1021/ef9004779

Zamboni, A., Shah, N., Bezzo, F., 2009b. Spatially Explicit Static Model for the Strategic Design of Future Bioethanol Production Systems. 1. Cost Minimization. Energy Fuels 23, 5121–5133. https://doi.org/10.1021/ef900456w

Zemek, O., 2009. Biomass and Carbon stocks Inventory of Perennial vegetation in the Chieng Khoi Watershed, NW Vietnam. Dep. Agroecol. Plant Prod. Trop. Subtrop. Univ. Hohenh. M.Sc. Thesis,.

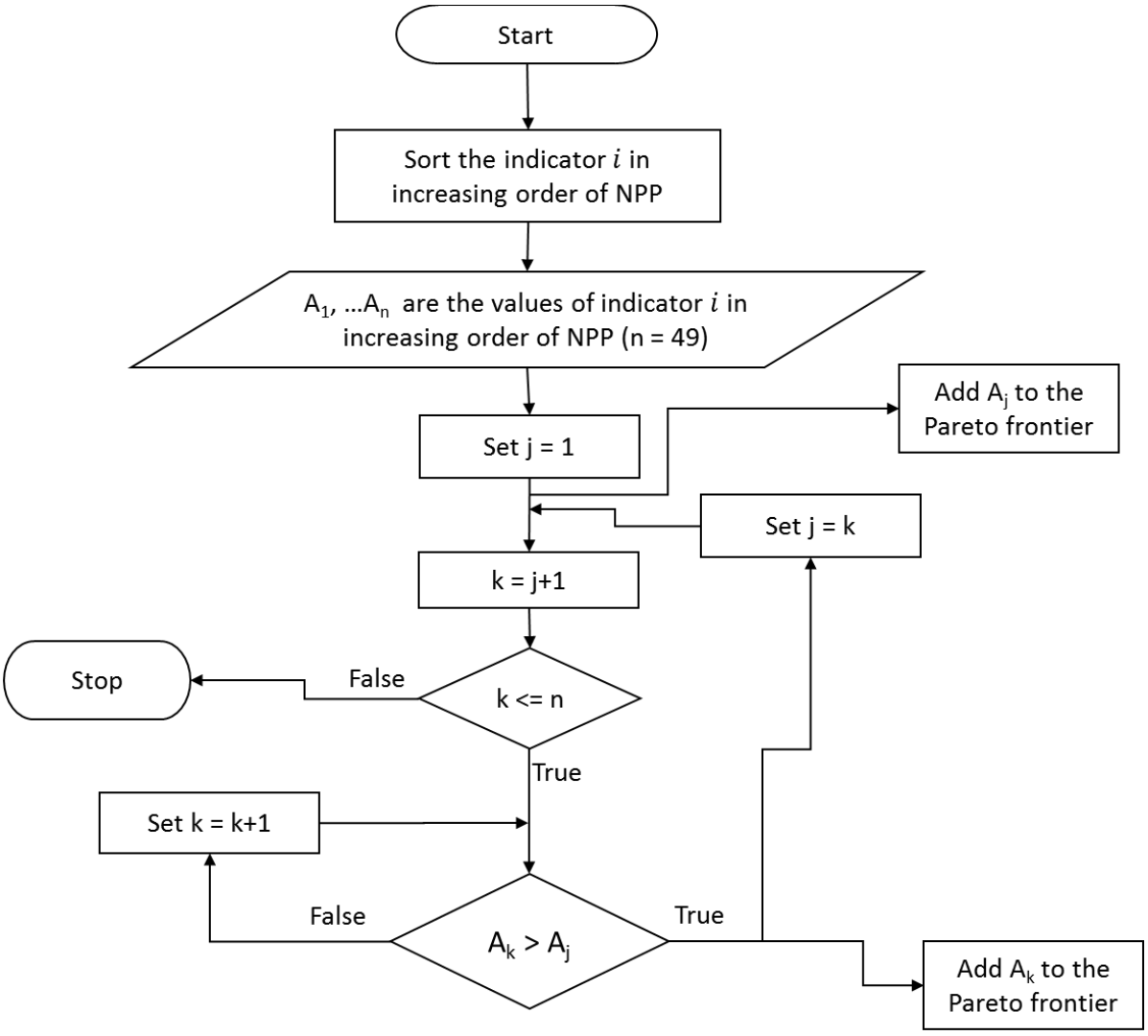
Zhang, W., Ricketts, T.H., Kremen, C., Carney, K., Swinton, S.M., 2007. Ecosystem services and dis-services to agriculture. Ecol. Econ. 64, 253–260. https://doi.org/10.1016/j.ecolecon.2007.02.024

Zhang, X., Izaurralde, R. c., Manowitz, D., West, T.O., Post, W.M., Thomson, A.M., Bandaru, V.P., Nichols, J., Williams, J.R., 2010. An integrative modeling framework to evaluate the productivity and sustainability of biofuel crop production systems. GCB Bioenergy 2, 258–277. https://doi.org/10.1111/j.1757-1707.2010.01046.x

Zhang, Y., 2016. Simulating canopy dynamics, productivity and water balance of annual crops from field to regional scales (PhD dissertation). Colorado State University.

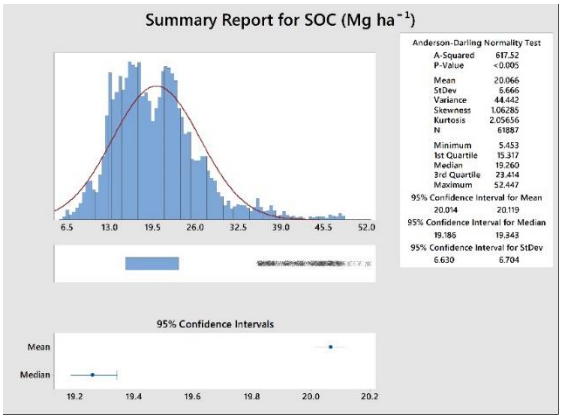
APPENDIX A

A.1 Non-dominated sorting for Pareto frontiers.

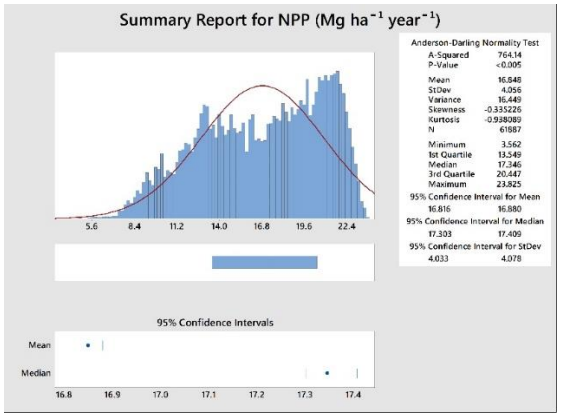


A.2 Summary statistics of ES and EDS indicators.

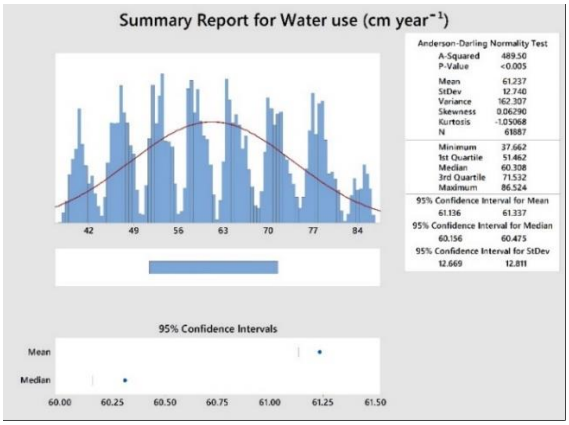
SOC



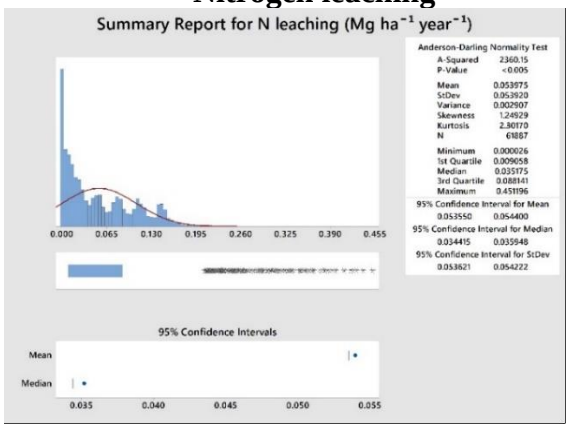
NPP



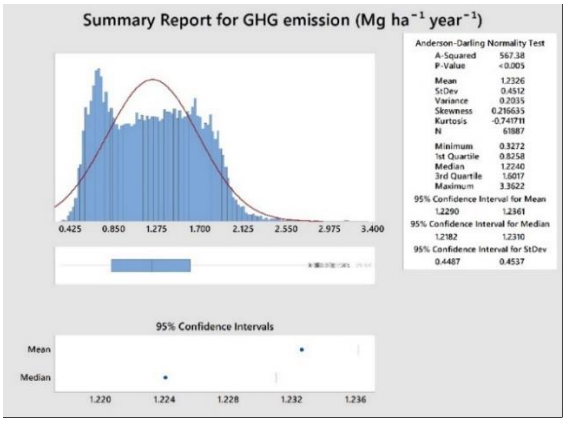
Water use



Nitrogen leaching



GHG



A.3 One-way ANOVA.

Method:

Null hypothesis: All means are equal
Alternative hypothesis: At least one mean is different
Significance level: $\alpha = 0.05$
Equal variances were not assumed for the analysis.

