CHANGING AGRICULTURAL SUITABILITY: MODEL DEVELOPMENT AND APPLICATIONS IN THE PAST AND FUTURE

by

Clarice Lee Bidez

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

The University of Utah

May 2018

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Clarice Lee Bidez**

has been approved by the following supervisory committee members:

Simon Brewer	, Chair	**5 January 2018** Date Approved
Duncan Metcalfe	, Member	**5 January 2018** Date Approved
Mitchell Power	, Member	**5 January 2018** Date Approved

and by **Andrea Brunelle** , Chair/Dean of

the Department/College/School of **Geography**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Modeling the relationships between climate and crop growth is an important tool to study agriculture in the past and future. Agricultural suitability models, which estimate the probability of rain-fed cultivation in a particular location, provide an opportunity to predict agricultural conditions for multiple scenarios in space and time. This thesis presents a methodological approach to developing a simple agricultural suitability model with machine learning techniques and applying it to two scenarios. First, to predict the potential impacts of climate change, the model was used to project agricultural suitability at the end of the 21st century at the global scale. Second, the model was used to predict agricultural suitability in Range Creek Canyon, an important archeological site in east-central Utah, in the 10th -13th centuries. The agricultural suitability model was supplemented with a streamflow model of Range Creek to address the potential for irrigation. The predictive power of random forest and XGBoost were compared and random forest was found to build the stronger model in this application. Applied to the future, this model predicts a net increase of 13.8-28.2% of agriculturally suitable land by AD 2100. The broad patterns agree with past suitability models, with the largest gains of suitability distributed in the high-northern latitudes. Decreases in suitability are projected in some regions, particularly in the current most intensely cultivated regions including Midwestern America. A sensitivity analysis revealed that the main driver of these shifts was changing growing season length and intensity, while changing soil moisture had a

limited effect. Applied to the past in Range Creek Canyon, the model predicted very low agricultural suitability, providing evidence that rain-fed agriculture in Range Creek Canyon from AD 900-1200 was nearly impossible. Modeled Range Creek streamflow predicted that mean streamflow from AD 900-1200 was 6.34% greater than modern streamflow. The area of maize fields that could be irrigated with this amount of streamflow was approximately 30.6 hectares. The agricultural suitability model applied to two scenarios differing in scales of time and space demonstrate how it can provide meaningful insights to a broad range of past and future scenarios.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES vi

LIST OF FIGURES vii

ACKNOWLEDGMENTS viii

Chapters

1. INTRODUCTION 1

- Climate Impacts on Agriculture 3
- Overview of Crop Models 5
- Impacts of Climate Change on Future Agriculture and Food Security 7
- Impacts of Climate on Prehistoric Agriculture in Range Creek Canyon 9

2. AGRICULTURAL SUITABILITY MODEL 13

- Methodology 13
- Future Agricultural Suitability and Sensitivity 21

3. AGRICULTURAL SUITABLITY AND STREAMFLOW IN RANGE CREEK CANYON, UTAH 34

- Introduction 34
- Data 35
- Methods 37
- Results 40
- Discussion 42

4. CONCLUSION 52

REFERENCES 54

LIST OF TABLES

Tables

1. Sources and descriptions of data used to build the model 27

2. Sources and descriptions of data used to project future agricultural suitability 30

3. The area of agriculturally suitable (AgS) land and changes compared to predicted modern agricultural suitability 31

4. Sources and descriptions of data used to project past agricultural suitability and streamflow 46

5. Differences between modern modeled means and the means of past scenarios 49

LIST OF FIGURES

Figures

1. Flow chart of the model-building process. 28

2. Comparison of agricultural suitability projections. 29

3. Distribution and residuals of future agricultural suitability. 32

4. Sensitivity analysis of future projections. 33

5. Linear regression of streamflow 45

6. Time series of Loess-smoothed monthly α values in Range Creek Canyon from AD 900-1200. 47

7. Streamflow calibration results. 48

8. Boxplots of Range Creek Canyon agricultural suitability, α, and GDD. 50

9. Boxplots of Range Creek Canyon streamflow and irrigation potential. 51

ACKNOWLEDGMENTS

Thank you to my advisor Simon Brewer. It has been an honor to work with and learn from you the last 3 years. I always left our meetings inspired and excited to try new ideas. Thank you to Duncan Metcalfe for recruiting me to participate in the incredible research surrounding Range Creek Canyon. Thank you to Mitchell Power for contributing your grand perspective to this project. Thank you to the amazing staff of the Department of Geography including Pam Mitchell and Lisa Clayton. Thank you to Christy Clay for starting me on this path, and to Yoshi Maezumi for pointing me in the right direction at a critical time. A special thank you to Andy and to my family, Patty, Earle, and Dylan. Additionally, thank you to the Natural History Museum of Utah for funding my Graduate Assistantships.

CHAPTER 1

INTRODUCTION

Models describing the relationship between crop growth and climate have been widely used to better understand critical components of food systems and to predict potential outcomes to future conditions. Amid threats from anthropogenic climate change, crop models are increasingly important tools to equip scientists, policy-makers, and farmers with vital information for an uncertain future. Additionally, crop models can provide valuable insight to researchers studying past agriculture of prehistoric civilizations.

Most crop models produce estimates of crop yield as an output, but an interesting and much less studied approach is to estimate agricultural suitability. Agricultural suitability predicts the probability that a particular location will be cultivated and provides a spatial domain for the crop/climate relationship. There is now an opportunity to improve past iterations of agricultural suitability models with advanced statistical modeling techniques and updated data. Additionally, agricultural suitability models can be applied to research questions at a broad range of scales in time and space and can easily be used in tandem with additional models to explore specific research questions related to climate/crop relationships.

The first objective of this thesis is to develop a global agricultural suitability

model, building on the framework outlined by Ramankutty, Foley, Norman, and McSweeney (2002). Two machine learning methods were used to build the model, and the resulting models were evaluated and compared by predictive skill. The second objective of this thesis is to apply the more robust model at the global scale in a snap-shot of time under two future climate scenarios. Future climate change has the potential to have net negative impacts on crop-yield, which would subsequently impact food security (Porter et al., 2014). Projecting future shifts in agricultural suitability can inform how to adapt to climate change and mitigate potential losses in food security. The third research objective is to apply the agricultural suitability model to prehistoric farming in the past over a 300-year-long time series and compliment the model with a simple streamflow model. Projecting changing agricultural suitability in Range Creek Canyon, UT during the Fremont occupation from AD 900-1200 provides the opportunity to understand agricultural limitations of these prehistoric Native American farmers. Irrigation likely played an important role in maize cultivation for the Fremont in Range Creek Canyon (Boomgarden, 2015) and therefore, understanding streamflow is a crucial component to understand past conditions.

This thesis presents a methodological approach to modeling climate/crop relationships at broad scales in space and time. The remainder of the first chapter presents a literature review of key topics: climate impacts on agriculture, an overview of crop models, future impacts of climate change on agriculture and food security, and relevant background of Range Creek Canyon and the Fremont farmers that occupied the canyon. Chapter 2 focuses exclusively on the agricultural suitability model and will present its development and applications in the future. Finally, Chapter 3 covers the implementation

and results of the agricultural suitability model and supplemental streamflow model in Range Creek Canyon from AD 900-1200.

Climate Impacts on Agriculture

Agricultural outcomes are greatly determined by climate, and precipitation and temperature are the two most important climate parameters that directly affect crop growth. All crops require access to water and a minimum number of days within an optimum temperature range, although the range of climate requirements varies depending on crop species and cultivar. Water availability is thought to be the most important variable for crop growth as drought is a leading cause of crop failure in rain-fed systems (i.e., Baro & Deubel, 2006; Mall, Singh, Gupta, Srinivasan, & Rathore, 2006), but temperature is also a vital control through qualities discussed below.

Water availability to plants is determined by precipitation and temperature through evapotranspiration and soil qualities that determine water retention. Crop modelers have found that an adequate proxy of soil moisture available to crops is the ratio of actual evapotranspiration to equilibrium evapotranspiration, referred to as the Priestley-Taylor coefficient (Cramer & Solomon, 1993; Priestly & Taylor, 1972), or the ratio of actual evapotranspiration to potential evapotranspiration, α (Ramankutty et al., 2002), which are determined by both precipitation and temperature. Equilibrium evapotranspiration considers only local conditions whereas potential evapotranspiration accounts for advection of dry air above a surface. Growing season intensity and length is more important to agriculture than temperature values, and is often measured for a location with growing degree days (GDD) by summing the annual total of degrees for

days above a base temperature, commonly 5°C. GDD and α have been shown to be adequate representations of the cold and dry limits of global agricultural lands (Cramer & Solomon, 1993).

Climate has the potential to indirectly impact crop growth as well, such as through pests and disease (Porter et al., 2014; Tubiello, Soussana, & Howden, 2007). Just as climate greatly determines the distribution and success of vegetation (Prentice, Bartlein, & Webb, 1991), climate also controls the success and geographic distribution of weeds that compete for resources, destructive insects, and pathogens (Porter et al., 2014). For example, the geographic distribution of *Myzus persicae*, an aphid that is considered a major pest of sugar beets and potatoes, was found to negatively correlate with precipitation and positively correlate with temperature (Cocu, Harrington, Rounsevell, Worner, & Hullé, 2005).

Crop growth is also influenced by the climate-related variables of atmospheric concentration of carbon dioxide (CO_2) and ozone (O_3). Increased CO_2 has been shown to increase photosynthesis, and therefore yield, in C3 plants, which include soybean and wheat, by reducing the instance of photorespiration (Ainsworth & Rogers, 2007; Long, 1991). This effect is not found in C4 plants such as maize and sorghum, as the unique photosynthetic pathway of these plants eliminates photorespiration (von Caemmerer & Furbank, 2003). A second way increased CO_2 impacts crop growth is by reducing stomatal conductance, the rate that CO_2 enters or water vapor exits a leaf through its stomata (Ainsworth & Rogers, 2007). Both C3 and C4 plants benefit from reduced stomatal conductance by allowing plants to conserve water while maintaining sufficient CO_2 for photosynthesis (Ainsworth & Rogers, 2007). It is important to note, however,

that increased CO_2 causes a significant decrease in plant protein in all crops investigated so far, including barley, rice, wheat, soybean, and potato (Taub, Miller, & Allen, 2008). Another important consideration is the detrimental impact of atmospheric O_3, a phytotoxic pollutant that inhibits photosynthesis (Ainsworth & Rogers, 2007; Morgan, Ainsworth, & Long, 2003). Important crops sensitive to O_3 include soybean, wheat, peanut, and cotton (J. L. Hatfield et al., 2011), and a meta-analysis of the effects on enclosed and controlled-environment soybean showed decreased yield after exposure to O_3 (Morgan et al., 2003).

Overview of Crop Models

There are two important approaches to crop yield models: process-based crop growth models and statistical models (Rosenzweig et al., 2013). Process-based growth models (also called biogeochemical or mechanical models) attempt to incorporate as much relevant detail about the biology of plant growth as possible in order to accurately predict crop yield (Brisson et al., 2003; Challinor, Wheeler, Craufurd, Slingo, & Grimes, 2004; Del Grosso, Mosier, Parton, & Ojima, 2005; Keating et al., 2003). These models are crop-specific and require detailed climate data, soil type and nutrients data, and crop management information, but have the advantage of being more flexible and can account for impacts of dynamic CO_2. Alternatively, statistical models construct crop yield regression equations using meteorological inputs (i.e., temperature and precipitation) trained on historic data, and have been shown to be comparable to process-based crop models at broad spatial scales (Lobell & Burke, 2010; Makowski et al., 2015). The advantage of using statistical models is their ability to predict with few key variables,

measure uncertainty, and to analyze large and irregular datasets across the globe (Lobell & Burke, 2010; Rosenzweig et al., 2013). Additionally, they are able to identify temperature and CO_2 thresholds that result in significant yield gains and losses through interpolating between data points (Makowski et al., 2015). However, important limitations of statistical models are a lack of consideration of variables related to plant biology as well as potential bias due to spatial correlation of climate variables (Lobell, Schlenker, & Costa-Roberts, 2011; Schlenker & Roberts, 2009).

While crop yield models are valuable, particularly because crop yields can be used as inputs for economic models, they lack the ability to explore shifting distributions of agriculturally suitable land, which likely impacts global crop yields. Agricultural suitability models (also called land suitability or agricultural land availability models) use statistical methods to explore the relationship between climate to cropland distribution (i.e., Ramankutty et al., 2002; Zhang & Cai, 2011). Agricultural suitability can be defined and measured as the proportion of land in a grid cell that is able to successfully grow crops. While traditional crop models ask the question "How much crop yield is expected under an assigned set of conditions for a specific crop," agricultural suitability models ask the question "How likely is it that land will be successfully cultivated for any crop under an assigned set of climate conditions?"

Ramankutty et al. (2002) created a pioneering global agricultural suitability model to examine sensitivity to climate change. To build the model, the authors empirically derived probability density functions to describe the relationships between the probability of cultivation and the climate variables GDD and α, and the soil quality parameters carbon density and soil pH. The resulting model captures important global patterns of

current cropland distribution but overestimates agricultural suitability. The authors argue that this result is realistic and represents the underutilization of suitable cropland due to competing factors such as grazing and forestry. Broadly, the study found that agricultural suitability is expected to decrease in tropical regions and increase in high northern latitudes under future climate change scenarios (Ramankutty et al., 2002).

There are trade-offs to using a model as simple as the above agricultural suitability model. Limitations include the exclusion of potentially important variables and processes such as disturbance and certain adaptive farming practices such as irrigation. Benefits of a simple model opposed to a complex model include fewer data requirements (important for regions where data are scarce), decreased computing time, and the ability to calibrate models on observed data. Additionally, simple agricultural suitability models implicitly account for some adaptive farming practices such as cultivar/species selection and changing cultivation and sowing times.

Impacts of Climate Change on Future Agriculture and Food Security

Understanding the effects of anthropogenic climate change on agriculture is imperative to ensure future food security of a growing population. Food security is met by physical and economic access to food, and involves all elements of complex food systems including production (i.e., crop growth) and nonproduction (i.e., packaging, transport, storage, retail, and income) elements (Porter et al., 2014). In their fifth report, the Intergovernmental Panel on Climate Change (IPCC) concluded with high confidence that climate change has the potential to affect all production and nonproduction aspects of food security (Porter et al., 2014), including through changes in crop yields and

consequences of climate-related disasters (J. Hatfield et al., 2014). If crop yields decrease, it is expected that food prices will increase in response, potentially resulting in decreased physical and economic food access (J. Hatfield et al., 2014; Porter et al., 2014).

A major concern associated with a potential decrease in future food security as a result of climate change is how to minimize these effects to feed a rapidly growing population. The people most vulnerable to food insecurity are the poorest populations of countries affected by conflict, violence, and fragility (FAO, IFAD, UNICEF, WFP, & WHO, 2017). The World Bank estimates that these populations may increase from 17% today to 50% of the total global population by 2030 due to the population growth (The World Bank, 2017). In 2016, an estimated 815 million people were chronically undernourished compared to 777 million people the previous year, the first increase in more than a decade (FAO et al., 2017). The causes of this decrease in food security are complex, and include increased conflict in Africa and the Near East and economic challenges in parts of Latin America (FAO et al., 2017). Many of these regions suffered concurrent incidence of drought and flood, exacerbating the problem (FAO et al., 2017).

Adapting food systems to mitigate potential losses will be essential moving forward, and an important consideration will be the changing geographic distribution of land suited to successfully grow crops. As part of the fifth Coupled Model Intercomparison Project (CMIP5), researchers have used a suite of general circulation models (GCMs) to simulate future climate scenarios (Taylor, Stouffer, & Meehl, 2012). Using the CMIP5 output data, the IPCC concluded that the mean global temperature for 2081-2100 is likely to increase between 0.3°C and 4.8°C relative to 1986-2005 depending on the CO_2 emissions scenario (Collins et al., 2013). Precipitation is expected

to increase globally as temperatures rise, although models show much greater uncertainty when projecting precipitation compared to temperature (Knutti & Sedláček, 2012). The spatial pattern of precipitation increase is expected to resemble a "rich get rich, poor get poorer" scenario in which already wet regions in high latitudes will experience most of this increase while normally arid regions dry out even further (Collins et al., 2013). These changes have the potential to directly affect agriculture in two important ways: growing season length and water availability. In some tropical regions, crop yields are predicted to decrease as extreme heat shortens the growing season and causes water loss via increased evaporation (Jones & Thornton, 2003; Porter et al., 2014). However, croplands are expected to expand at high northern latitudes as increasing temperatures cause the growing season to lengthen (Ramankutty et al., 2002; Zhang & Cai, 2011). Overall, climate change is expected to have a net negative impact on global crop yield through direct (i.e., temperature) and indirect (i.e., pests and diseases) impacts (Porter et al., 2014).

Impacts of Climate on Prehistoric Agriculture in Range Creek Canyon

Range Creek Canyon in east-central Utah on the Tavaputs Plateau is the site of an intense Fremont occupation from AD 900-1200. The Fremont were an agricultural society as evidenced by maize cobs, maize starch on tools, and maize storage structures (Boomgarden, 2015). The cause of the civilization's collapse in AD 1200 is not well understood, but a reasonable hypothesis is that changing climate reduced agricultural suitability and thus food security, potentially increasing already present internal strife (Boomgarden, Metcalfe, & Springer, 2014). Similar collapses of Native American

societies in the 11th through 13th centuries across the American Southwest have been correlated with severe drought, including the Anasazi in the Four Corners region and other Fremont populations in the Uinta Basin, eastern Great Basin, and Southern Colorado Plateau in AD 1150 (Benson et al., 2007). Modeled prehistoric maize productivity for southwestern Colorado showed that during a major drought in AD 1130-1180, precipitation would only have been adequate for dry maize farming at elevations above 2200 m, where growing seasons are considered to be unsuitable for prehistoric agriculture (Benson, Ramsey, Stahle, & Petersen, 2013), a challenge which may have contributed to the decline of the local Anasazi population.

Experimental maize farms in the Range Creek Canyon in 2013 and 2014 revealed that modern maize farming is nearly impossible without irrigation (Boomgarden, 2015). A precipitation reconstruction from the Tavaputs Plateau indicates that the climate during the Fremont occupation was similar to the modern climate and experienced significant dry periods, though on average received greater mean annual precipitation than modern years (Boomgarden, 2015; Knight, Meko, & Baisan, 2010). Using the Tavaputs Plateau precipitation reconstruction, Boomgarden (2015) calculated that dry farming maize would have been possible for at most 38 years out of the 300-year occupation, assuming seasonal precipitation patterns were not drastically different during that time frame. This means that irrigation would have played a key role in the civilization's ability to farm, and the streamflow of Range Creek is important to understand.

In an effort to estimate the area of maize fields possible to irrigate with available streamflow (hereafter called "irrigation potential"), Potter (2016) developed a method that considers irrigation efficiency and water requirements for maize during the critical

maize reproductive phase. Simple surface irrigation systems such as diversion dams result in considerable water loss due to spillage, percolation, and evaporation. Therefore, the amount of streamflow successfully converted to soil moisture is a small proportion of the total streamflow even when 100% of streamflow is diverted. Potter (2016) estimates that this proportion, the irrigation efficiency, can be no greater than 38% in Range Creek Canyon with simple technology. Water requirements for maize peak at 0.81 cm per day during the reproductive phase (Kranz, Irmak, Van Donk, Yonts, & Martin, 2008), and thus water availability during reproduction is arguably the greatest limiting factor in irrigation potential. This method for calculating irrigation potential does not consider complicating factors such as the location(s) of irrigation diversion and possible subsequent downstream recharge via groundwater contribution, but is a starting-point to quantify how streamflow translates to maize production along Range Creek.

The Range Creek watershed is located on the western side of the Tavaputs Plateau and has an area of ~375 km^2. It is a perennial stream that flows for 60 km and discharges into the Green River. Precipitation contributes to streamflow via surface runoff and snowmelt, and it is thought that a large proportion of Range Creek streamflow is groundwater contribution, though the groundwater/streamflow relationship is not yet well understood (Potter, 2016). Streamflow measurements recorded in the summers of 2015 and 2016 ranged from 1037 m^3 to 9245 m^3 per day, and included a measurement of 3,542 m^3 per day during the maize reproductive phase (Potter, 2016), enough water to irrigate 16.6 hectares of maize per day. This irrigation potential seems low considering the intensity of the Fremont habitation; however, a single measurement has limited accuracy to predict irrigation potential presently or during the Fremont occupation. While future

studies of Range Creek Canyon will yield more streamflow data and a better understanding of groundwater and maize farming, it is currently possible to model past conditions to provide preliminary insight into the agricultural potential and challenges for the Fremont farmers in Range Creek Canyon.

CHAPTER 2

AGRICULTURAL SUITABILITY MODEL

The objective of this chapter was to develop a simple global agricultural suitability model and apply it to future scenarios. The first section describes the construction and evaluation of two candidate models. In the second section, the more robust model was used to project agricultural suitability at the end of the century and calculate predicted gains and losses in agricultural suitability. A sensitivity analysis was also conducted to parse the individual contributions of changing GDD and α to changing agricultural suitability.

Methodology

The framework for this agricultural suitability model was derived from the model presented by Ramankutty et al. (2002). Ramankutty et al. (2002) used GDD, α, soil pH, and soil carbon content as predictors of fractional cropland area and created an agricultural suitability index by combining probability density functions of each variable. The resulting projection of modern agricultural suitability depicted important geographic patterns of intensely cultivated regions in the United States, Russia, and China, but greatly overestimated agricultural suitability compared to the actual distribution of cropland. The authors justify the overestimation as a realistic representation because not

all agriculturally suitable land is utilized as cropland, and instead is used for grazing, forestry, or is protected (Ramankutty et al., 2002).

Like Ramankutty et al. (2002), this model used GDD, α, and soil quality as drivers of agricultural suitability. Additionally, slope was included as a variable to account for topography, an influence that Ramankutty et al. (2002) identified as necessary to incorporate in future model iterations. The distribution of cropland used to train this model includes rain-fed cropland only, which represents more than 70% of modern agriculture (Porter et al., 2014), and allows the effects of climate to be distinguished from the effects of irrigation.

The machine learning statistical approaches used in the construction of this model differ substantially from previous agricultural suitability models. While the Ramankutty et al. (2002) model is empirically derived, the rules and weights of the relationship between variables and cropland distribution is oversimplified and ultimately programmed by the authors. This strategy potentially impacts model fit and predictive power, and it fails to account for nonlinear relationships between climate variables and agricultural suitability. Supervised machine learning algorithms are able to address these issues by iteratively building, analyzing, and revising models based on the given data to produce final dynamic models capable of making highly accurate predictions. They are completely data-driven and are able to model nonlinearities and other complex relationships found within the data without being explicitly programmed. The two machine learning techniques used in this study are random forest and XGBoost, described in detail below.

Data

This study utilized a global dataset from the Global Agro-Ecological Zones Assessment for Agriculture (GAEZ) of the distribution of modern rain-fed cultivated land, which is derived from six geographic datasets with the purpose of categorizing global land cover (IIASA & FAO, 2012). The units are the proportion of cropland per raster cell provided at a 5-minute resolution, which were aggregated to a 10-minute resolution. These data were used to train the random forest and XGBoost models.

WorldClim 2 (Fick & Hijmans, 2017) is a spatially interpolated monthly terrestrial data set that provides climate variables averaged across 1970-2000. This study used 10-minute mean monthly temperature and mean monthly sum of precipitation from this data set to represent modern climatology. Mean monthly cloud cover fraction was used from the Climatic Research Unit Climatology version 2.0 (CRU CL v. 2.0) (New, Lister, Hulme, & Makin, 2002), a mean monthly global climatology data set averaged across 1961-1990. Precipitation, temperature, and cloud cover fraction were inputs used to calculate α, the ratio of potential evapotranspiration over actual evapotranspiration, using SPLASH methodology (Davis et al., 2016). Monthly mean temperature was interpolated into daily mean temperature to calculate growing degree days (GDD), the annual sum of degrees for every day above surface air temperature 5°C, a proxy of growing season length and intensity.

As part of the Harmonized World Soil Database (HWSD), a data set of seven soil quality parameters important to crop growth was created (FAO, IIASA, ISRIC, ISSCAS & JRC, 2012). The parameters are nutrient availability, nutrient retention capacity, rooting conditions, oxygen availability to roots, excess salts, toxicity, and workability.

These data are provided at a 5-minute resolution on a 4-tier categorical scale ranging from “very severe constraints” to “no or slight constraint,” with the additional categories “mostly nonsoil,” “permafrost area,” and “water bodies.” These categories were converted into continuous numerical data by assigning the three nonsoil categories the value 0, and the remaining categories integer values 1-4, with 4 as the highest quality. These data were then aggregated to a 10-minute scale by taking the mean of 2 x 2 adjacent grid cells, a process that transformed the integers to continuous numerical values from 0 to 4. Because the soil quality parameters were highly correlated and to make the model as parsimonious as possible, a principal component analysis (PCA) was run on the seven variables. The first principal component, which accounted for 91% of the variability, was used as a soil quality index to build the agricultural suitability model.

To incorporate topographical features that may constrain successful crop cultivation, slope (degrees) was also included in the model. Slope was calculated with the landsat8 package in R (dos Santos, 2017) using an elevation raster from the HWSD (FAO et al., 2012). The data used to build and assess the models are summarized in Table 1, and a flow chart describing the data and model-building process is presented in Figure 1.