Factor Information:

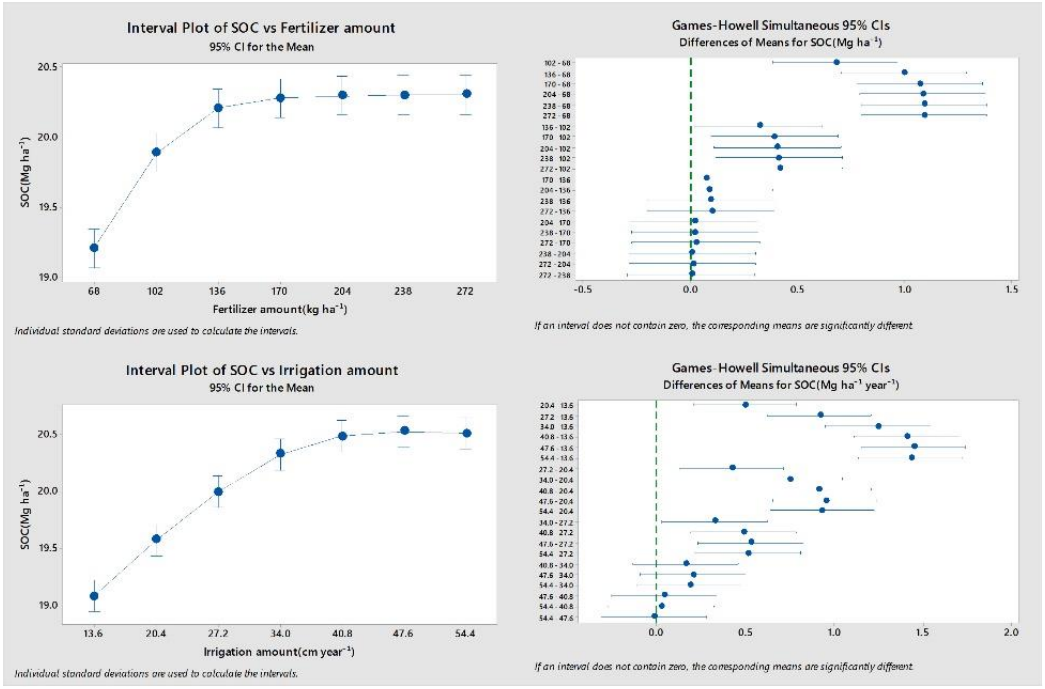
Factor	Levels	Values
Fertilizer amount (kg ha^{-1})	7	68, 102, 136, 170, 204, 238, 272
Irrigation amount (cm year^{-1})	7	13.6, 20.4, 27.2, 34.0, 40.8, 47.6, 54.4

Welch's Tests

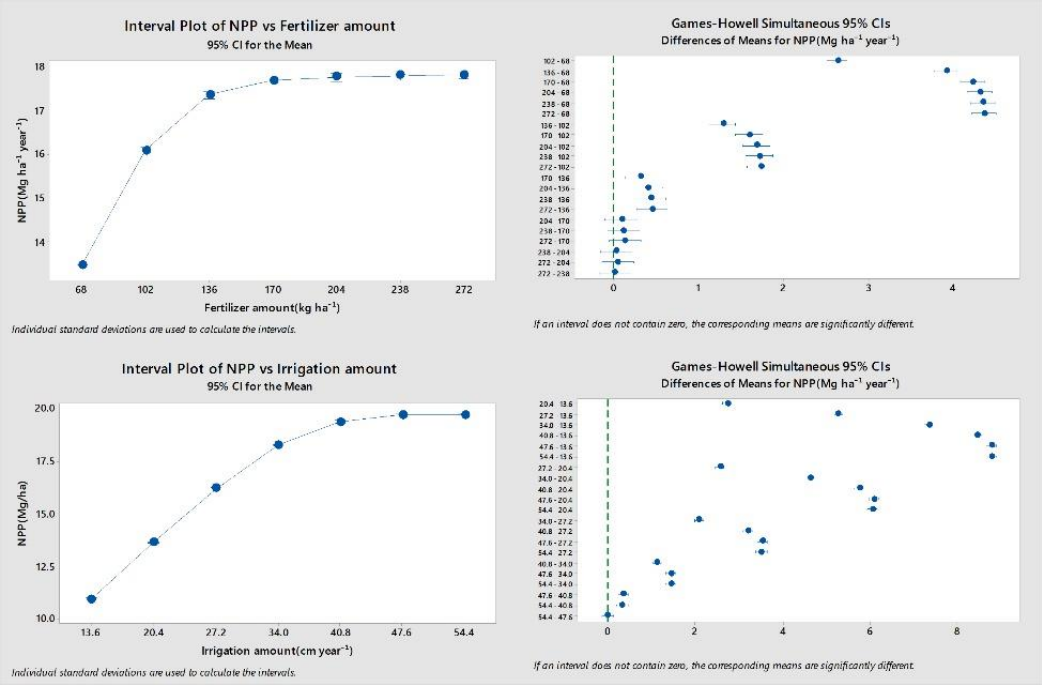
ES/EDS indicator	Fertilizer	Irrigation
SOC	$F(6, 27501) = 34.12$, $p < 0.000$	$F(6, 27502) = 63.02$, $p < 0.000$
NPP	$F(6, 27168) = 3414.91$, $p < 0.000$	$F(6, 27434) = 18946.82$, $p < 0.000$
Water use	$F(6, 27502) = 0.00$, $p = 1.000$	$F(6, 27495) = 149594.60$, $p < 0.000$
N leaching	$F(6, 27289) = 6666.53$, $p < 0.000$	$F(6, 27397) = 2231.40$, $p < 0.000$
GHG	$F(6, 27488) = 48677.28$, $p < 0.000$	$F(6, 27477) = 171.85$, $p < 0.000$

Interval plots (left) and Post-hoc comparison using Games-Howell method (right)

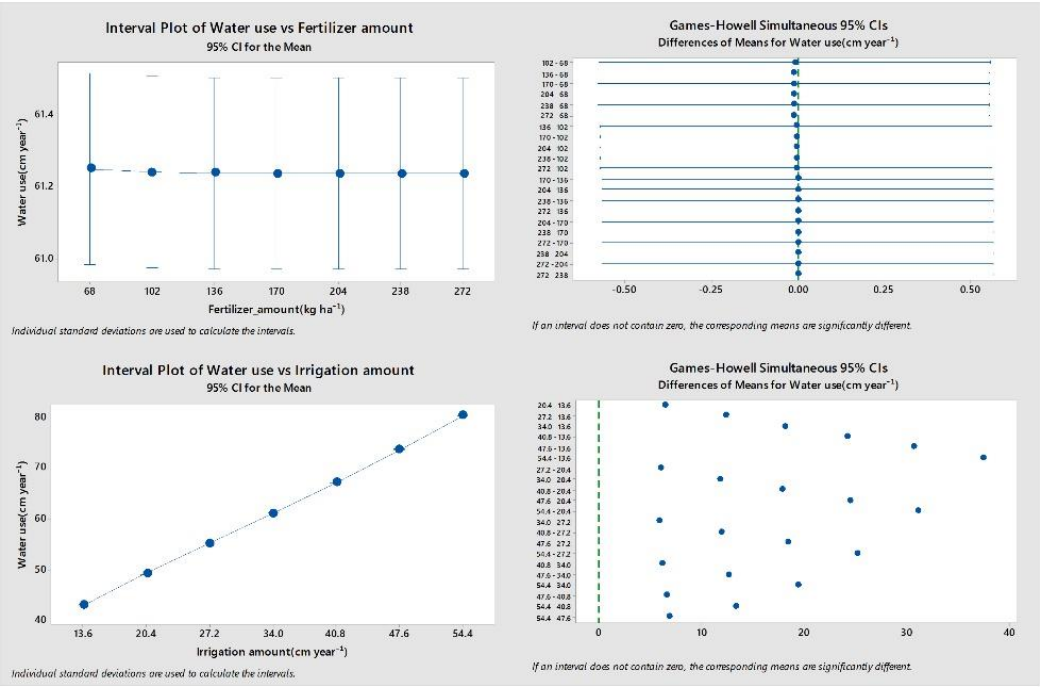
SOC:



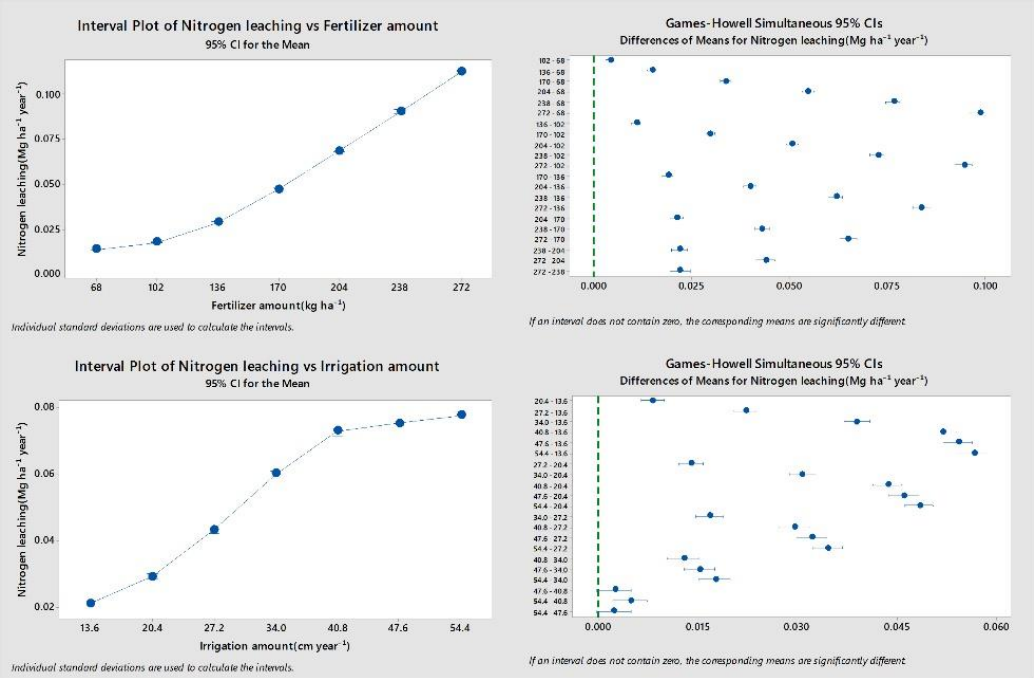
NPP:



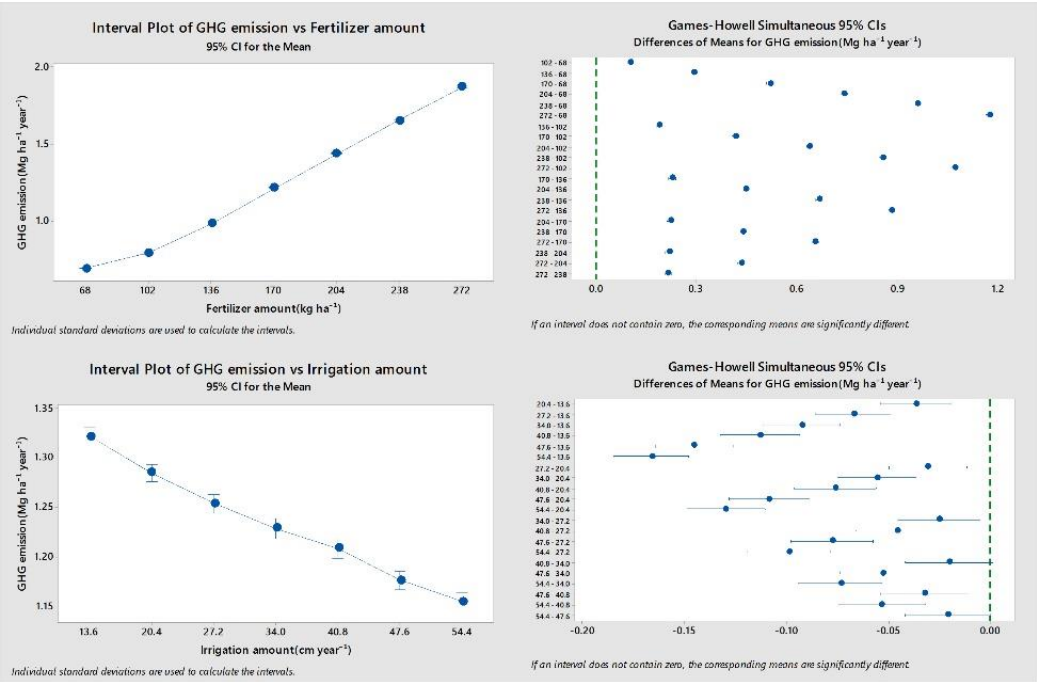
Water use:



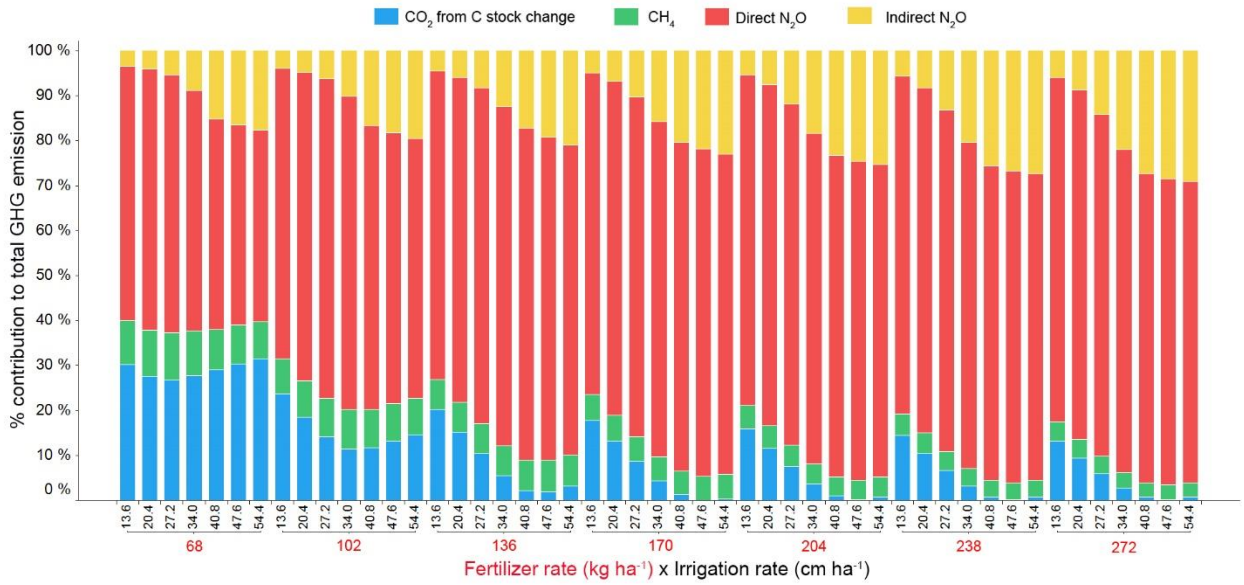
Nitrogen leaching:



GHG emissions:



A.4 Contribution of GHG components to total GHG emissions.



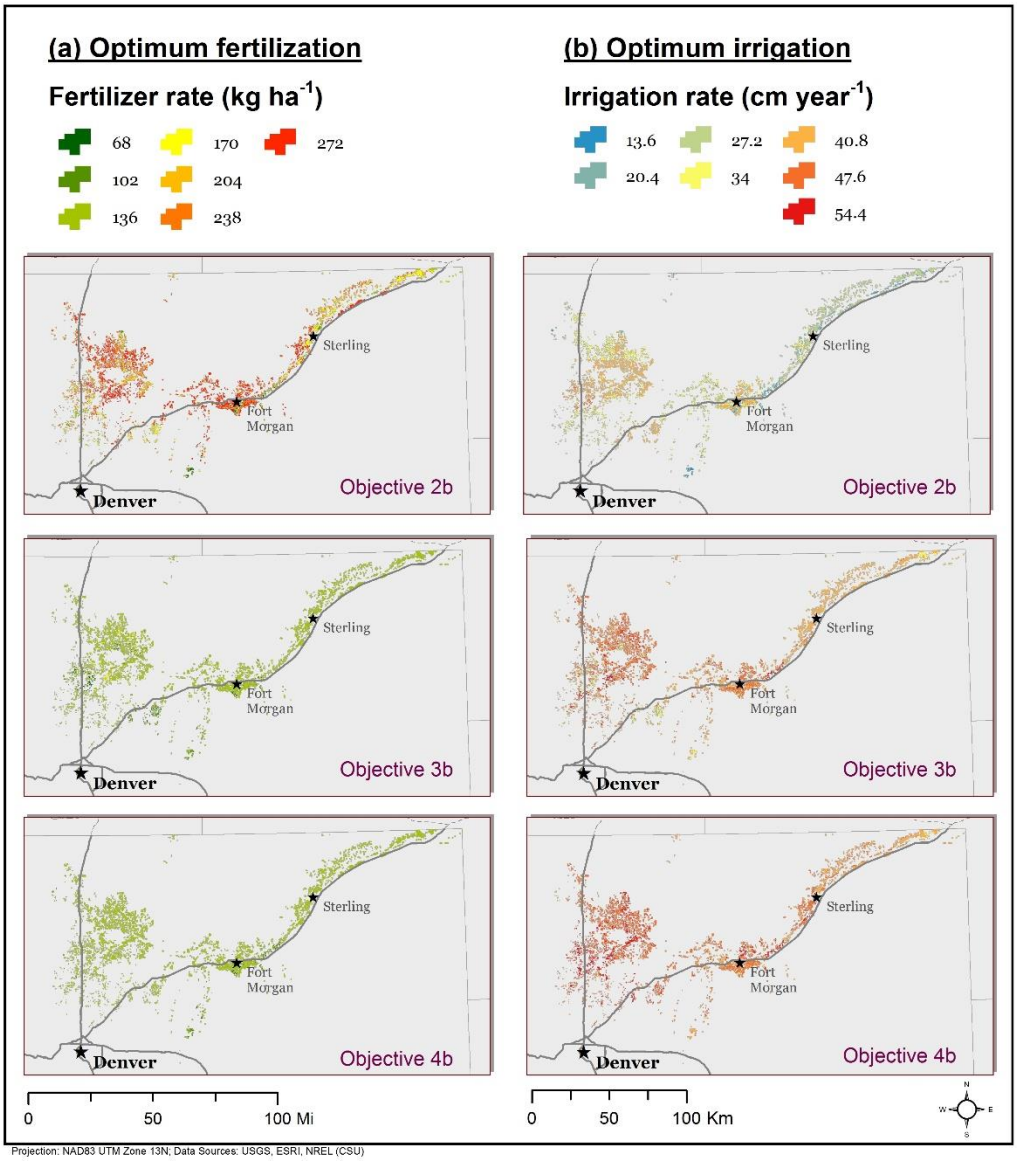
A.5 Optimum fertilizer and irrigation rates and EDS footprints.

A.5.1 Color representation using Jenk natural breaks classifications.

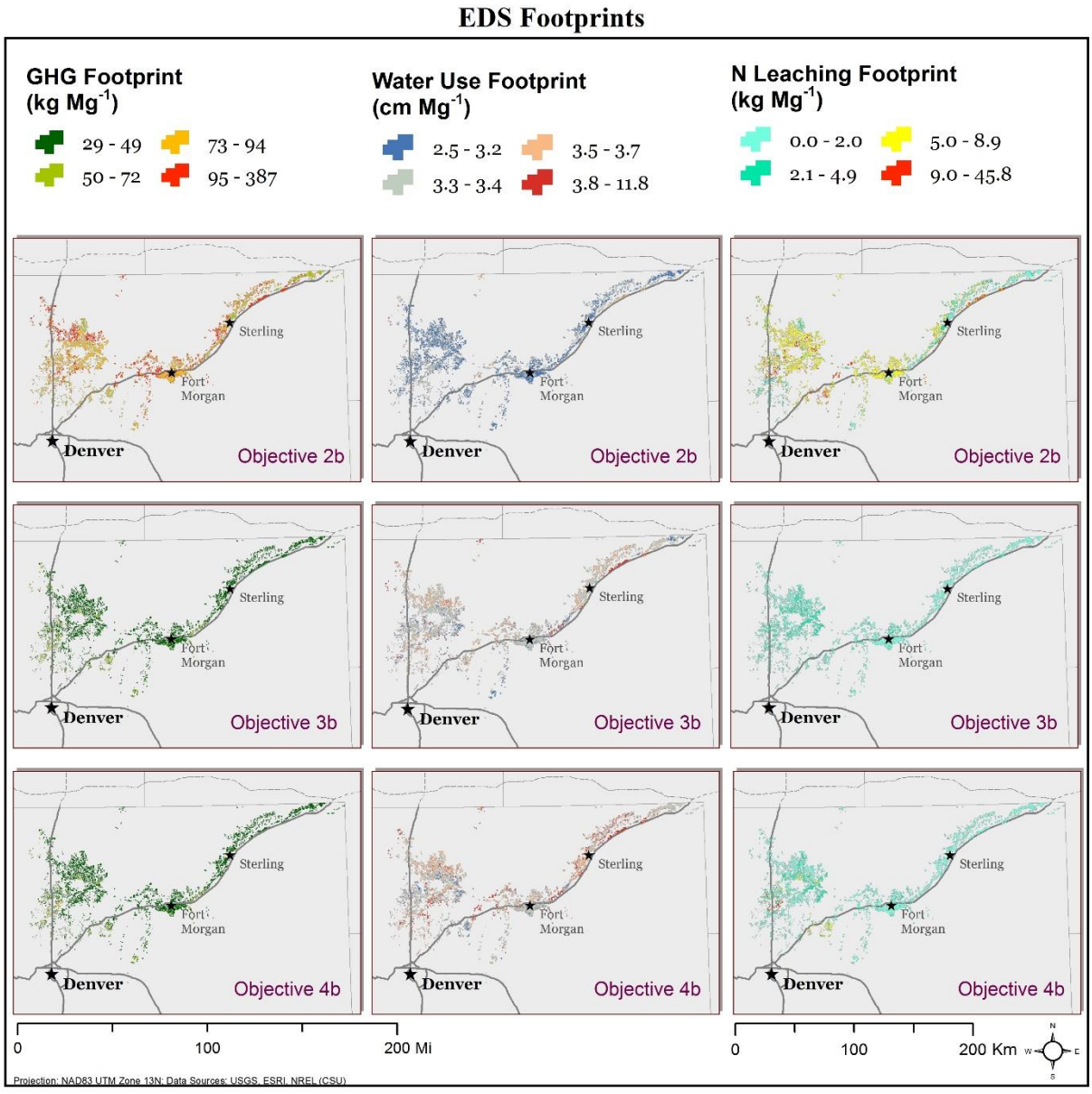
For better visualization and comparison of continuous data (e.g., EDS footprints) on a series of maps, we used the Jenk natural breaks classification method (color scaling) which seeks to reduce the variance within classes and maximize the variance between classes (Jenks, 1967). Jenk natural breaks classifications were performed on EDS footprint data using the PySAL library in Python (https://pypi.python.org/pypi/PySAL).