Methods

All analyses were conducted with R (R Core Team, 2017). Random forests (Breiman, 2001) and XGBoost (Chen & Guestrin, 2016) are both tree-based ensemble methods. Tree-based methods refer to classification and regression tree algorithms that build a network of nodes and splits intended to maximize variance of points between nodes and minimize the variance within nodes. For regression trees, the type of trees used

in this study, the measure of variance is the sum of squared errors (S). Each tree begins with a "root" node that contains all of the points in the dataset. The data in that node is then split into two parent nodes in a position that will produce the maximum reduction of S using one of the explanatory variables (i.e., all points < GDD 3000 belong to parent node 1, all points ≥ GDD 3000 belong to parent node 2). Each parent node is split in the same way until a stopping criterion is met. The mean of the observed outcomes (response variable) of the points in each terminal node is the value subsequently assigned to predicted outcomes. The resulting model is a decision tree that predicts an outcome based on given explanatory variables. Ensemble tree methods build hundreds or thousands of these trees and take votes or an average of all trees to produce the final model. They are capable of handling large, irregular datasets with missing data (Breiman, 2001; Chen & Guestrin, 2016), all of which recommends them as great options for this application.

Random forest is an algorithm that utilizes a "bagging" method. Individual trees are grown with bootstrapped samples of data with the same distribution, and each node is split with a random selection of variables. In this way, trees are independent from each other, allowing many trees to be built (100s to 1000s) without overfitting the model to the data (Breiman, 2001), and also allowing variable importance to be calculated. Random forest averages the ensemble of trees to create a predictor (Breiman, 2001). This study used random forest regression available through the randomForest package in R (Liaw & Wiener, 2002). The model was built with 200 trees that considered two variables at each node.

XGBoost (eXtreme Gradient Boosting) utilizes a "boosting" method. Trees are built sequentially, and each tree models the errors of the previous combination of trees.

Unlike random forest, the trees are dependent on each other and the final model is a combination of the weighted trees. XGBoost differs from traditional gradient boosting in a few major ways. The algorithm builds full trees down to individual data points then prunes them from the bottom-up, and employs a regularization parameter that penalizes model complexity to avoid overfitting (Chen & Guestrin, 2016). Lastly, XGBoost is able to run in parallel, dramatically decreasing computation time (Chen & Guestrin, 2016). This study used the logistic regression option in the XGBoost package in R (Chen, He, Benesty, Khotilovich, & Tang, 2017) and tuned parameters using the mlrMBO package (Bischl et al., 2017).

Random forest and XGBoost agricultural suitability models were built as a function of GDD, α, soil quality, and slope, and trained with modern rain-fed cropland distribution (Figure 1). The models were used to predict modern agricultural suitability and were evaluated with the root mean squared error (RMSE). They were further examined by mapping the residuals of the proportion of rain-fed cropland subtracted from the projected agricultural suitability.

Results

The random forest proved to be a more robust model than XGBoost for this application. The RMSE of the random forest's projection of modern agricultural suitability was 0.044 and the model reported to explain 68.38% of the variance. The RMSE of XGBoost's modern projection was 0.099 and did not capture extreme values as well as the random forest model (Figure 2). The models underestimate and overestimate in the same geographic pattern, but the magnitude of the residuals is more extreme with

the XGBoost model (Figure 2). The random forest model was used for further projections and analysis in the remainder of this thesis.

The models tend to underestimate agricultural suitability in the most intensely cultivated regions such as the U.S. Midwest, India, Southern Australia, sub-Saharan Africa, and Eurasia. They overestimate suitability in regions with very little rain-fed cultivation such as the American Southwest and Northern Australia but predict the lack of rain-fed cultivation in the Sahara, the Amazon, Western China, and high-latitudes of Eurasia, North America, and Greenland very well. These regions are marked with poor soil quality or permafrost, which may be a prominent influence in the predictions of these regions. Although magnitude of agricultural suitability in regions with intense cultivation and regions with existing but little cultivation are not perfectly projected, the geographic pattern of suitability is very accurate.

Discussion

This agricultural suitability model predicts the distribution of rain-fed cropland accurately, especially considering its simplicity. This model is intended to be as simple as possible while still providing valuable insight to the geography of climate/agriculture relationships. The benefits of simple models include ease of implementation and interpretation; however, they require sacrifices in precision and must be discussed within the context of their limitations.

Variables omitted from this model that are known to impact agriculture are changing atmospheric CO_2 and O_3 and changing distributions of pests and insects. Although results of CO_2 fertilization studies are inconsistent (Porter et al., 2014),

increased CO_2 is known to positively impact yield (McGrath & Lobell, 2011) and negatively impact nutritional value (J. L. Hatfield et al., 2011). While these effects are not relevant to agricultural suitability, increased CO_2 may positively impact suitability in regions with low α values due to decreased stomatal conductance. On the other hand, any positive effects of CO_2 fertilization may be counteracted by the damaging effects of increasing atmospheric O_3 (Porter et al., 2014). Climate can also impact the geographic distribution of pests and diseases, but there is a shortage of studies on the topic (Porter et al., 2014). Because the individual and interactive effects of these factors on agriculture are not well understood, they were left out of this study.

Another limitation of this model that it does not address extreme climatic events, such as drought and floods, that are projected to occur more often and intensely by the end of the century (Seager et al., 2007; Sheffield & Wood, 2008). Drought and extreme heat were demonstrated to have a considerable damaging effect on global agriculture from 1964-2007 (Lesk, Rowhani, & Ramankutty, 2016), so it can be assumed that they will in the future as well. Extreme climatic events have been identified as particularly difficult to model because their rarity makes model calibration nearly impossible (Porter et al., 2014).

In a very general way, this model takes human management and adaptation into consideration. The distribution of cropland used to calibrate the model includes land managed with fertilizer, cultivar selection, and other adaptive practices that are therefore implicit to the model. This means that if a farm is projected to experience a locally novel climate in the future, if that climate has a modern analog somewhere on the globe where land is cultivated, the farm will be deemed suitable for agriculture. This is a strength of

the model where adaptive management practices are available and possible, but it is a limitation where they are not. To take advantage of the model in this way, further efforts to identify modern analogs and research their management practices would be required.

Future Agricultural Suitability and Sensitivity

Climate change has the potential to negatively impact global food security in the coming century (Porter et al., 2014). Decreased crop yields due to increased temperature and aridity are expected, and adaptation to these changes will be a vital response to maintain food security. Agricultural suitability models can identify at-risk regions as well as regions that may become more agriculturally suitable due to climate change and are therefore a useful tool for managers and policy-makers.

This section utilizes the agricultural suitability model described in the previous section to project agricultural suitability at the end of the 21st century. Future agricultural suitability is compared to modern suitability to identify regions of gains and losses. Additionally, a sensitivity analysis is conducted to parse the individual contributions of changing GDD and changing α on the outcome.

Data

Future mean monthly temperature and mean monthly total precipitation data simulated by the Goddard Institute for Space Studies (GISS) ModelE2 GCM, the model used in CMIP5 (Nazarenko et al., 2015; Taylor et al., 2012), were used to predict agricultural suitability at the end of the 21st century. To estimate a range of potential outcomes, low (RCP2.6) and high (RCP8.5) CO_2 emissions scenarios were used. The

RCP2.6 scenario represents a "peak and decay" emissions pattern in which radiative forcing increases until the middle of the 21st century, then falls to a level of 2.6 W m^{-2} by 2100 (Taylor et al., 2012). The RCP8.5 scenario represents a high emissions scenario in which radiative forcing increases throughout the 21st century, reaching 8.5 W m^{-2} by 2100 (Taylor et al., 2012).

As with the modern climate data, mean monthly temperatures were interpolated and used to calculate GDD. Mean monthly precipitation, temperature, and cloud cover fraction from CRU CL v. 2.0 (New et al., 2002) were used to calculate α with SPLASH (Davis et al., 2016). Modern cloud cover was used because it is one of the only estimates of high-resolution cloud cover available, although this is not an ideal solution as recent evidence indicates that cloud distribution and shape is changing as a result of climate change (Witze, 2016). Soil quality and slope were assumed to be unchanging on this relatively small time-step, so the same data sets from the HWSD (FAO et al., 2012) were used to predict future suitability. Projections of agricultural suitability were debiased using the change-factor approach (Wilby et al., 2004) by subtracting future modeled agricultural suitability from modern modeled agricultural suitability, then adding the resulting residuals to the modern observed rain-fed cropland (IIASA & FAO, 2012). The data sources and descriptions used in this section are summarized in Table 2

Methods

Experiment 1: Future Agricultural Suitability

The agricultural suitability model built with a random forest described in the first section was used to project future suitability under RCP2.6 and RCP8.5 scenarios. Future

GDD, future α, modern soil quality, and slope were the inputs into the model. Residuals for each projection were calculated by subtracting the projected modern agricultural suitability from future agricultural suitability, and the final future agricultural suitability distributions were calculated by adding the residuals to the modern observed distribution of rain-fed cropland.

Experiment 2: Sensitivity Analysis

A sensitivity analysis was conducted by running the model with a) future GDD and modern α, and b) modern GDD and future α. The future scenario used for the sensitivity analysis was RCP8.5. The same process as above was used to calculate residuals and final results for the sensitivity analysis. Predicted global change in agricultural suitability was calculated for each future and sensitivity scenario.

Results

Experiment 1: Future Agricultural Suitability

The model predicted a 13.8% and 28.2% increase in the global area of agriculturally suitable land for RCP2.6 and RCP8.5, respectively (Table 3). Figure 3 shows the distribution of projected future agricultural suitability as well as the residuals between the future scenarios and modern projected agricultural suitability, which can be interpreted as gains (positive values) and losses (negative values). The land surface area projected to improve suitability by at least 1% is 38.8% for RCP2.6 and 45.6% for RCP8.5. Although the model predicts a net gain in suitability, a considerable proportion of land will experience a suitability loss of at least 1% compared to projected modern

values: 18.8% for RCP2.6 and 18.6% for RCP8.5.

Experiment 1: Future Agricultural Suitability

The sensitivity analysis shows that future change in GDD has the most influential impact on the positive outcomes compared to future change in α. A simulation run with RCP8.5 GDD values and modern α values resulted in a projection with a 27.8% increase in suitability compared to modern projected suitability (Table 3). Alternately, simulation with RCP8.5 α and modern GDD resulted in a -0.6% change, indicating a minor net loss in suitability. Figure 4 shows the distribution of the residuals between the two sensitivity projections and projected modern agricultural suitability.

The area of land projected to experience a loss of agricultural suitability greater than or equal to 1% is approximately the same for each RCP and sensitivity scenario, ranging from 17.9% to 18.9%. This indicates that changes in GDD and α are approximately equally important in areas projected to experience decreasing agricultural suitability.

Discussion

Generally, the most agriculturally suitable regions projected for the end of the 21st century closely resemble the most intensely rain-fed cultivated regions today. The American Midwest, Central America, Eastern Argentina, regions directly south of the Sahara, Eastern Africa, Southern Australia, India, Indonesia, Mediterranean countries, central Europe, Southwest Russia, and Eastern China stand out as the most agriculturally suitable regions currently and in the future (Figures 2A, 3A, 3B). However, while these

regions are expected to remain suitable, they are also projected to experience the greatest relative loss of agriculturally suitable land in the future (Figure 3C, 3D). The sensitivity analysis reveals that the losses in these regions are caused mostly by changes of GDD (Figure 4B), although changes of α are also a contributing factor (Figure 4A).

Future projections predict a net increase in agriculturally suitable land, much of which can be attributed to increases in suitability in high northern latitudes of Canada and Russia (Figure 3A, 3B). The sensitivity analysis shows that this increase is due almost entirely to increased GDD, as short growing seasons are currently a major limiting factor in these cold regions (Figure 4A, 4B). The American Southwest, Northern Australia, much of sub-Saharan Africa, and South America are predicted to increase in suitability as well, mostly due to changes in GDD.

The projected future increase in agricultural suitability in high Northern latitudes and the American Southwest, as well as projected future decreases along the Mediterranean and directly south of the Sahara, agrees with previous work (Ramankutty et al., 2002; Zhang & Cai, 2011). However, this model predicts a much more heterogeneous distribution of future gains and losses compared to previous studies, likely due to a higher resolution. Additionally, this model predicts gains and losses in some regions where previous models do not: losses in Europe, the American Midwest, and China are unique to this model, as are gains in the Amazon and Northern Australia. These disparities may be caused by the differing climate models used for future projections, or by the differences in models.

The relative importance of GDD changes compared to α changes to agricultural suitability as demonstrated by the sensitivity analysis agrees with past work analyzing the

importance of temperature relative to precipitation on agricultural outcomes (Lobell & Burke, 2008). Lobell and Burke (2008) demonstrated that despite the importance of precipitation (and therefore soil moisture) to annual crop yields, changes in temperature had a greater impact on agricultural projections. This was found to be caused by the greater magnitude of climate change-induced temperature increase compared to changes in precipitation. The diminished contribution of α to the projected outcomes compared to GDD likely reflects this same pattern.

While this model predicts a 13.8%-27.8% net increase in agriculturally suitable land by 2100, major obstacles must be addressed to ensure food security of the future population. The projected global population increase to 9.8 billion people by 2050 (United Nations, Department of Economic and Social Affairs, 2017) is expected to increase agricultural demand by 50% (FAO, 2016). The challenge to meet this increasing agricultural demand could be helped by expanding agriculture to newly suitable land identified by this model; however, it is unknown if increasing the area of cultivated land will result in a directly proportional increase in food production. Additionally, expanding cultivated lands comes with severe environmental costs including deforestation, which leads to decreased biodiversity, species loss, and greatly adds to greenhouse gas emissions (FAO, 2016).

Table 1. Sources and descriptions of data used to build the model

Variable	**Source**	**Description**
Modern rain-fed cropland	GAEZ	Rainfed cropland distribution c. 2000 (mean proportion/cell)
Modern temperature	WorldClim 2	Monthly 1970-2000 mean terrestrial (mean °C)
Modern precipitation	WorldClim 2	Monthly 1970-2000 mean terrestrial (total mm)
Modern cloud cover	CRU CL v. 2.0	Monthly 1961-1990 mean terrestrial (mean percent)
Soil quality	HWSD	Seven parameters described on a categorical scale
Slope	HWSD	Mean rate of elevation change (°)

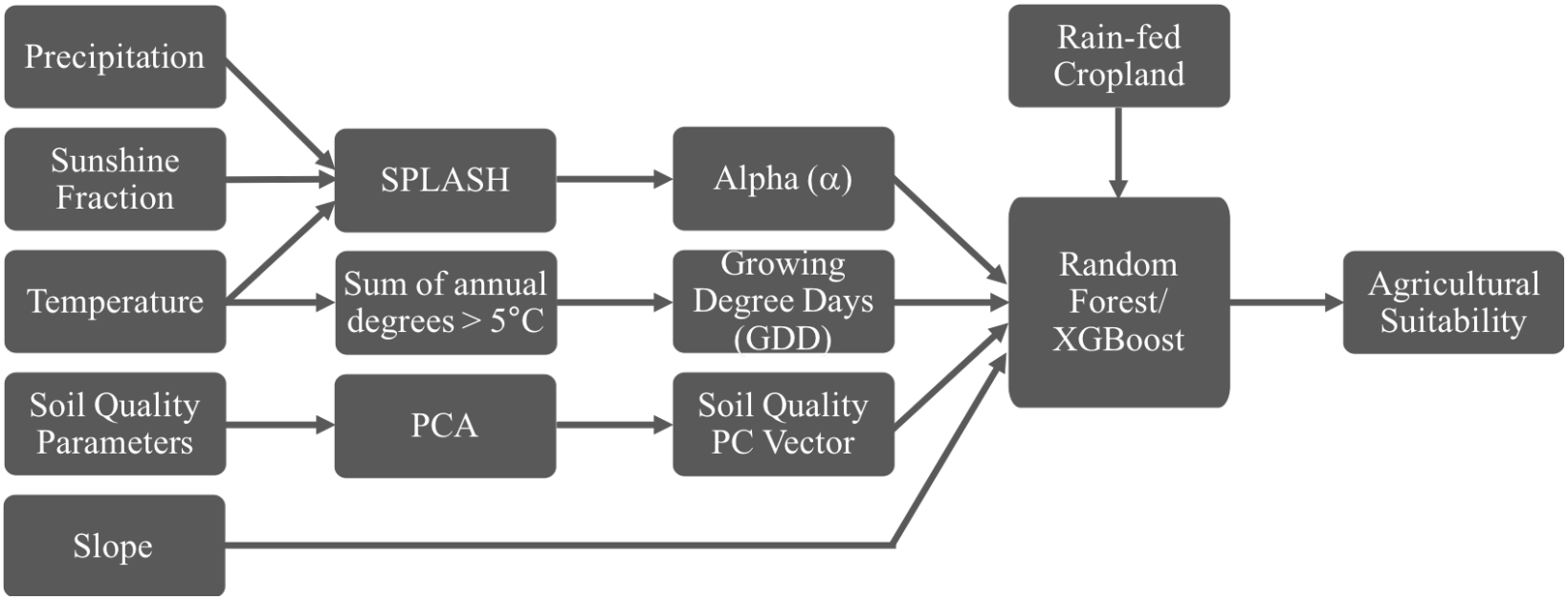

Figure 1. Flow chart of the model-building process. Raw data inputs are represented in the left-hand column, and processes used to calculate drivers are in the following column. Final drivers used as model inputs are in the middle column. The next column represents the two models built as well as the rain-fed cropland data used to train the model. The final output, agricultural suitability, is represented in the right-hand column.

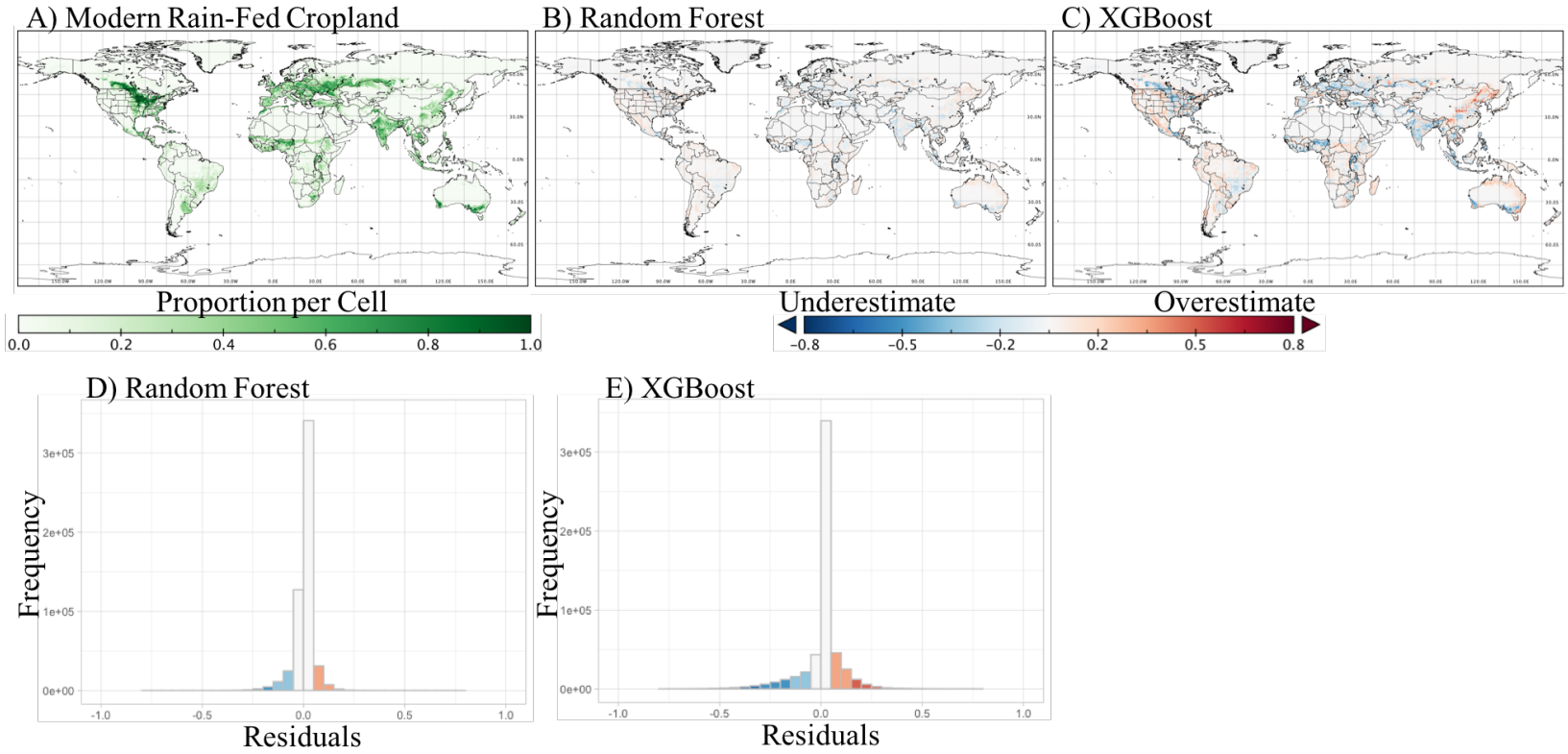

Figure 2. Comparison of agricultural suitability projections. A) A map of the distribution of modern rain-fed cropland. B) A map of the residuals between random forest projected modern agricultural suitability and modern rain-fed cropland. C) A map of the residuals between XGBoost projected modern agricultural suitability and modern rain-fed cropland. D) A histogram of the random forest residuals. E) A histogram of the XGBoost residuals. In (B-E), blue represents underestimation and red represents overestimation.

Table 2. Sources and descriptions of data used to project future agricultural suitability

Variable	**Source**	**Purpose**	**Description**
Future temperature	GISS 2.6 and 8.5	Future agricultural suitability	Projected monthly 2081-2100 mean terrestrial for RCP2.6 and RCP8.5 (mean °C)
Future precipitation	GISS 2.6 and 8.5	Future agricultural suitability	Projected monthly 2081-2100 mean terrestrial for RCP2.6 and RCP8.5 (total mm)
Modern cloud cover	CRU CL v. 2.0	Future agricultural suitability	Monthly 1961-1990 mean terrestrial (mean percent)
Modern temperature	WorldClim 2	Debias GISS temperature	Monthly 1970-2000 mean terrestrial (mean °C)
Modern precipitation	WorldClim 2	Debias GISS precipitation	Monthly 1970-2000 mean terrestrial (total mm)
Soil quality	HWSD	Future agricultural suitability	Seven parameters described on a categorical scale
Slope	HWSD	Future agricultural suitability	Mean rate of elevation change (°)

Table 3. The area of agriculturally suitable (AgS) land and changes compared to predicted modern agricultural suitability

	Suitable Land (km^2)	**Suitable Land (%)**	**AgS Change (%)**	**AgS Increase ≥ 1% (km^2)**	**AgS Increase ≥ 1% (%)**	**AgS Decrease ≥ 1% (km^2)**	**AgS Decrease ≥ 1% (%)**
Actual Rain-fed Cropland	1.28E+07	9.59%					
Predicted Modern AgS	1.29E+07	9.61%					
RCP2.6	1.46E+07	10.94%	13.82%	5.19E+07	38.79%	2.53E+07	18.88%
RCP8.5	1.65E+07	12.32%	28.20%	6.10E+07	45.60%	2.49E+07	18.60%
RCP8.5 GDD	1.64E+07	12.28%	27.76%	6.15E+07	45.96%	2.40E+07	17.92%
RCP8.5 α	1.28E+07	9.55%	-0.63%	3.22E+07	24.05%	2.53E+07	18.92%