For scenario 1 and 'a' scenarios	For 'b' scenarios
Method: Python library: pysal.esda.mapclassify Population of the classification: 66,604 (16,651 polygons x 4 scenarios) Initial solution generation: 100 Target class: 4	**Method:** Python library: pysal.esda.mapclassify Population of the classification: 49,953 (16,651 polygons x 3 scenarios) Initial solution generation: 100 Target class: 4
Natural breaks of GHG footprint (kg CO_2 Mg^{-1}) Lower Upper Count x[i] <= 37.569 14987 37.569 < x[i] <= 43.294 27593 43.294 < x[i] <= 54.447 18204 54.447 < x[i] <= 129.141 5820	**Natural breaks of GHG footprint (kg CO_2 Mg^{-1})** Lower Upper Count x[i] <= 49.129 29956 49.129 < x[i] <= 71.536 5306 71.536 < x[i] <= 93.923 8102 93.923 < x[i] <= 387.061 6589
Natural breaks of N leaching footprint (kg N Mg^{-1}) Lower Upper Count x[i] <= 0.397 22413 0.397 < x[i] <= 1.006 26179 1.006 < x[i] <= 2.031 13211 2.031 < x[i] <= 8.231 4801	**Natural breaks of N leaching footprint (kg N Mg^{-1})** Lower Upper Count x[i] <= 2.031 28523 2.031 < x[i] <= 4.858 8379 4.858 < x[i] <= 8.888 11274 8.888 < x[i] <= 45.754 1777
Natural breaks of Water footprint (cm Mg^{-1}) Lower Upper Count x[i] <= 3.132 19484 3.132 < x[i] <= 3.338 24044 3.338 < x[i] <= 3.675 18873 3.675 < x[i] <= 10.371 4203	**Natural breaks of Water footprint (cm Mg^{-1})** Lower Upper Count x[i] <= 3.154 15544 3.154 < x[i] <= 3.376 16518 3.376 < x[i] <= 3.662 12861 3.662 < x[i] <= 11.816 5030

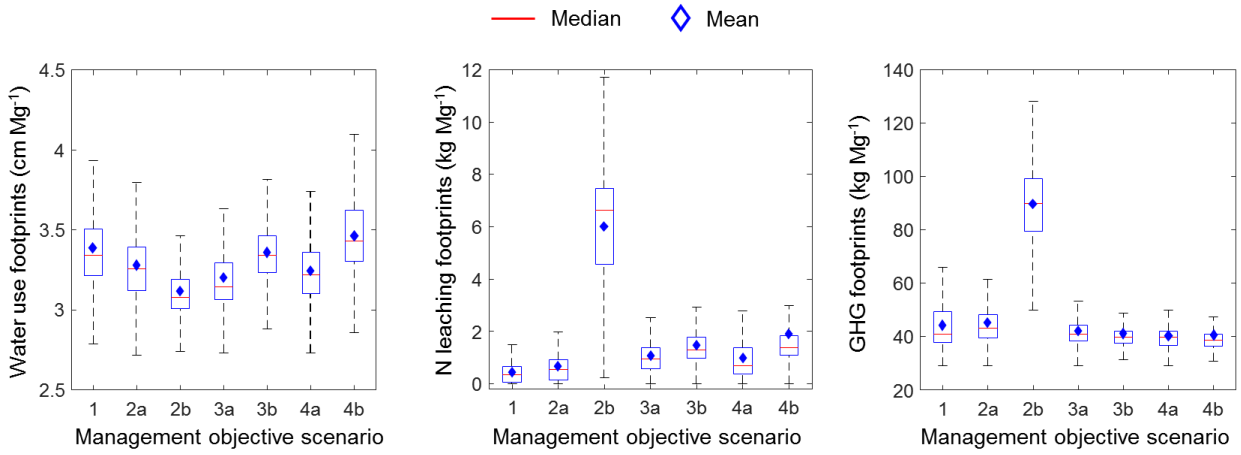
A.5.2 Optimum fertilizer and irrigation rates for "b" scenarios.



A.5.3 Distribution of EDS footprints for "b" scenarios.



A.5.4 Boxplots of EDS footprints.



In general, average N leaching footprints tended to be lower for equal optimization of all indicators (scenario 1) or constrained optimization of NPP and water use (scenario 2a). Lower average water use footprints were obtained under unconstrained optimization of NPP and water use (scenario 2b) and constrained optimization of NPP and N leaching (scenario 3a), whereas lower average GHG footprints were achieved when optimizing for NPP and GHG emissions (scenarios 4a and 4b).

APPENDIX B

B.1 Supply chain budgeting parameters.

Item	Unit	Conventional tillage	Reduced tillage	No-till
Fixed costs				
Pre-harvest machinery	$ ha^{-1}	224.8	184.2	115.6
Seeds	$ ha^{-1}	281.1	281.1	281.1
Other fertilizers and Lime	$ ha^{-1}	117.4	117.4	117.4
Herbicide and Insecticide	$ ha^{-1}	46.2	47.8	67.7
Non-Machinery labor	$ ha^{-1}	35.0	35.0	35.0
Irrigation Fixed cost	$ ha^{-1}	228.7	228.7	228.7
Miscellaneous	$ ha^{-1}	25.4	25.4	25.4
Fix costs total	$ ha^{-1}	**958.7**	**919.7**	**871.1**
Variable costs				
Fertilizers	$ kg^{-1}		1.27	
Irrigation Variable cost	$ cm^{-1}		1.31	
Corn Drying	$ Mg^{-1}		5.26	
Corn harvest cost	$ Mg^{-1}		11.73	
Corn insurance	$ Mg^{-1}		3.29	
Stover harvest cost	$ Mg^{-1}		15.43	
Stover storage cost	$ Mg^{-1}		16.46	
Interest on 1/2 non-land costs	%		6.5%	
Land- cash rent equivalent	% of total revenue		20%	
Corn hauling per mile	$ Mg^{-1} $mile^{-1}$		0.20	
Stover loading and unloading cost	$ Mg^{-1}		5.87	
Stover transport cost per mile	$ Mg^{-1} $mile^{-1}$		0.37	

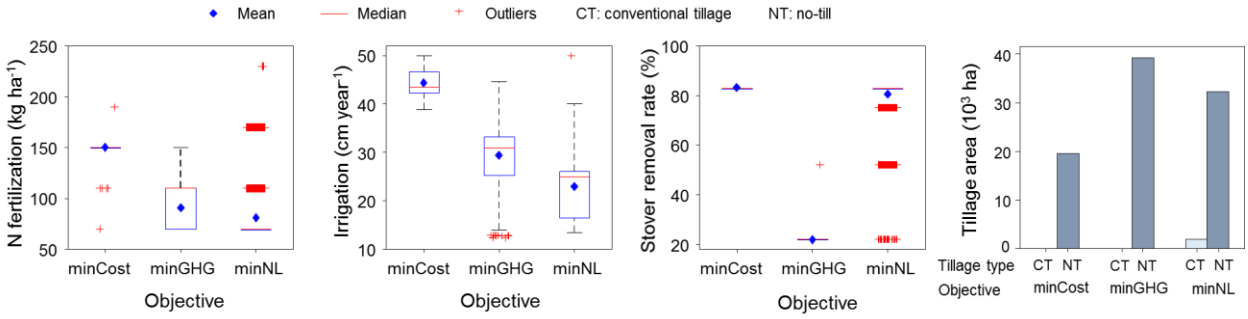
B.2 Energy use for farm operations.

Machinery item	Average diesel use (L ha^{-1})	Energy* (Btu ha^{-1})
Moldboard plowing (MP)	34.2	1,281,771
Chisel plowing (CP)	12.0	448,901
Planting for tilled system (PT)	6.9	259,920
Planting for no-till system (PN)	9.8	368,204
Fertilization (F)	8.0	301,957
Herbicide and insecticide application for tilled system (HT)	2.0	75,442
Herbicide and insecticide application for no-till system (HN)	4.0	150,885
Grain harvest (G)	9.9	371,582
Stover harvest (S)	9.0	337,802
Total farm energy use		
Conventional tillage	70.0	2,628,475
Reduced tillage	47.8	1,795,606
No-till	40.8	1,530,431

Note: Conventional tillage = MP + PT + F + HT + G + S; Reduced tillage = CP + PT + F + HT + G + S; and No-till = PN + F + HN + G + S.