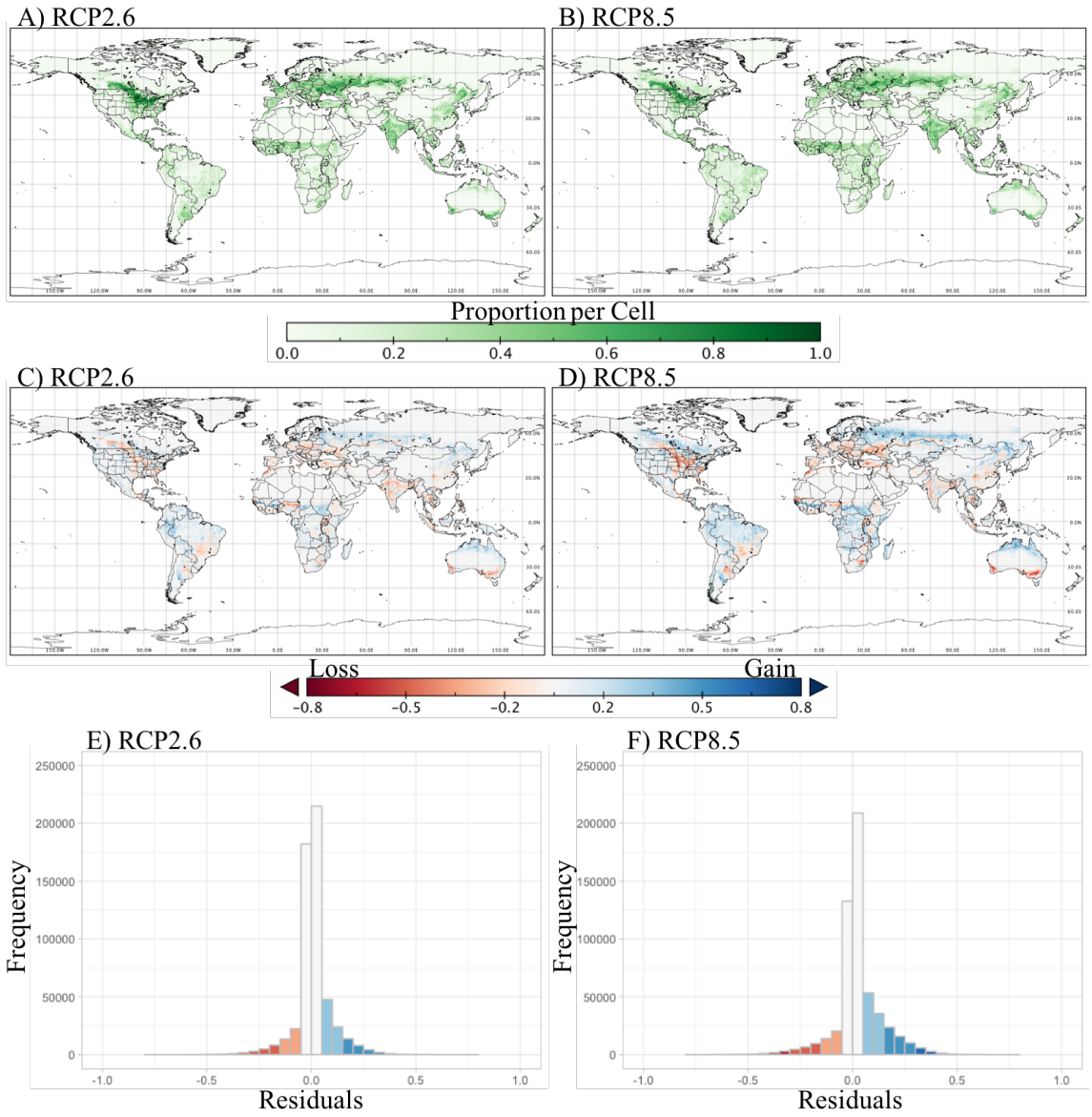

Figure 3. Distribution and residuals of future agricultural suitability. A) A map of the distribution of projected agricultural suitability under RCP2.6. B) A map of the distribution of projected agricultural suitability under RCP8.5. C) A map of the residuals between projected RCP2.6 agricultural suitability and modern projected agricultural suitability. D) A map of the residuals between projected RCP8.5 agricultural suitability and modern projected agricultural suitability. E) A histogram of the projected RCP2.6 agricultural suitability residuals. F.) A histogram of the projected RCP2.6 agricultural suitability residuals. In (C-F), red represents losses of agricultural suitability and blue represents gains.

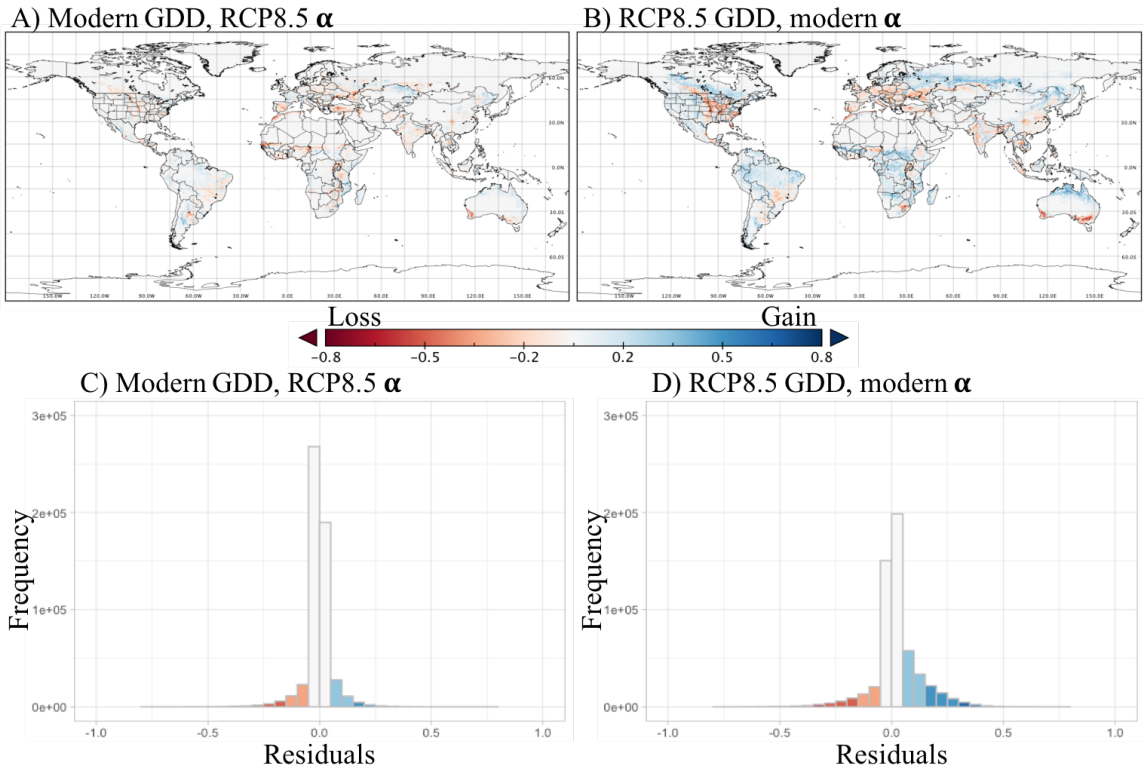

Figure 4. Sensitivity analysis of future projections. A) A map of the residuals between agricultural suitability projected with modern values of GDD and RCP8.5 values of α and modern projected agricultural suitability. B) A map of the residuals between agricultural suitability projected with RCP8.5 values of GDD and modern values of α and modern projected agricultural suitability. C) A map of the residuals between modern GDD/RCP8.5 α and modern agricultural suitability. D) A map of the residuals between RCP8.5 GDD/modern α and modern agricultural suitability.

CHAPTER 3

AGRICULTURAL SUITABLITY AND STREAMFLOW IN RANGE CREEK CANYON, UTAH

Introduction

Agriculture in Range Creek Canyon, UT was likely a challenge for the Fremont farmers that occupied the canyon from AD 900-1200. Today, with a mean GDD of 1823 and a mean α of 0.6, the semi-arid climate and short growing season severely limits agriculture to the point that maize farming is impossible without irrigation (Boomgarden, 2015). Predicting the agricultural suitability of Range Creek Canyon during the time of the Fremont occupation will provide an estimate of the probability of successful rain-fed cultivation. Supplementing the agricultural suitability model with a streamflow model will allow the irrigation potential to be estimated, and together, these models can give quantifiable insight to the agricultural conditions faced by the Fremont farmers.

The objective of this chapter is to apply the agricultural suitability model described in Chapter 2 along with a streamflow model to study how climate may have impacted agriculture in Range Creek Canyon during the time of the Fremont occupation. Because water is known to be the major limiting factor in maize farming in Range Creek Canyon today, three scenarios were used to investigate a range of conditions the Fremont farmers may have experienced: a “wet” scenario, a “dry” scenario, and one scenario

representing the mean-state from AD 900-1200. These scenarios were compared to modern conditions for context.

Data

The Community Earth System Model-Last Millennium Ensemble (CESM-LME) (Otto-Bliesner et al., 2016) provides fully forced global daily temperature and precipitation simulations from AD 850-2005 compiled from proxy records worldwide at a 2° resolution. This data set is the most appropriate option for this study because it provides modeled temperature data that are otherwise unavailable for the time period of interest. Because all models present some form of bias and projections differ from observations, best-practice is to debias modeled climate data (Wilby et al., 2004). The CESM-LME data were debiased using the change-factor approach (Wilby et al., 2004) by subtracting CESM modeled modern data from CESM modeled paleo data, then added to the WorldClim 2 modern observed climatology (Fick & Hijmans, 2017). By debiasing the modeled climate data, the WorldClim 2 climatology acts as a correct baseline and spatial pattern, and the CESM-LME data provide changes in the mean-state and variability. The modern portion of the debiased time series had greater variability than modern observations, but the means were very similar.

Temperature and precipitation from CESM-LME were used to calculate a monthly time series of GDD and α and PET (for use in the streamflow model) with the SPLASH methodology (Davis et al., 2016) for Range Creek Canyon (-110.25 E, 39.48 N). As in Chapter 2, modern cloud cover from CRU CL v. 2.0 (New et al., 2002) was used and repeated for the length of the time series as it is one of the only estimates of

high-resolution sunshine fraction available. Soil quality and slope were assumed to be unchanging, so the same data sets from the HWSD (FAO et al., 2012) were used to predict past suitability.

Because modern streamflow data from Range Creek Canyon are currently very sparse, modeled streamflow was calibrated with observed streamflow from 1986-2016 collected from the Green River at Green River site (-110.09 E, 38.59 N), the nearest downstream USGS data collection site from the confluence of Range Creek and Green River (U.S. Geological Survey, 2016). These data were aggregated from daily to monthly values, and the units were converted from m^3/day to mm/month by dividing mm^3/month by the area of the Range Creek water basin (3.75e+14 mm^2). The data were then scaled with the available Range Creek discharge data (Potter, 2016), consisting of 10 data points recorded in the summers of 2015 and 2016. The scalar was the slope of linear regression between the two data sets with the y-intercept set to 0 (Figure 5). Modern mean monthly temperature, total monthly precipitation, and total monthly PET were used from CRU TS vs. 4.01 (Harris, Jones, Osborn, & Lister, 2014), debiased with WorldClim2 climatology. To account for the effects of snow for past and modern streamflow, a method described by Gerten et al., (2004) was used to calculate "rain," the combination of snowmelt and precipitation runoff available to enter the stream. Rain was used in place of precipitation in the models for calibration and simulation. In the future, this model will greatly benefit from calibration with longer and more consistent streamflow data from Range Creek. The data sources and descriptions used in this section are summarized in Table 4.

Methods

All analyses were conducted in R (R Core Team, 2017). A time series of α from AD 900-1200 extracted from the coordinates -110.25 E, 39. 48 N was constructed with a Loess smoother spanning 3% of the data. The minimum and maximum of the smoothed time series were identified (Figure 6), and 10-year time series surrounding them served as the dry and wet scenarios (AD 1051-1060 and AD 1174-1183, respectively). A 10-year time series surrounding the mean-state was also selected (AD 992-1001), and the three scenarios were compared to the most recent 10-year time series in the CESM data set (AD 1996-2005).

Experiment 1: Agricultural Suitability in Range Creek Canyon

The agricultural suitability model built with a random forest described in Chapter 2 was used to predict a monthly time series of agricultural suitability values from AD 850-2005 using past GDD (annual values repeated each month of each respective year), past α, modern soil quality, and slope as inputs. Agricultural suitability during the modern, mean, dry, and wet scenarios was isolated. To compensate for the greater variability in the past modeled data compared to observed data, the proportional difference between the mean of the modern scenario and the means of the mean, wet, and dry scenarios were calculated and used to scale modern agricultural suitability to produce estimates of all scenarios. This method was also used to scale modern GDD and α for all scenarios.

Experiment 2: Streamflow and Irrigation Potential in Range Creek Canyon

Past streamflow in the mean, dry, and wet scenarios was simulated and compared to the modern scenario. The proportional differences between the mean of the modern scenario and the means of the mean, wet, and dry scenarios were calculated and used to scale modern observed streamflow for each scenario to offset the greater variability found in past modeled data. The resulting streamflow scenarios were used to calculate July irrigation potential, the month Boomgarden (2015) found that the maize reproductive phase occurred in the experimental plots in Range Creek Canyon.

Streamflow models use climate variables and relevant data about a watershed to predict streamflow (also called discharge) at the basin outlet. Many models used to predict streamflow run on a daily time-step and require detailed data related to catchment morphology and initial state (Devia, Ganasri, & Dwarakish, 2015). However, detailed data for the study site and time period are not available, so a reasonable option for predicting past streamflow is a monthly model with few data requirements. The GR2M model (Mouelhi, Michel, Perrin, & Andréassian, 2006) implemented in this study runs on a monthly time-step, requires three inputs (temperature, precipitation, and potential evapotranspiration), and can be calibrated on modern streamflow.

The GR2M model was empirically derived in a step-wise manner using streamflow data from 410 basins across the world, including semi-arid basins in the American Southwest (Mouelhi et al., 2006). It was developed with the intention to improve existing models while reducing their complexity, and ultimately to produce a simple monthly model with broad applicability. GR2M requires total monthly precipitation and total monthly PET and has two free parameters that should be

calibrated: maximum soil storage capacity and the outside exchange coefficient. The authors describe the model as possessing "three-fold lumping" consisting of spatial, temporal, and conceptual lumping. Spatial lumping occurs because precipitation and PET are averaged across the entire basin, so relative wet and dry areas within the basin are not considered separately. Temporal lumping occurs because precipitation and PET are summed into an entire month so that individual precipitation events are not considered individually. Conceptual lumping occurs because biophysical processes involved are not considered individually. Lumping sacrifices accuracy to some extent but allows the complexity of the model to be greatly reduced.