*1 L diesel = 37,533.56 Btu

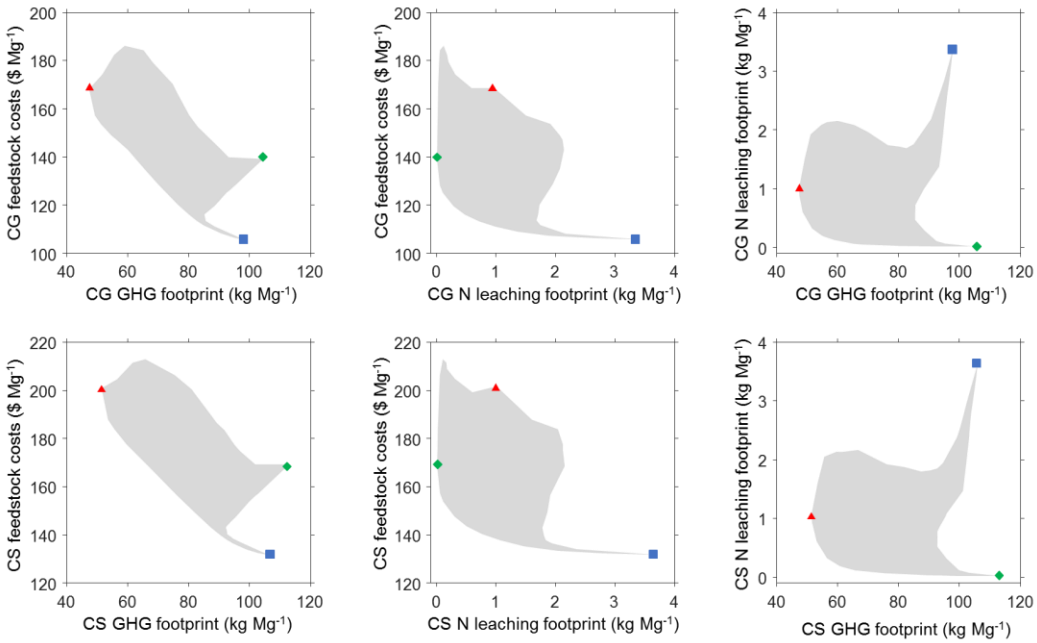
B.3 Optimum management inputs for the corner solutions.



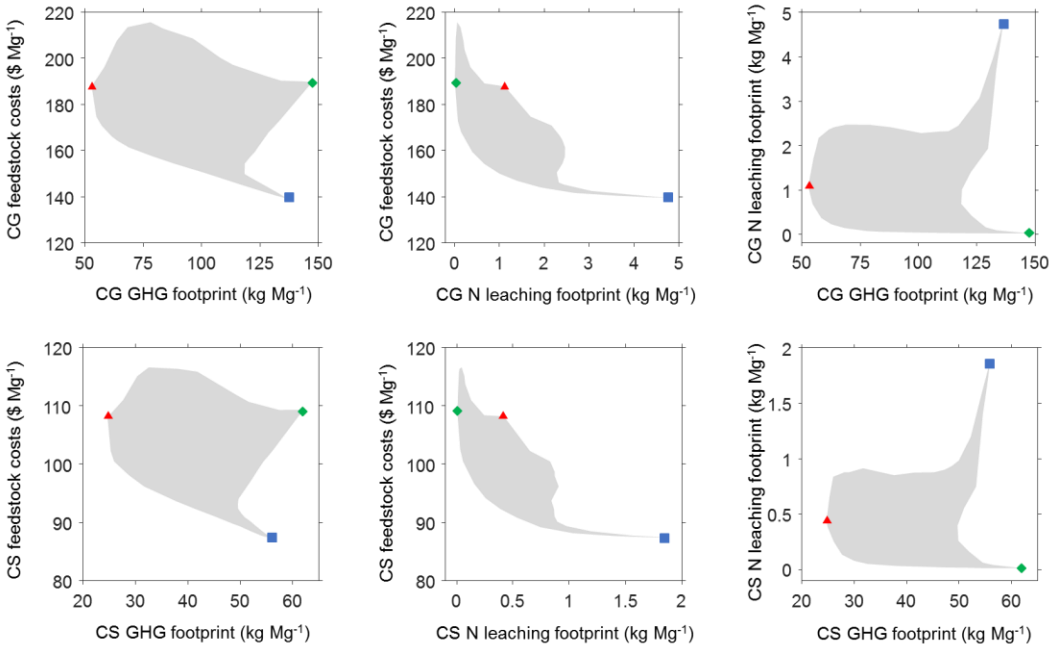
B.4 Footprints of production costs, GHG emissions, and N leaching.

The average feedstock costs and GHG and N leaching footprints of a feedstock were calculated by dividing the corresponding objective value (i.e., production costs, GHG emissions, and N leaching) attributed to the production of that feedstock over the total amount of the feedstock used for ethanol production. The production costs, GHG emissions, and N leaching were attributed to corn grain and corn stover using process-level allocation method based on energy contents or market values. The energy contents in terms of lower heating values (LHV) used for corn grain and corn stover in our study were 15.01 and 16.22 MBTU Mg^{-1}, respectively. For market value allocation, we used a 10-year averaged U.S. corn grain price of \$180 Mg^{-1} reported by Ycharts (https://ycharts.com/) for the period from 2007 to 2017 and an average stover price of \$70 Mg^{-1} reported by Thompson and Tyner, (2011).

(a) Energy content allocation.



(b) Market value allocation.



Legend

Point by: ■ minCost ▲ minGHG ◆ minNL

Fill by: Feasible Regions

Abbreviation by: CG: Corn grain CS: Corn Stover

APPENDIX C

C.1 Input variables used for metamodeling.

Table C.1.1. Bounds and statistics of input variables.

Input	Management-specific input variable bounds	Site-specific input variable statistics			
		Min	Max	Median	StdDev
Fertilizer (kg N ha^{-1})	[0, 300]				
Irrigation (cm $year^{-1}$)	[0, 70]				
Residue removal rate (%)	[0, 90]				
Number of soil layer		5	14	13	1.22
Bulk density (g cm^{-3})		1.21	1.85	1.42	0.13
Sand (%)		5.7	96.0	47.7	24.4
Silt (%)		2.0	70.3	29.0	17.6
BAU's soil organic carbon (Mg C ha^{-1})		7.6	48.9	22.3	6.2
BAU's soil nitrogen (Mg N ha^{-1})		0.9	11.2	3.0	1.4
30-year-average precipitation (cm $year^{-1}$)		24.2	38.0	33.4	3.8
30-year-average maximum temperature (^{o}C)		22.5	28.2	25.2	1.1
30-year-average minimum temperature (^{o}C)		11.1	14.6	11.9	0.8

Note: Values between brackets are the minimum and maximum

C.2 Calculation of objective and constraint metrics.

Farm profits were calculated as the difference between the revenues from corn grain and corn stover and the corresponding farm production costs. To calculate revenues, we used a 10-year averaged U.S. corn grain price of \$180 Mg^{-1} reported by Ycharts (https://ycharts.com/) for the period from 2007 to 2017 and an average stover price of \$70 Mg^{-1} reported by (Thompson and Tyner, 2011). Farm production costs included tillage-based fixed costs and variable costs of N fertilization and irrigation, land rent, and feedstock-dependent costs for harvest, grain drying, storage, and handling. The costs were calculated based on the economic data obtained from different sources (e.g., (Thompson and Tyner, 2011; Vadas and Digman, 2013; Russell et al.,

2016; Ibendahl et al., 2015) (Table C.2.1). All the monetary values acquired from the literature were adjusted for inflation to 2017 US dollars.

The values of soil organic carbon (SOC) and nitrogen (N) leaching were taken as direct DayCent outputs while greenhouse gas (GHG) emissions were calculated from soil-related emissions and on-farm machinery emissions. Soil GHG emission estimates included modeled outputs from DayCent (annual values of net CO_2 emission, CH_4, N_2O emission, NOx efflux, volatilized NH_3, and N leaching) as well as additional models (e.g., for indirect N_2O emissions) from the U.S. Department of Agriculture entity-scale greenhouse gas inventory guidelines (Eve et al. 2014). The on-farm machinery emissions were calculated using the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation Model (GREET) (UChicago Argonne, 2015) based on the avereage farm energy use for different tillage types in the region (Table C.2.2) and the GREET's default emission factors (Table C.2.3).

Table C.2.1. Economic parameters for farm production costs calculation

Item	Unit	Conventional tillage	No-till
Fixed costs			
Pre-harvest machinery	$ ha^{-1}	224.8	115.6
Seeds	$ ha^{-1}	281.1	281.1
Other fertilizers and Lime	$ ha^{-1}	117.4	117.4
Herbicide and Insecticide	$ ha^{-1}	46.2	67.7
Non-Machinery labor	$ ha^{-1}	35.0	35.0
Irrigation Fixed cost	$ ha^{-1}	228.7	228.7
Miscellaneous	$ ha^{-1}	25.4	25.4
Fix costs total	$ ha^{-1}	**958.7**	**871.1**
Variable costs			
Fertilizers	$ kg^{-1}	1.27	
Irrigation Variable cost	$ cm^{-1}	1.31	

Corn Drying	$ Mg^{-1}	5.26
Corn harvest cost	$ Mg^{-1}	11.73
Corn insurance	$ Mg^{-1}	3.29
Stover harvest cost	$ Mg^{-1}	15.43
Stover storage cost	$ Mg^{-1}	16.46
Interest on 1/2 non-land costs	%	6.5%
Land-cash rent equivalent	% of total revenue	20%

Table C.2.2. Energy use for farm operations

Machinery item	**Average diesel use (L ha^{-1})**	**Energy* (Btu ha^{-1})**
Moldboard plowing (MP)	34.2	1,281,771
Chisel plowing (CP)	12.0	448,901
Planting for tilled system (PT)	6.9	259,920
Planting for no-till system (PN)	9.8	368,204
Fertilization (F)	8.0	301,957
Herbicide and insecticide application for tilled system (HT)	2.0	75,442
Herbicide and insecticide application for no-till system (HN)	4.0	150,885
Grain harvest (G)	9.9	371,582
Stover harvest (S)	9.0	337,802
Total farm energy use		
Conventional tillage	70.0	2,628,475
No-till	40.8	1,530,431

Note: Conventional tillage = MP + PT + F + HT + G + S; and No-till = PN + F + HN + G + S.

*1 L diesel = 37,533.56 Btu

The energy use for irrigation was 107,242 Btu cm^{-1}, calculated based on the method presented in (Martin et al., 2010) assuming an average pumping lift of 80-feet and a pump discharge pressure of 45 pounds per square inch.