Described generally, each month's streamflow is calculated starting with the previous month's excess soil moisture storage and reservoir storage. Water is added to the system in the form of precipitation, and water is removed from the system in the form of potential evapotranspiration. Some water may remain in the system as determined by the maximum store capacity (a free parameter) and the reservoir storage (set to 60 mm), and water is either added or subtracted to the system by the outside exchange coefficient (a free parameter) which represents groundwater contribution or loss. Excess water exits the system in the form of streamflow at the basin outlet.

The two free parameters in the GR2M model were optimized with a genetic algorithm using the GA package in R (Scrucca, 2013). Genetic algorithms are designed to mimic natural selection to find the optimal set of parameters given a "fitness" function, and are commonly used to calibrate hydrological models (i.e., Wang, 1991). The genetic algorithm first creates multiple parameter sets by selecting random values for each parameter and uses them to run the model. It then evaluates the outputs compared to

actual values using the fitness function. The highest performing parameter sets are kept (parent parameter sets), and a series of new parameter sets (child parameter sets) are created from combinations of the parent parameter sets. The child parameter sets are "mutated" by randomly changing one or more parameter values, and the process is repeated until a stopping criterion is met.

The fitness function used for the genetic algorithm in this experiment is the root mean squared error (RMSE), and the optimal values were found to be 752.08 mm for the maximum soil store capacity, and 0.92105 for the outside exchange coefficient. The resulting simulated streamflow had a RMSE of 0.309. Although the simulated streamflow does not fully capture the amplitude of the peaks of the observed streamflow, the model adequately captures annual fluctuation and is a great approximation considering the observed streamflow is proxy data (Figure 7).

Irrigation potential was calculated via the methods described by Potter (2016). The streamflow data were first reduced to 38% of the total to account for losses to evapotranspiration and soil percolation during simple open irrigation. The values were then converted to ha/day·0.81 cm. This unit provides the hectares the daily streamflow could irrigate to a depth of 0.81 cm, the daily water requirement for maize during the reproductive stage.

Results

Experiment 1: Agricultural Suitability in Range Creek Canyon

The proportional differences between modeled modern scenarios and modeled mean, wet, and dry scenarios used to scale modern observed data are summarized in

Table 5. Past mean agricultural suitability was very low in each scenario examined, ranging from 0.0492 to 0.0644 (Figure 8A). Surprisingly, the wet scenario resulted in the lowest agricultural suitability, while the modern scenario resulted in the highest agricultural suitability. This was unexpected because α was hypothesized to be the limiting factor of suitability in Range Creek Canyon. However, the result can be explained by examining α and GDD across the scenarios: Increased α correlates with decreased GDD, a variable shown to have a larger impact on agricultural suitability by the sensitivity analysis in Chapter 2. The relative difference in mean α between the wet (0.676) and dry (0.597) scenarios is small compared to the difference in GDD (1537 and 1747, respectively) (Figure 8B and 8C). The modern scenario had a greater mean GDD than the dry scenario (1823) as well as a greater mean α (0.612) than the dry scenario, which resulted in the highest agricultural suitability. The mean agricultural suitability for the mean scenario was 0.0591, right between the dry and wet scenario means.

Experiment 2: Streamflow and Irrigation Potential in Range Creek Canyon

The proportional differences between the modeled modern scenario and modeled mean, wet, and dry scenarios used to scale modern observed streamflow data are summarized in Table 5. The medians of the streamflow scenarios ranged from 0.208 mm/month (dry) to 0.603 mm/month (wet) (Figure 9A). The median streamflow of the mean scenario was 0.383 mm/month, and at 0.234 mm/month, the median modern streamflow closely resembled the dry scenario streamflow.

These streamflow values resulted in median irrigation potentials ranging from 16.6 ha (dry) to 48.1 ha (wet) (Figure 9B). Just like the streamflow medians, the median

irrigation potential of the modern scenario (18.7 ha) resembled the dry scenario, and the median irrigation potential of the median scenario fell between modern and wet at 30.6 ha. The medians of streamflow and irrigation potential were considered rather than the means because the means were influenced by a particularly high outlier in the observed streamflow data. Although the outlier is an example of real variability, proportionally increasing the data (as in the mean and wet scenarios) may have compounded the outlier beyond normal variability.

Discussion

The results of Experiment 1 indicate that not only does the agricultural suitability model agree with the experimental results of Boomgarden (2015) that modern rain-fed maize agriculture is impossible in Range Creek Canyon, it also suggests that rain-fed maize agriculture was unlikely to succeed from AD 900-1200 even during anomalously wet periods. The mean scenario resulted in a mean agricultural suitability value of 0.0591, indicating that the mean probability of rain-fed cultivation in Range Creek Canyon throughout the Fremont occupation was 5.91%. The highest value of agricultural suitability across the three past scenarios was 0.175 in the dry scenario, still a very small probability of rain-fed cultivation.

While the projected agricultural suitability values were too similar to meaningfully differentiate between scenarios, the results do reveal meaningful information about rain-fed agriculture in Range Creek Canyon. Interestingly, the wet scenario produced low agricultural suitability despite some α values reaching 1.0 (indicating fully saturated soil). This result emphasizes the severe limitations imposed by

low GDD, and suggests that short growing seasons in Range Creek Canyon may have been at least as much of an obstacle to successful cultivation as soil moisture. Additionally, this result points out a limitation of the agricultural suitability model that can be addressed in the future.

Modern streamflow compared to the past mean scenario provides further evidence that current conditions are dry compared to mean conditions during the Fremont occupation, and are more comparable to streamflow in the dry scenario. The current irrigation potential is therefore a low-end estimate of past irrigation potential. However, even the median irrigation potential of the wet scenario, at 30.6 ha, seems like a small area considering the scale of the Fremont occupation. The mean irrigation potential of the mean scenario suggests that Range Creek had the potential to irrigate a median of only 30.6 ha annually from AD 900-1200.

The estimates of past streamflow and irrigation potential may be low due in part to weaknesses in the streamflow model calibration. Using scaled streamflow from Green River at Green River as a proxy for Range Creek streamflow certainly resulted in inaccuracies. This obstacle can easily be overcome in the future by calibrating the model with actual Range Creek streamflow data. Additionally, groundwater is a vital component of streamflow that is not yet understood in Range Creek. The optimal value of the outside exchange coefficient, the parameter that accounts for groundwater gains or losses, in the GR2M model was found to be 0.92105 when calibrated with the scaled Green River streamflow data. Because the value is less than one, it works as an overall loss of streamflow to groundwater that acts in contradiction to the current understanding of Range Creek streamflow (Potter, 2016). Additionally, it is important to consider that

groundwater contribution changes over time as a function of decadal-scale climate changes. Future work that defines this relationship in Range Creek could be utilized to calculate a dynamic outside exchange coefficient and greatly improve past streamflow and irrigation potential estimates.

Combining the insights gained from the agricultural suitability model and the streamflow model suggests that maize farming in Range Creek Canyon during the Fremont occupation was an arduous task, even with irrigation. A long and hot enough growing season was likely at least as important as soil moisture to maximizing crop success, an occurrence that correlates with lower soil moisture, streamflow, and irrigation potential. The ideal scenario for a Fremont farmer was likely an anomalously wet, warm and long summer, or an anomalously warm and long summer with enough groundwater contribution to the streamflow to counteract corresponding decreased soil moisture.

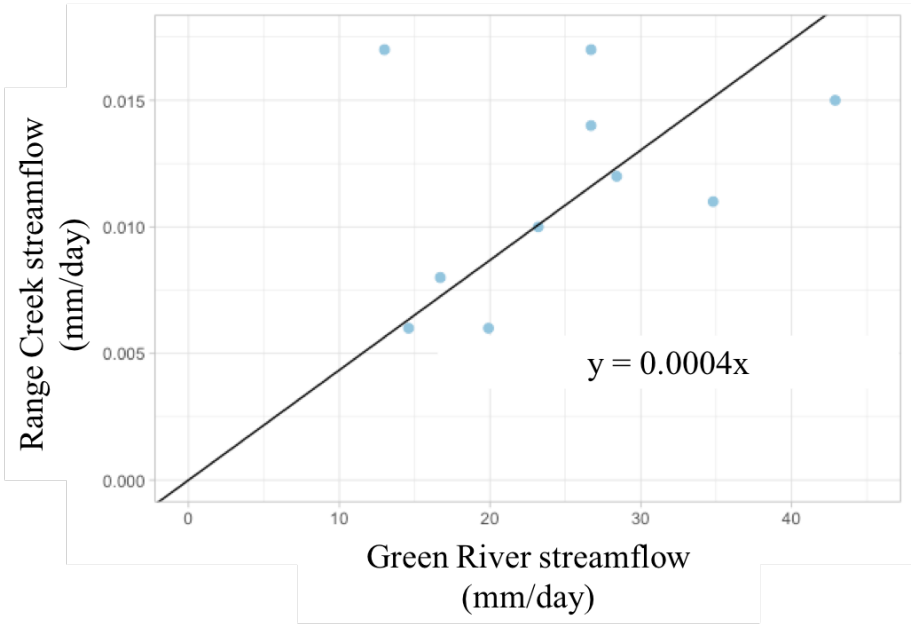

Figure 5. Linear regression of streamflow. Green River at Green River streamflow is on the x-axis and Range Creek streamflow is on the y-axis and the y-intercept is set to 0. The slope of the regression line, 0.0004, is used to scale Green River streamflow data for use in GR2M calibration.

Table 4. Sources and descriptions of data used to project past agricultural suitability and streamflow

Variable	**Source**	**Purpose**	**Description**
Past temperature	CESM-LME	Past agricultural suitability and streamflow	Projected daily AD 850-2005 mean terrestrial temperature (°C)
Past precipitation	CESM-LME	Past agricultural suitability and streamflow	Projected daily AD 850-2005 total precipitation (mm)
Modern temperature	CRU TS v. 4.01	GR2M modern streamflow calibration and modern agricultural suitability	Daily mean terrestrial (°C)
Modern precipitation	CRU TS v. 4.01	GR2M modern streamflow calibration and modern agricultural suitability	Daily total terrestrial (mm)
Modern PET	CRU TS v. 4.01	GR2M modern streamflow calibration	Daily total terrestrial (mm)
Modern cloud cover	CRU CL v. 2.0	Past PET and α	Monthly 1961-1990 mean terrestrial (mean percent)
Modern temperature climatology	WorldClim 2	Debias CESM-LME and CRU TS v. 4.01 temperature	Monthly 1970-2000 mean terrestrial (mean °C)
Modern precipitation climatology	WorldClim 2	Debias CESM-LME and CRU TS v. 4.01 precipitation	Monthly 1970-2000 mean terrestrial (total mm)
Soil quality	HWSD	Past and modern agricultural suitability	Seven parameters described on a categorical scale
Slope	HWSD	Past and modern agricultural suitability	Mean rate of elevation change (°)
Green River streamflow	USGS	Modern streamflow and irrigation potential	Daily streamflow 1986-2016 (m^3/day)
Range Creek streamflow	Potter (2015)	Scale Green River streamflow	Streamflow from 10 days in 2015 and 2016 (m^3/day)

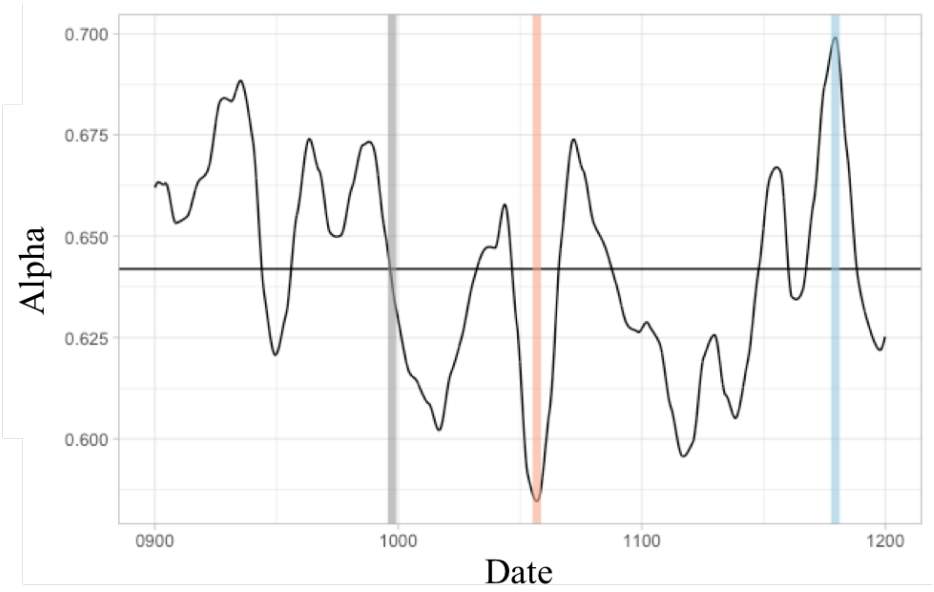

Figure 6. Time series of Loess-smoothed monthly α values in Range Creek Canyon from AD 900-1200. The red vertical line is centered on the series minimum in AD 1056, blue vertical line is centered on the series maximum in AD 1179, and the gray vertical line is centered on AD 997, the selected example of mean conditions during this time series.