Table C.2.3. GREET's emission factors for farm operation

Emission source	**Unit**	**Emission factor**
Farm energy	g CO_2e MJ^{-1}	90
N fertilizer	g CO_2e g^{-1}	3.9

P fertilizer	g CO_2e g^{-1}	1.5
Herbicide	g CO_2e g^{-1}	20
Insecticide	g CO_2e g^{-1}	23
Storage and handling of stover	g CO_2e Mg^{-1}	1167

Table C.2.4. Extra farm management inputs

Item	Unit	Conventional tillage	No-till
Phophorus application rate	kg ha^{-1}	76.21	76.21
Herbicide	g Mg^{-1}	7	10.29
Insecticide	g Mg^{-1}	0.06	0.09

C.3 Extra statistics of the MLP surrogate used in the optimization.

Table C.3.1. Configuration of weights and biases of the Multilayer perception

		Input Layer													Output Layer		
	Neu-ron	BAU's SOC	Residue removal	Fertilizer	Irrigation	BAU's soil N	Silt	Max temp	Soil layer	Precip	Sand	Bulk density	Min temp	Bias	SOC	Stover	Grain
Hidden Layer	H1	-0.42	0.10	1.53	-0.33	0.72	-0.06	0.04	-0.07	-0.02	-0.09	-0.03	-0.06	2.32	0.96	-1.95	0.81
	H2	1.41	0.00	-0.74	0.76	-2.71	-0.64	0.00	0.14	0.08	-0.59	-0.05	0.07	-1.41	-0.71	-0.67	-1.65
	H3	-0.26	0.17	2.72	0.14	0.13	0.04	-0.12	0.07	0.06	0.12	0.02	-0.07	0.90	0.01	0.03	0.14
	H4	0.43	-0.16	0.12	-0.70	-0.33	1.58	-0.03	0.17	0.20	2.50	1.08	-0.13	-3.38	-0.42	-0.52	-0.66
	H5	0.64	-0.01	0.69	2.46	-1.10	0.16	-0.80	0.84	0.38	-0.56	0.20	0.03	-1.02	0.01	0.03	0.12
	H6	0.36	0.40	0.01	0.76	-0.29	-0.21	-0.06	0.16	0.05	-0.11	0.05	0.11	-1.20	-0.40	1.11	-0.11
	H7	0.68	0.06	-0.62	1.03	-1.23	-0.41	-0.20	0.00	0.21	-0.30	0.12	0.10	-0.74	1.02	0.35	1.63
	H8	-0.80	0.03	0.84	-1.04	0.96	0.35	0.22	0.01	-0.22	0.24	-0.27	-0.12	0.37	1.58	0.62	3.23
	H9	0.89	-0.05	0.33	-0.36	-0.60	-0.30	0.01	0.25	-0.16	-0.27	-0.42	0.00	-1.37	-0.14	-0.11	-0.59
	H10	0.00	-0.04	-0.01	-0.22	0.12	-0.05	0.00	0.11	-0.03	-0.09	0.15	0.19	0.23	0.63	1.72	2.50
	H11	-0.32	0.00	1.45	-1.25	0.68	0.23	0.31	-0.02	-0.20	0.17	-0.13	-0.10	0.12	-0.16	-0.57	-0.79
	H12	-0.03	0.00	-0.10	0.00	0.17	-0.15	0.07	-0.06	-0.11	-0.14	-0.04	0.18	0.08	-1.40	-1.45	-2.55
	H13	-0.33	-0.27	0.04	-0.98	0.01	0.34	0.18	-0.05	-0.16	0.14	-0.14	-0.16	-0.29	0.42	-0.76	0.32
	H14	-0.08	0.08	-0.08	0.23	0.02	-0.03	0.08	0.00	-0.11	0.09	-0.41	0.00	-0.18	1.47	1.35	2.62
	H15	-0.19	-0.02	-0.24	-0.23	0.85	0.16	-0.16	0.82	-0.08	0.56	-0.38	0.09	0.03	-0.04	-0.23	-0.41
	H16	-0.93	0.00	-0.21	-0.31	0.19	1.00	0.07	0.02	-0.17	0.53	-0.59	-0.19	0.53	-1.35	-0.77	-1.44

	H17	-0.22	0.01	0.01	-0.05	-0.01	0.03	0.02	0.02	-0.01	0.06	-0.08	-0.05	-0.24	-3.63	-1.90	-1.91
	H18	-1.75	0.04	0.85	-0.74	2.58	0.70	0.00	-0.10	-0.13	0.62	-0.27	-0.11	1.01	-0.67	-0.75	-1.78
	H19	0.90	0.00	0.24	0.41	-0.25	-1.07	-0.11	-0.02	0.21	-0.58	0.56	0.21	-0.69	-1.21	-0.60	-1.25
	H20	-0.43	0.24	-0.18	-0.72	0.07	0.32	0.11	-0.14	-0.08	0.19	-0.15	-0.19	0.43	-0.09	1.12	-0.18
	H21	-0.72	0.18	0.97	-1.07	0.70	0.30	0.25	-0.02	-0.20	0.19	-0.25	-0.13	0.16	-0.45	0.77	-0.41
	H22	-0.41	-0.14	1.47	-0.36	0.70	0.02	0.04	-0.08	-0.02	-0.04	-0.07	-0.07	2.27	-0.42	3.07	0.53
	H23	-0.64	0.53	-0.01	0.10	0.07	-0.14	-0.02	-0.04	-0.06	-0.01	-0.11	-0.12	0.67	-0.17	0.31	-0.09
	H24	0.39	-0.05	-0.08	2.39	0.09	-0.59	-0.53	0.12	0.33	-0.44	0.36	0.26	1.21	0.05	0.04	0.60
	H25	0.15	-0.08	-1.65	0.35	-1.49	0.19	-0.04	0.09	0.02	0.25	0.14	0.01	-3.13	0.21	0.43	0.30
	Bias														-1.7	-2.03	-2.78

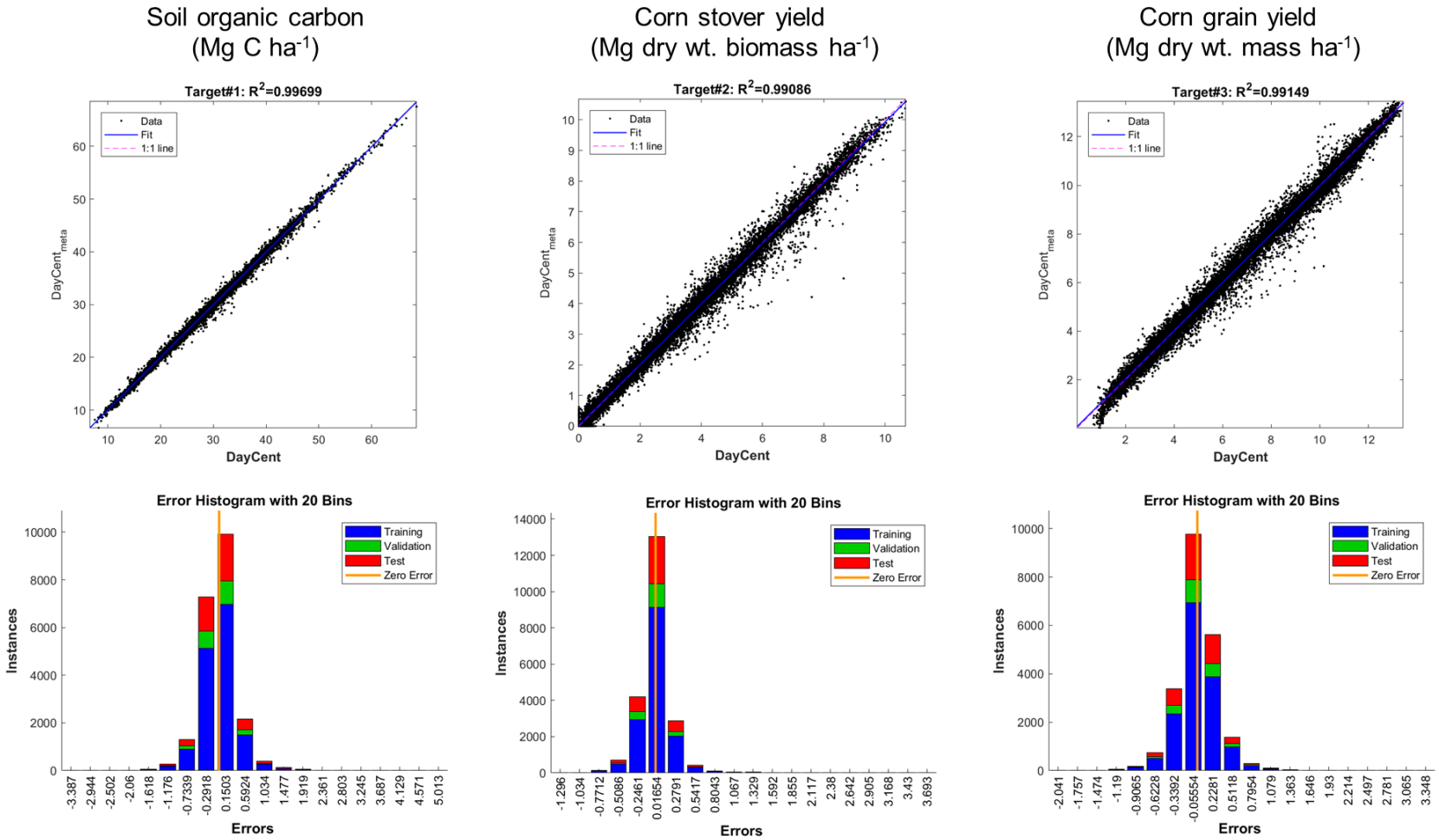


Fig. C.3.1. Regression plots and error histograms of the MLP surrogate model.

C.4 Extra dependence plots of input variables of the MLP surrogate

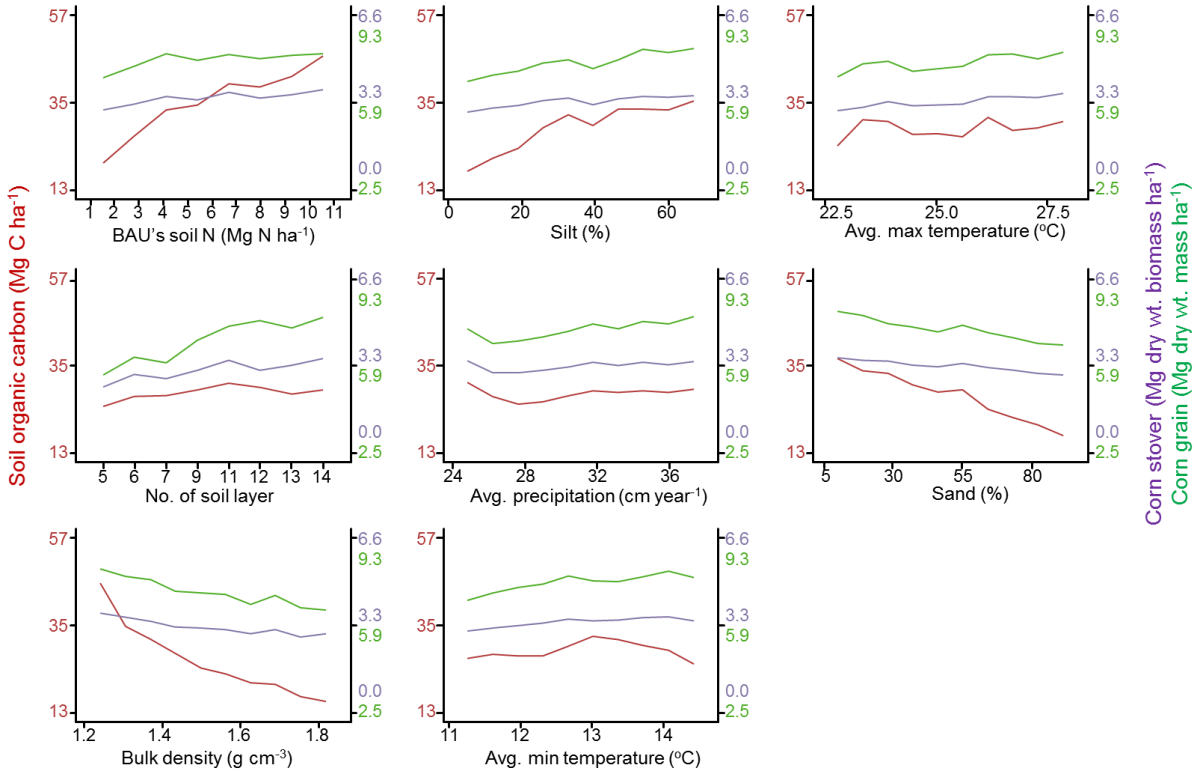


Fig. C.4.1. Dependence plot of the last eight input variables. The dependence plot quantified the ranges of MLP predicted outputs averaged across the entire landscape. The left y-axes were labeled with corn stover yield in Mg dry wt. biomass ha^{-1} (top in purple) and corn grain yield in Mg dry wt. mass ha^{-1} (bottom in green).

C.5 Exploratory analysis of root mean squared errors.

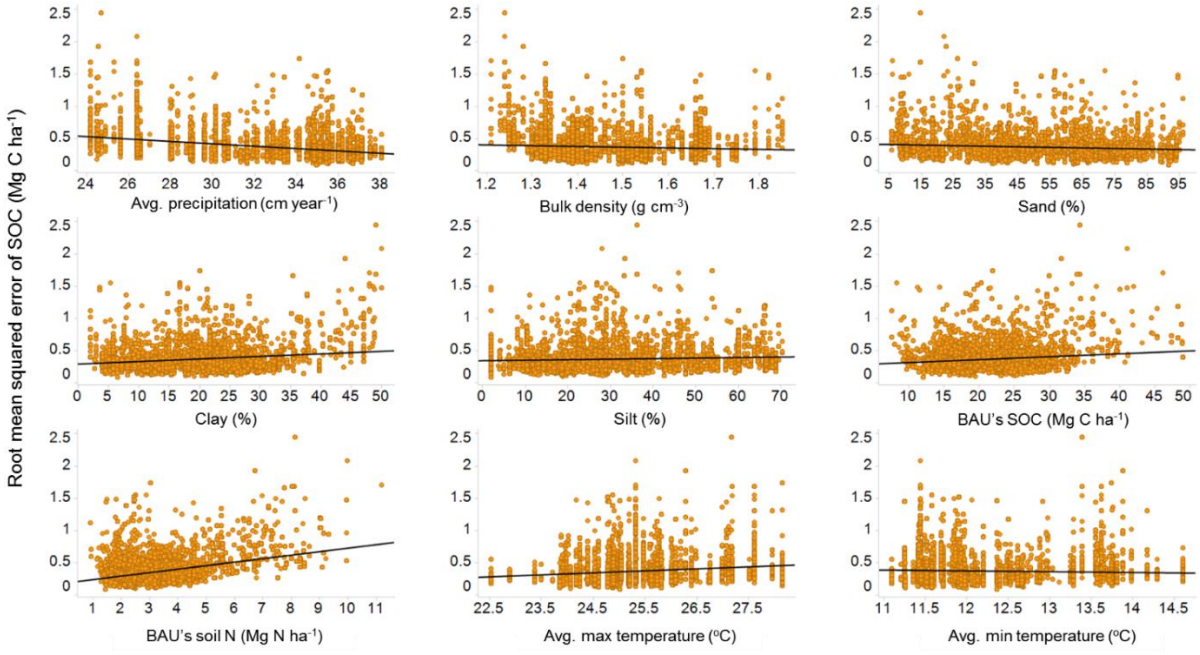


Fig. C.5.1. RMSE of SOC plotted against different site-specific input variables.

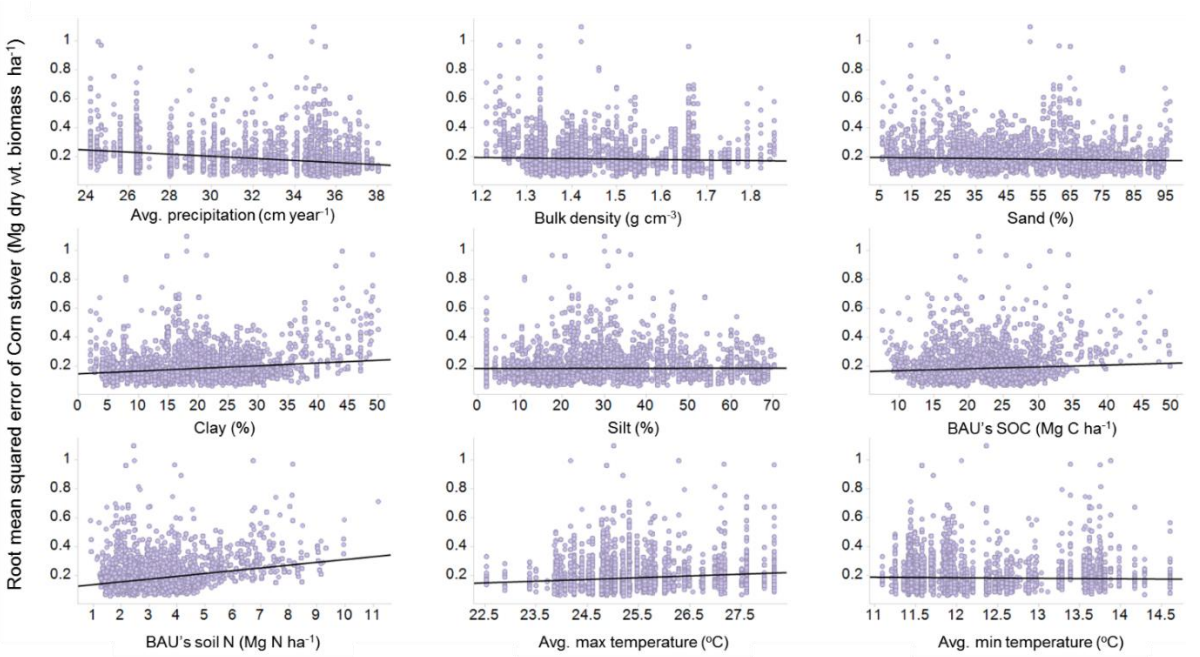


Fig. C.5.2. RMSE of corn stover plotted against different site-specific input variables.

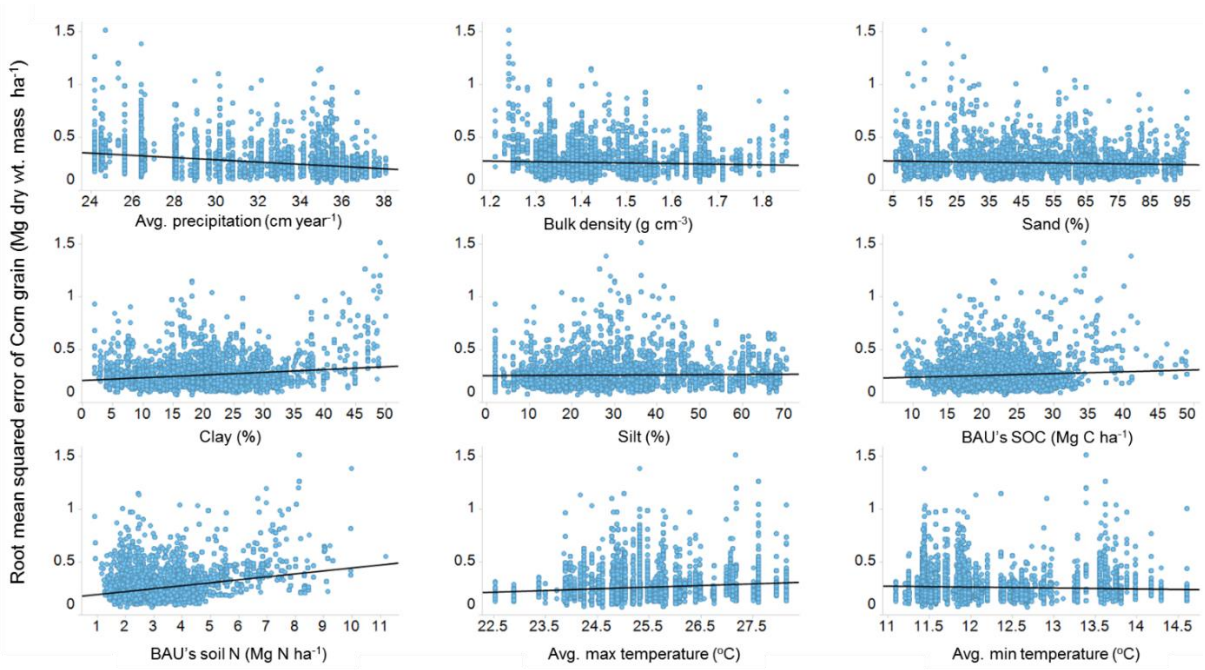


Fig. C.5.3. RMSE of corn grain plotted against different site-specific input variables.