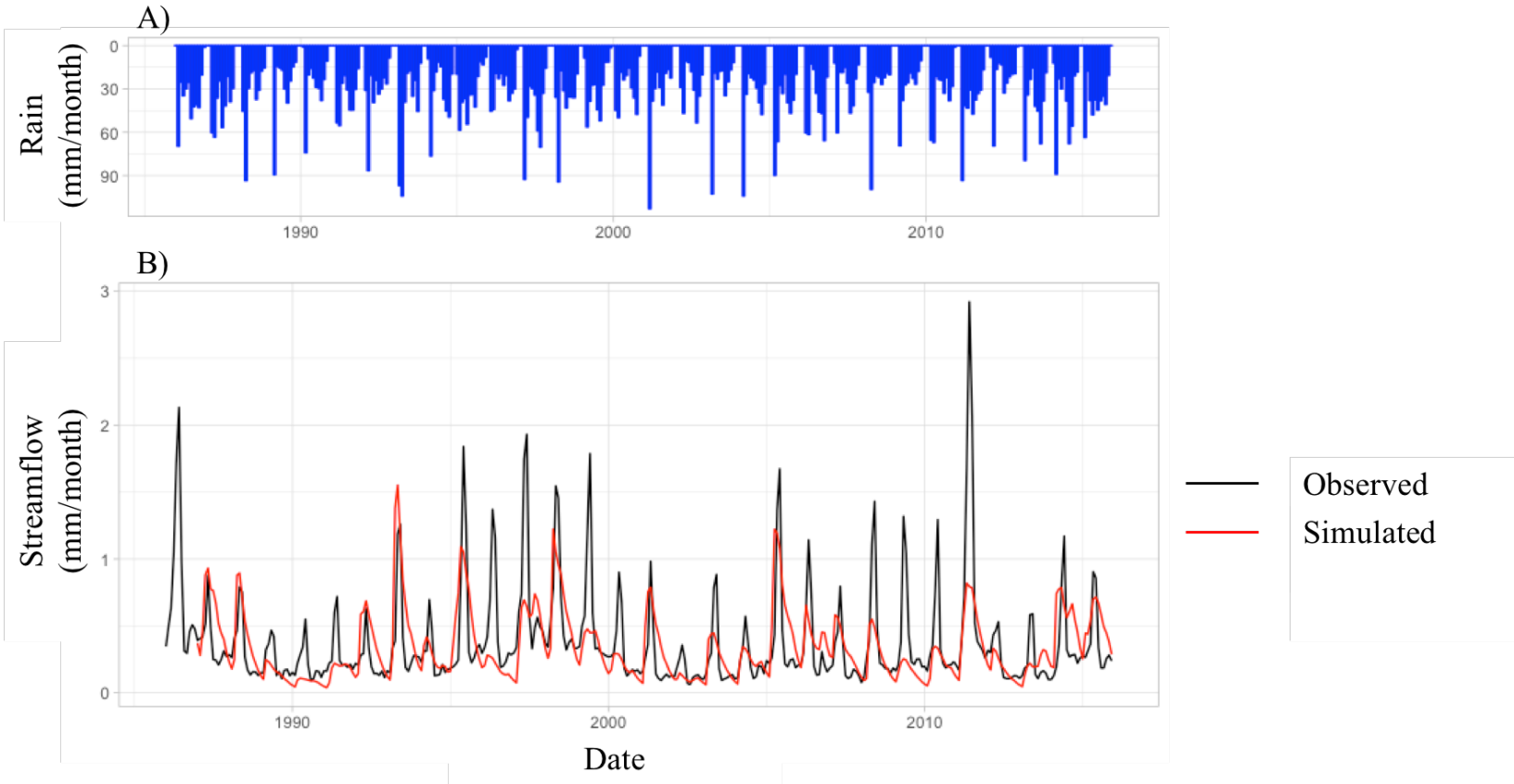

Figure 7. Streamflow calibration results. A) A time series of monthly rain values in Range Creek Canyon from AD 1986-2016. B) A time series of observed (black) and simulated (red) streamflow.

Table 5. Differences between modern modeled means and the means of past scenarios

	Modern AgS	**Modern α**	**Modern GDD**	**Modern Streamflow**
Past Mean	-8.27%	5.84%	-8.44%	63.4%
Past Wet	-23.5%	17.2%	-15.7%	157%
Past Dry	-6.55%	-2.58%	-4.17%	-11.4%

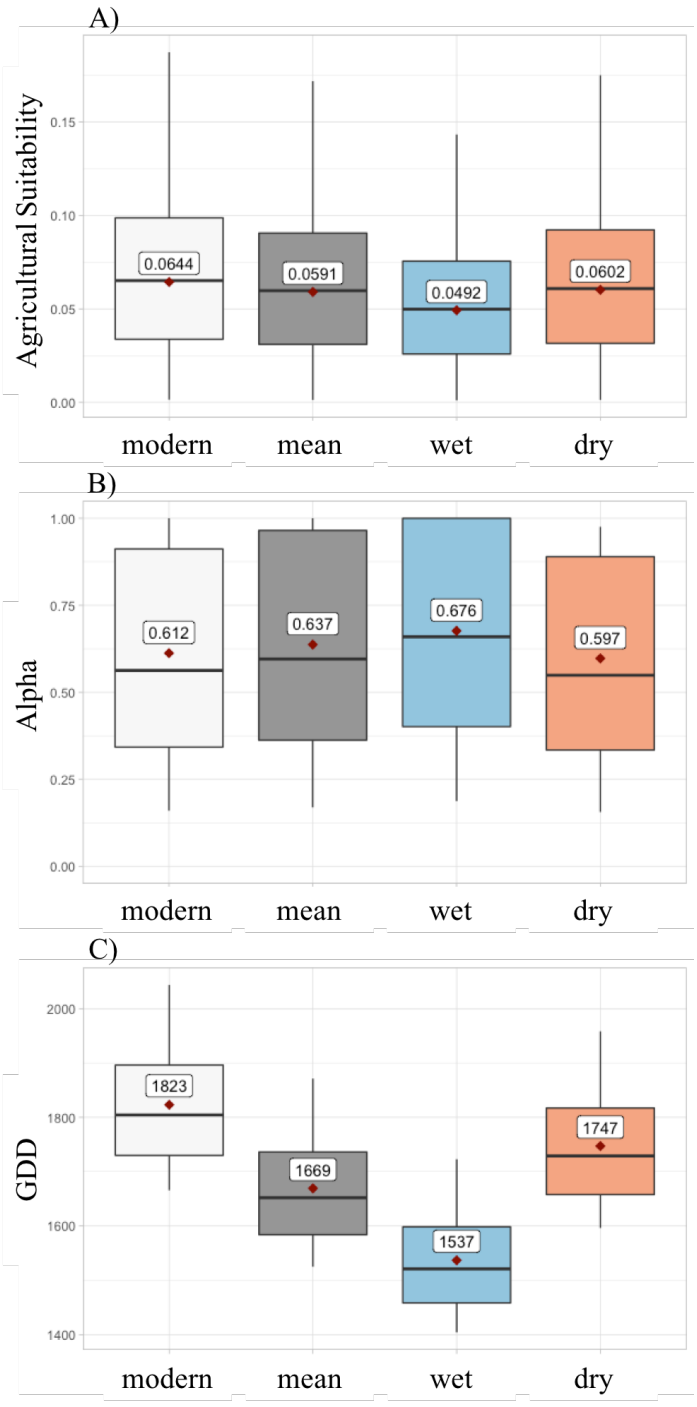

Figure 8. Boxplots of Range Creek Canyon agricultural suitability, α, and GDD. The boxes in A) agricultural suitability, B) α, and C) GDD show for four scenarios: modern (displayed in white), mean (displayed in gray), wet (displayed in blue), and dry (displayed in red). The dark red points and text in each box indicate the respective means.

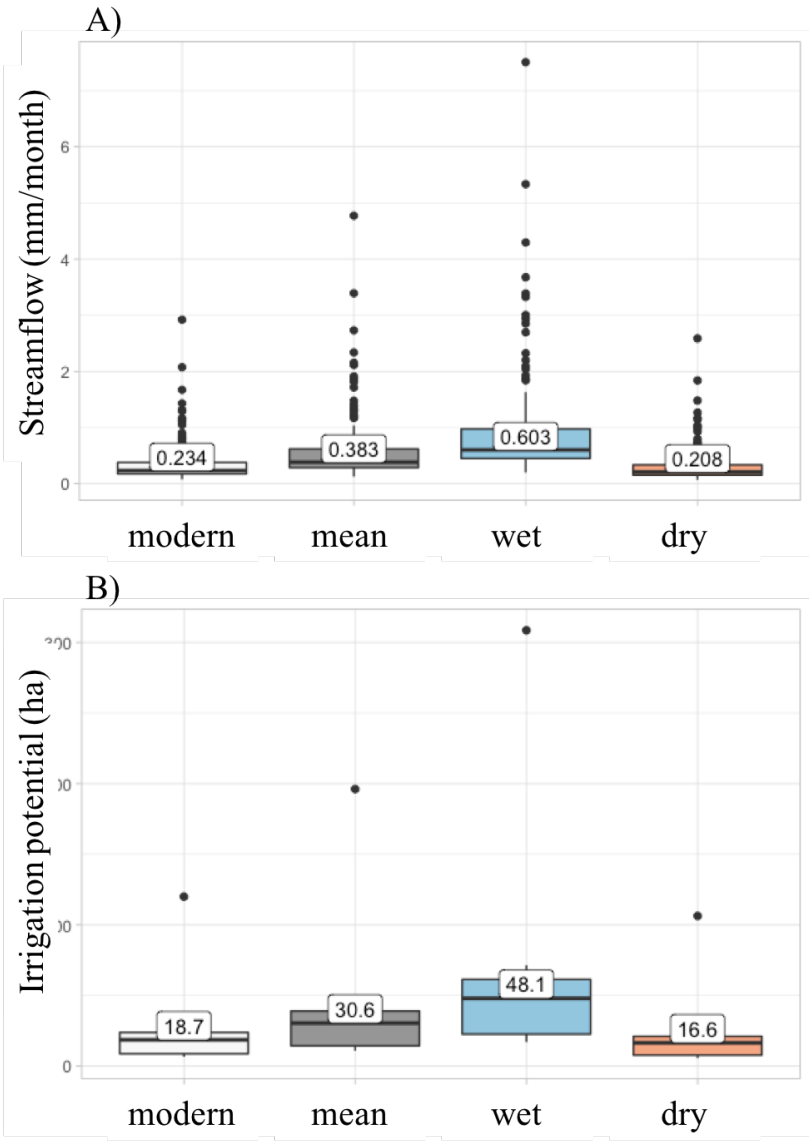

Figure 9. Boxplots of Range Creek Canyon streamflow and irrigation potential. The boxes in A) streamflow and B) July maize irrigation potential for Range Creek show four scenarios: modern (displayed in white), mean (displayed in gray), wet (displayed in blue), and dry (displayed in red). The text in each box indicates the respective medians.

CHAPTER 4

CONCLUSION

This thesis presented a methodological approach to the development and implementation of an agricultural suitability model. Using a random forest to build the agricultural suitability model along with updated, high-resolution climate, soil, and slope parameters increased the accuracy and resolution of the suitability model presented by Ramankutty et al. (2002) from which this model was derived. In the future, the model can certainly be improved by accounting for CO_2, and may be improved by using growing season α instead of mean annual α, and using stratified sampling to train the model.

Predicting the impacts of climate change on agricultural suitability is an important application of this model, which was demonstrated here at a global scale. Understanding how the distribution of suitable land is likely to shift provides a tool that could be used to adapt to climate change, which will be necessary to meet the increasing agricultural demands of the growing population. While the model predicted a net increase in agriculturally suitable land by the end of the century, a large proportion of land is expected to decrease in suitability as well. It will be important to consider the distribution of predicted gains and losses to guard against future food insecurity.

The second application of the agricultural suitability model demonstrated here was to predict prehistoric suitability at a single location. While statistical models are

known to present inaccuracies when used at a small spatial scale, a process-based model would be extremely difficult or impossible to implement in the study of prehistoric agriculture in Range Creek Canyon because of the limited data availability from AD 900-1200. This case study utilized the agricultural suitability model at a broad time scale instead of a broad spatial scale to provide insight to prehistoric farming conditions, an exercise that provided useful insights. Supplementing the agricultural suitability model with a streamflow model was an important addition to the study, as projected past agricultural suitability indicates that rain-fed agriculture would unlikely have been successful during the Fremont occupation. The results presented here provide the first quantifiable insight into past irrigation potential for Range Creek Canyon during the Fremont occupation, as well as a method to improve the results as knowledge of Range Creek streamflow improves.

REFERENCES

Ainsworth, E. A., & Rogers, A. (2007). The response of photosynthesis and stomatal conductance to rising [CO 2]: Mechanisms and environmental interactions. *Plant, Cell and Environment*, *30*(3), 258–270. https://doi.org/10.1111/j.1365-3040.2007.01641.x

Baro, M., & Deubel, T. F. (2006). Persistent hunger: Perspectives on vulnerability, famine, and food security in Sub-Saharan Africa. *Annual Review of Anthropology*, *35*(1), 521–538. https://doi.org/10.1146/annurev.anthro.35.081705.123224

Benson, L. V., Berry, M. S., Jolie, E. A., Spangler, J. D., Stahle, D. W., & Hattori, E. M. (2007). Possible impacts of early 11th, middle 12th and late 13th century droughts on western Native Americans and the Mississippian Cahokians. *Quaternary Science Reviews*, *26*, 336–350.

Benson, L. V., Ramsey, D. K., Stahle, D. W., & Petersen, K. L. (2013). Some thoughts on the factors that controlled prehistoric maize production in the American Southwest with application to southwestern Colorado. *Journal of Archaeological Science*, *40*(7), 2869–2880. https://doi.org/10.1016/j.jas.2013.03.013

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J., & Lang, M. (2017). mlrMBO: A modular framework for model-based optimization of expensive black-box functions. *arXiv preprint arXiv:1703.03373*.

Boomgarden, S. A. (2015). Experimental maize farming in Range Creek Canyon, Utah (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing.

Boomgarden, S. A., Metcalfe, D., & Springer, C. (2014). Prehistoric archaeology in Range Creek Canyon, Utah: A summary of activities of the Range Creek field station. *Utah Archaeology*, *27*(1), 9–32. https://doi.org/10.1038/366388a0

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., … Sinoquet, H. (2003). An overview of the crop model STICS. *European Journal of Agronomy, 18*, 309–332. https://doi.org/10.1016/S1161-0301(02)00110-7

Challinor, A. J., Wheeler, T. R., Craufurd, P. Q., Slingo, J. M., & Grimes, D. I. F. (2004). Design and optimisation of a large-area process-based model for annual crops.

Agricultural and Forest Meteorology, *124*(1–2), 99–120. https://doi.org/10.1016/j.agrformet.2004.01.002

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. New York, NY. https://doi.org/10.1145/2939672.2939785

Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2017). xgboost: Extreme gradient boosting. Retrieved from https://cran.r-project.org/package=xgboost

Cocu, N., Harrington, R., Rounsevell, M. D. A., Worner, S. P., & Hullé, M. (2005). Geographical location, climate and land use influences on the phenology and numbers of the aphid, *Myzus persicae*, in Europe. *Journal of Biogeography*, *32*(4), 615–632. https://doi.org/10.1111/j.1365-2699.2005.01190.x

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., … Wehner, M. (2013). Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1029–1136. https://doi.org/10.1017/CBO9781107415324.024

Cramer, W. P., & Solomon, A. M. (1993). Climatic classification and future global redistribution of agricultural land. *Climate Research*, *3*, 97–110. https://doi.org/10.3354/cr003097

Davis, T. W., Prentice, I. C., Stocker, B. D., Whitely, R. J., Wang, H., Evans, B. J., … Cramer, W. (2016). Simple Process-Led Algorithms for Simulating Habitats (SPLASH v.1.0): Robust indices of radiation, evapotranspiration and plant-available moisture. *Geoscientific Model Development Discussions*, (April), 1–25. https://doi.org/10.5194/gmd-2016-49

Del Grosso, S. J., Mosier, A. R., Parton, W. J., & Ojima, D. S. (2005). DAYCENT model analysis of past and contemporary soil N2O and net greenhouse gas flux for major crops in the USA. *Soil and Tillage Research*, *83*(1 SPEC. ISS.), 9–24. https://doi.org/10.1016/j.still.2005.02.007

Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, *4*, 1001–1007. https://doi.org/10.1016/j.aqpro.2015.02.126

dos Santos, A. (2017). landsat8: Landsat 8 imagery rescaled to reflectance, radiance and/or temperature. Retrieved from https://cran.r-project.org/package=landsat8

FAO. (2016). *The future of food and agriculture – Trends and challenges.* https://doi.org/ISBN 978-92-5-109551-5

FAO, IFAD, UNICEF, WFP, & WHO. (2017). *The State of Food Security and Nutrition in the World 2017. Building resilience for peace and food security*. Rome, Italy: FAO. Retrieved from http://www.fao.org/state-of-food-security-nutrition/en/

FAO, IIASA, ISRIC, ISSCAS & JRC. (2012). *Harmonized world soil database (version 1.2)*. Rome, Italy: FAO and Laxenburg, Austria: IIASA.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302-4315.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., & Sitch, S. (2004). Terrestrial vegetation and water balance - Hydrological evaluation of a dynamic global vegetation model. *Journal of Hydrology*, *286*(1–4), 249–270. https://doi.org/10.1016/j.jhydrol.2003.09.029

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H., Izaurralde, R. C., Ort, D., … Wolfe, D. (2011). Climate impacts on agriculture: Implications for crop production. *Agronomy Journal*, *103*(1), 351–370. https://doi.org/10.2134/agronj2010.0303

Hatfield, J., Takle, G., Grotjahn, R., Holden, P., Izaurralde, R. C., Mader, T., … Liverman, D. (2014). Agriculture. In J. M. Melillo, T.C. Richmond, & G. W. Yohe (Eds.), *Climate change impacts in the United States: The third National Climate Assessment* (pp. 150-174). https://doi.org/10.7930/J02Z13FR.On

IIASA, FAO. (2012). Global Agro-ecological Zones (GAEZ v3. 0). Laxenburg, Austria: IIASA and Rome, Italy: FAO.

Jones, P. G., & Thornton, P. K. (2003). The potential impacts of climate change on maize production in Africa and Latin America in 2055. *Global Environmental Change*, *13*(1), 51–59. https://doi.org/10.1016/S0959-3780(02)00090-0

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., … Hargrea, J. N. G. (2003). An overview of APSIM , a model designed for farming systems simulation. *European Journal of Agronomy*, *18*, 267–288.

Knight, T. A., Meko, D. M., & Baisan, C. H. (2010). A bimillennial-length tree-ring reconstruction of precipitation for the Tavaputs Plateau, Northeastern Utah. *Quaternary Research*, *73*(1), 107–117. https://doi.org/10.1016/j.yqres.2009.08.002

Knutti, R., & Sedláček, J. (2012). Robustness and uncertainties in the new CMIP5

climate model projections. *Nature Climate Change*, *3*(October), 1–5. https://doi.org/10.1038/nclimate1716

Kranz, W. L., Irmak, S., Van Donk, S. J., Yonts, C. D., & Martin, D. L. (2008). Irrigation management for corn. *Neb Guide, University of Nebraska, Lincoln*, 1–4. https://doi.org/10.2489/jswc.69.3.67A

Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, *529*(7584), 84–87. https://doi.org/10.1038/nature16467

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(December), 18–22. https://doi.org/10.1177/154405910408300516

Lobell, D. B., & Burke, M. B. (2008). Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environmental Research Letters*, *3*(34007). https://doi.org/10.1088/1748-9326/3/3/034007

Lobell, D. B., & Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, *150*(11), 1443–1452. https://doi.org/10.1016/j.agrformet.2010.07.008

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science (New York, N.Y.)*, *333*(6042), 616–620. https://doi.org/10.1126/science.1204531

Long, S. P. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO2 concentrations: Has its importance been underestimated? *Plant, Cell & Environment*, *14*(8), 729–739. https://doi.org/10.1111/j.1365-3040.1991.tb01439.x

Makowski, D., Asseng, S., Ewert, F., Bassu, S., Durand, J. L., Li, T., … Zhu, Y. (2015). A statistical analysis of three ensembles of crop model responses to temperature and CO2 concentration. *Agricultural and Forest Meteorology*, *214–215*, 483–493.

Mall, R. K., Singh, R., Gupta, A., Srinivasan, G., & Rathore, L. S. (2006). Impact of climate change on Indian agriculture: A review. *Climatic Change*, *78*(2–4), 445–478. https://doi.org/10.1007/s10584-005-9042-x

McGrath, J. M., & Lobell, D. B. (2011). An independent method of deriving the carbon dioxide fertilization effect in dry conditions using historical yield data from wet and dry years. *Global Change Biology*, *17*(8), 2689–2696. https://doi.org/10.1111/j.1365-2486.2011.02406.x

Morgan, P. B., Ainsworth, E. A., & Long, S. P. (2003). How does elevated ozone impact soybean? A meta-analysis of photosynthesis, growth and yield. *Plant, Cell and*

Environment, *26*(8), 1317–1328. https://doi.org/10.1046/j.0016-8025.2003.01056.x

Mouelhi, S., Michel, C., Perrin, C., & Andréassian, V. (2006). Stepwise development of a two-parameter monthly water balance model. *Journal of Hydrology*, *318*(1–4), 200–214. https://doi.org/10.1016/j.jhydrol.2005.06.014

Nazarenko, L., Schmidt, G. A., Miller, R. L., Tausnev, N., Kelley, M., Ruedy, R., … Zhang, J. (2015). Future climate change under RCP emission scenarios with GISS ModelE2. *Journal of Advances in Modeling Earth Systems*, *7*(1), 244–267. https://doi.org/10.1002/2014MS000403

New, M., Lister, D., Hulme, M., & Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, *21*(1), 1–25.

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., … Strand, G. (2016). Climate variability and change since 850 ce an ensemble approach with the community earth system model. *Bulletin of the American Meteorological Society*, *97*(5), 787–801. https://doi.org/10.1175/BAMS-D-14-00233.1

Porter, J. R., Xie, L., Challinor, A. J., Cochrane, K., Howden, S. M., Iqbal, M. M., … Travasso, M. I. (2014). Food security and food production systems. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 485–533. https://doi.org/10.1111/j.1728-4457.2009.00312.x

Potter, D. (2016). *Range Creek Stream Flow*. University of Utah, Salt Lake City, Utah.

Prentice, I. C., Bartlein, P. J., & Webb, T. (1991). Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology*, *72*(6), 2038–2056. Retrieved from http://www.jstor.org/stable/1941558

Priestly, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *100*(2), 81–92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria. https://doi.org/ISBN 3-900051-07-0

Ramankutty, N., Foley, J. A., Norman, J., & McSweeney, K. (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography*, *11*(5), 377–392.

Rosenzweig, C., Jones, J. W., Hatfield, J. L., Ruane, A. C., Boote, K. J., Thorburn, P., … Winter, J. M. (2013). The Agricultural Model Intercomparison and Improvement

Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, *170*, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, *106*(37), 15594–15598. https://doi.org/10.1073/pnas.0906865106

Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, *53*(4), 1–37.

Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., … Naik, N. (2007). Model projections of an imminent transition to a more arid climate in southwestern North America. *Science*, *316*(5828), 1181–4. https://doi.org/10.1126/science.1139601

Sheffield, J., & Wood, E. F. (2008). Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Climate Dynamics*, *31*(1), 79–105. https://doi.org/10.1007/s00382-007-0340-z

Taub, D. R., Miller, B., & Allen, H. (2008). Effects of elevated CO2 on the protein concentration of food crops: A meta-analysis. *Global Change Biology*, *14*(3), 565–575. https://doi.org/10.1111/j.1365-2486.2007.01511.x

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

The World Bank. (2017). Fragility, Conflict, and Violence. Retrieved December 21, 2017, from http://www.worldbank.org/en/topic/fragilityconflictviolence/overview

Tubiello, F. N., Soussana, J.-F. J.-F., & Howden, S. M. (2007). Crop and pasture response to climate change. *Proceedings of the National Academy of Sciences*, *104*(50), 19686–19690. https://doi.org/10.1073/pnas.0701728104

U.S. Geological Survey. (2016). National Water Information System. Retrieved from https://waterdata.usgs.gov/nwis/

United Nations, Department of Economic and Social Affairs, P. D. (2017). *Key findings & advance tables*. *World Population Prospects: The 2017 Revision*. https://doi.org/10.1017/CBO9781107415324.004

von Caemmerer, S., & Furbank, R. T. (2003). The C4 pathway: An efficient CO2 pump. *Photosynthesis Research*, *77*(2), 191. https://doi.org/10.1023/A:1025830019591

Wang, Q. J. (1991). The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research*, *27*(9), 2467–2471.

Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., & Mearns, L. O. (2004). Guidelines for use of climate scenarios developed from statistical downscaling methods. *Analysis*, *27*(August), 1–27. https://doi.org/citeulike-article-id:8861447

Witze, A. (2016, July 11). Clouds get high on climate change. *Nature News*. https://doi.org/doi:10.1038/nature.2016.20230

Zhang, X., & Cai, X. (2011). Climate change impacts of global agricultural land availability. *Environmental Research Letters*, *6*(14014). https://doi.org/10.1088/1748-9326/6/1/014014