Table C.5.1. The SSURGO soil unit and their corresponding site-specific input variable related to higher RMSE of SOC, stover, and grain

SSURGO soil map unit	$RMSE_{SOC}$	$RMSE_{Stover}$	$RMSE_{Grain}$	# Soil lyr	Bulk density	Sand (%)	Silt (%)	BAU SOC	BAU N	Precip	Max temp	Min temp
94371	2.4	1.0	1.5	13	1.24	15	36	34.3	8.1	24.7	27.2	13.4
497580	2.1	0.6	1.4	13	1.24	22	28	41.0	10.0	26.4	25.3	11.4
95480	1.9	1.0	1.0	13	1.28	23	33	31.6	6.7	24.5	26.3	13.9
94904	1.7	0.7	0.7	14	1.5	26	54	21.9	3.0	34.1	28.2	13.7
95205	1.7	0.7	0.6	13	1.21	6	46	46.1	11.2	26.4	25.3	11.4
94371	1.7	0.8	1.2	13	1.24	15	36	34.3	8.1	25.3	27.2	13.4
94371	1.7	0.7	1.0	13	1.24	15	36	33.9	8.0	24.2	27.6	13.6
497593	1.7	0.4	1.0	13	1.33	31	33	29.9	7.7	26.4	25.3	11.4
95483	1.5	0.3	0.8	13	1.79	72	23	8.3	2.9	24.2	27.6	13.6
94777	1.5	0.5	0.9	13	1.54	56	27	20.2	2.5	35.5	24.9	11.6
94391	1.5	0.6	1.1	9	1.25	23	29	33.1	7.0	25.3	27.2	13.4
94777	1.5	0.6	0.8	13	1.54	56	27	19.8	2.4	35.4	24.8	11.9
95206	1.5	0.5	0.8	13	1.21	6	46	42.0	8.5	26.4	25.3	11.4
1585853	1.5	0.4	1.0	13	1.24	11	40	36.7	7.4	30.1	25.8	11.8
94718	1.5	0.7	0.4	13	1.82	94	2	11.2	1.5	24.2	27.6	13.6
94987	1.5	0.6	0.7	13	1.67	59	24	14.3	1.8	32.9	25.2	11.7
94971	1.5	1.0	1.1	6	1.42	52	30	21.6	2.5	34.8	24.2	12.1
497713	1.5	0.5	0.8	13	1.24	22	28	41.0	10.0	26.4	25.3	11.4
94971	1.5	1.1	1.2	6	1.42	52	30	21.4	2.4	34.9	25.0	12.4
94573	1.5	0.5	0.6	13	1.53	56	23	19.3	2.4	33.4	24.8	11.3
95503	1.4	0.6	0.7	13	1.28	9	46	31.0	7.7	24.5	26.3	13.9
95408	1.4	0.5	1.1	13	1.24	10	42	40.0	6.7	30.1	24.8	12.9
94718	1.4	0.5	0.6	13	1.82	94	2	11.6	1.5	24.5	27.1	13.5

110441	1.4	0.7	1.0	13	1.33	26	31	28.3	4.5	29.0	24.4	11.4
94775	1.4	1.0	1.0	13	1.66	65	21	18.1	2.2	35.5	24.9	11.6
497725	1.4	0.6	0.9	13	1.33	31	32	29.6	7.2	26.4	25.3	11.4
94234	1.4	0.8	0.9	13	1.46	81	11	18.7	2.6	26.6	27.0	13.8
497724	1.4	0.5	0.8	13	1.33	31	32	29.6	7.2	26.4	25.3	11.4
95222	1.3	0.4	0.4	13	1.33	44	27	29.6	5.5	26.4	25.3	11.4
94842	1.3	1.0	1.0	14	1.33	61	18	25.4	3.9	32.1	28.1	13.7
94716	1.3	0.6	1.3	13	1.24	27	27	34.4	8.1	24.2	27.6	13.6
94904	1.3	0.7	0.8	14	1.5	26	54	21.8	3.0	34.5	28.0	14.2
1537056	1.3	0.4	0.8	13	1.25	24	29	36.1	7.4	28.1	24.9	11.9

C.6 Constraints for the landscape optimization.

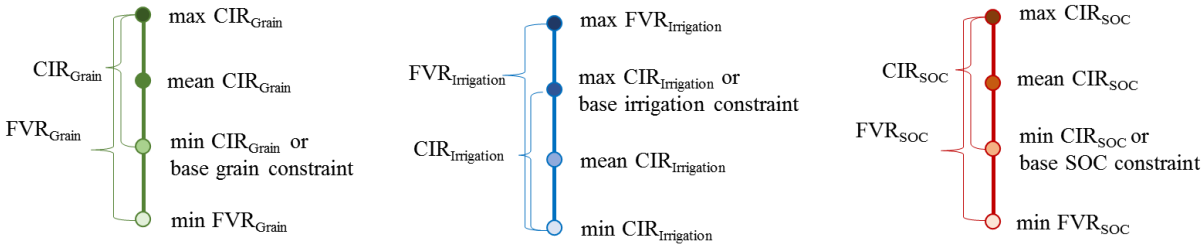


Fig. C.6.1. Illustration of Feasible Value Ranges (FVR) and Constraint Improvement Regions (CIR) for different landscape constraint metrics. FVR defines the range between min and max values of a constraint metric. CIR defines the range between the base and max values for corn grain and SOC, or between the base and min values for irrigation water use.

Table C.6.1. Constraint values used for different constraint scenarios

Constraint scenario	Corn grain (10^6 Mg dry wt. $year^{-1}$)	SOC (10^6 Mg C)	Irrigation (10^9 m^3)	Symbols on Fig. 5.7
Base	3.34	8.19	1.36	
Mean CIR_{grain}	3.39	8.19	1.36	
Max CIR_{grain}	3.45			
Mean CIR_{SOC}	3.34	9.17	1.36	
Max CIR_{SOC}		10.15		
Mean $CIR_{irrigation}$	3.34	8.19	1.32	
Max $CIR_{irrigation}$			1.29	

Note: The min, mean, and max CIR are scenarios where the subscripted constraint metric was improved to the min, mean, and max of its Constrain Improvement Region (CIR), respectively.

Table C.6.2. Average production constraints and optimization objectives of landscape trade-off Pareto frontiers

Pareto trade-off	Constraint scenario	Production Constraints			Optimization objectives				Percentage of total area (%)
		Grain ($\times 10^6$ Mg year^{-1})	Irrigation ($\times 10^9$ m^3 year^{-1})	SOC ($\times 10^6$ Mg C year^{-1})	Stover ($\times 10^6$ Mg biomass year^{-1})	GHG ($\times 10^4$ Mg CO_2e year^{-1})	NL ($\times 10^3$ Mg N year^{-1})	Profits ($\times \$10^6$ year^{-1})	
	Max grain	**3.45**	1.39	9.91	1.08	38.8	14.2	87	100
	Min irrigation	3.34	**1.29**	9.78	0.94	34.8	12.5	82	100
	Max SOC	3.44	1.36	**10.15**	0.89	35.8	13.8	83	97
Stover vs. N leaching	Base	3.35	1.33	8.66	1.86	48.9	13.0	98	99
	meanCIR$_{grain}$	**3.39**	1.35	8.94	1.74	47.5	13.4	97	100
	meanCIR$_{irrigation}$	3.34	**1.32**	8.82	1.69	46.2	12.7	95	98
	meanCIR$_{soc}$	3.38	1.33	**9.24**	1.47	43.1	12.9	92	99
GHG vs. N leaching	Base	3.34	1.30	9.86	0.87	33.8	12.4	81	97
	meanCIR$_{grain}$	**3.39**	1.32	9.99	0.89	34.7	12.9	83	99
	meanCIR$_{irrigation}$	3.34	**1.30**	9.86	0.87	33.8	12.4	81	97
	meanCIR$_{soc}$	3.34	1.30	**9.86**	0.87	33.8	12.4	81	97
Stover vs. GHG	Base	3.34	1.32	9.05	1.52	43.4	12.7	92	98
	meanCIR$_{grain}$	**3.39**	1.34	9.26	1.49	43.5	13.3	93	99
	meanCIR$_{irrigation}$	3.34	**1.31**	9.17	1.41	41.7	12.6	90	98
	meanCIR$_{soc}$	3.35	1.31	**9.45**	1.23	39.1	12.6	88	98

Note: The average of a Pareto frontier was computed from all solutions in that Pareto frontier. The Pareto trade-offs and constraint scenarios are corresponded to those in Fig. 5.7. Each column is colored based on its value range with white color being the min value. The min, mean, and max CIR are scenarios where the subscripted constraint metric was improved to the min, mean, and max of its Constrain Improvement Region (CIR), respectively. The bolded numbers are values of the varied constraint metrics.

C.7 Examples of county-level optimum stover removal rates.

Table C.7.1. County-level optimum stover removal rates at different stover production targets for irrigated corn systems in Colorado

Stover production ($\times 10^6$ Mg biomass year^{-1})	0.83	0.96	1.60	2.35
County name	**Stover removal rate (%)**			
Yuma	37	42	63	82
Weld	36	42	59	74
Kit Carson	15	16	43	72
Logan	24	26	66	77
Morgan	37	47	59	78
Phillips	26	27	53	82
Baca	30	32	39	70
Sedgwick	21	23	57	77
Prowers	20	22	38	69
Washington	27	31	60	77
Cheyenne	21	21	33	71
Otero	20	34	53	79
Bent	20	22	29	57
Pueblo	12	13	26	73
Larimer	22	32	44	72
Lincoln	13	14	47	62
Adams	30	41	48	60
Elbert	21	29	45	81
Crowley	34	47	56	80
Kiowa	28	28	30	53
Arapahoe	17	20	25	49
El Paso	33	34	68	74
Las Animas	24	24	34	85
Boulder	11	67	67	89
Fremont	22	22	38	50
Huerfano	23	28	45	67

Note: The table is colored from blue to yellow to red based on its value range with blue color being the min value.