The Impacts of Climate Change via Robust Optimization: Two Applications in Land Investment and Electricity Storage Systems

by

Zhenggao Wu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Jianxin Guo Associate Professor, Institutes of Science and Development, Chinese Academy of Sciences
Supervisor(s):	Stanko Dimitrov Professor, Dept. of Management Science and Engineering, University of Waterloo
	J. Michael Pavlin Associate Professor, Lazaridis School of Business and Economics, Wilfrid Laurier University
Internal Member:	Qi-Ming He Professor, Dept. of Management Science and Engineering, University of Waterloo
	Xuan Zhao Professor, Lazaridis School of Business and Economics, Wilfrid Laurier University
Internal-External Member:	Christopher G. Fletcher Associate Professor, Dept. of Geography and Environmental Management, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Effectively adapting to a changing climate involves making appropriate operational decisions based on long-term climate forecasts. This dissertation presents a comprehensive framework that combines climate data, regression models, and robust optimization models to examine the decision-making process for adapting to climate change over long time horizons. The research includes two projects: one focuses on studying land investment decisions, and the other investigates the operations of electricity storage systems, both considering the impacts of climate change.

Project 1: Climate change affects agricultural inputs, like temperature and precipitation, and further affects the economic output of farmland. In this study, we focus on formulating effective policies to aid various stakeholders, including investors and farmers, in adapting to the climate-induced impacts on farmland investment in the Mississippi River Basin (MRB) by using well-known climate models. Each climate model generates a unique climate forecast, and based on these forecasts, we compute a range of farmland values for the MRB. Utilizing these ranges, we apply a robust optimization model to study the optimal investment policies under varying levels of conservatism, representing the extent to which farmland assets are constrained to adopt worst-case values. We show that the optimization model can be linearized and can scale to long time frames, about 50-plus years, and sets of assets. The case study of investment in the MRB covers the years 2023-2090 and uses trajectories of land values determined for each climate scenario using a regression model. Our empirical study shows that there is a disagreement between popular climate forecasts that influence land investment and may affect the most profitable land

investments.

Project 2: The effects of climate change on energy markets are diverse, encompassing changes in demand patterns and supply dynamics, particularly concerning the increasing penetration of renewable energy. These changes impact the dynamics of energy supply from renewable sources, such as wind and solar, leading to increased intermittency. Battery energy storage systems (BESSs) present a promising solution to effectively manage this intermittency from renewable energy sources. However, their profitability and incentive to participate in markets under climate change are susceptible to both the magnitude and frequency of price variation. This project investigates the impact of climate change on a BESS operating in a North American deregulated electricity market. We propose a robust optimization model to determine the operating policy of a BESS over 80 years (from 2021 to 2100) under different climate projections. We reformulate the robust optimization model to an equivalent linear program that allows us to numerically explore the operations of the BESS over the time horizon. Our empirical study analyzes the optimal arbitrage operations of the BESS in the Midcontinent Independent System Operator market in the United States, using the proposed robust model and trajectories of electricity prices determined for each climate scenario by a regression model. Additionally, we introduce a downscaling method to adjust climate scenarios to the desired resolutions for predicting electricity prices through the regression model. The results of the robust model reveal significant variations in the operating incomes of the BESS across different geographical locations and climate scenarios, highlighting the need for tailored strategies adapting to climate-induced variations in energy markets.

The findings from both projects underscore the critical significance of considering a

wide range of climate scenarios, encompassing detailed temporal and spatial data when assessing climate adaptation decisions.

Note that I used ChatGPT to proofread my dissertation, primarily for grammar and spelling checks. I do not use ChatGPT to generate content for the dissertation, as it is prohibited by the University of Waterloo's policy.

Acknowledgements

I would first like to acknowledge my supervisors, Prof. Stan Dimitrov and Prof. Michael Pavlin, who support me, guide me, and make this dissertation possible. I would also like to express my appreciation to Prof. Xuan Zhao for reading my dissertation and providing helpful suggestions. I would like to thank Prof. Chris Fletcher for helping me understand climate areas and for supplying useful comments. I would like to appreciate Prof. Qi-Ming He for his support and mathematics training. I am also grateful to my external examiner, Prof. Jianxin Guo of the Chinese Academy of Science, for his insightful comments. During my academic journey, the entire faculty of the Management Science and Engineering department has fostered a welcoming atmosphere, offered invaluable support, and generously awarded scholarships, all of which I deeply appreciate. I would like to express my gratitude for the financial support from both the China Scholarship Council and the PhD program at the Department of Management Science and Engineering. I would also like to express my appreciation to Prof. Haiyan Xu, who made my PhD program available.

I would like to express my gratitude to my family. Thank you for the mental support from my parents, grandmother, brother Zhenghuan Wu, and cousins Qingchang Chen and Xiaoguo Pan. I would like to express my appreciation to my friends who have brought fun into my life and have helped me, including Prof. Ginger Ke, Chenyu Xi, Hao Fu, Haokun Chen, Jialin Xu, Jiawen Zheng, Mengqi Lu, Wei Liang, Xiaoyue Zhu, Xin Wang, Yu Han, and Zhichao Zhang. Thanks to Dana Hospitality at St. Jerome's University for providing healthy and nutritious food, which I like most of the time. Finally, I would like to express appreciation for myself, who enjoyed the PhD period.

Dedication

This dissertation is dedicated to my beloved parents Meizhu Pan and Shoucheng Wu.

Table of Contents

List of Figures

List of Tables

Chapter 1

Thesis Introduction

Adapting to climate change involves making informed decision sequences over a long-term horizon. This dissertation delves into these decision sequences for climate adaptation through a framework comprising climate data, regression models, and robust optimization. The following sections provide detailed introductions to climate adaptation and the framework.

1.1 Climate Change Adaptation

Climate change, as defined by the IPCC, encompasses any alteration in climate patterns over time, whether stemming from natural variability or resulting from human activities (IPCC, 2007).

Multiple scientific assessments have outlined the varied and substantial effects of climate

change on natural systems. Climate change can affect rainfall patterns, humidity, sea level, irregular seasons, floods, droughts, and storms (Rajkhowa and Sarma, 2021).

These climate changes bring about various challenges and opportunities for human society. On the challenging side, climate change has been linked to severe health issues (Kotcher et al., 2021), significant economic losses (Brás et al., 2021), sea level rise (Gregory et al., 2019), food and water insecurity (Fitton et al., 2019), social instability (Richards et al., 2021), ecosystem loss (Turner et al., 2020), among other issues. Conversely, it may also offer benefits in some respects. For example, lands near polar regions may experience increased productivity due to global warming (Su et al., 2021). Therefore, it is crucial to consider climate change in human operations and production to adapt to and exploit these challenges and opportunities (Santos et al., 2020; Parmesan et al., 2022).

Climate change adaptation, according to the IPCC, is defined as "*adjustment in natural or human systems in response to actual or expected climatic stimuli or their effects, which moderates harm or exploits beneficial opportunities*" (IPCC, 2007). Furthermore, this adaptation should encompass actions over more than a 50-year time horizon, which is necessary to account for many possible climate change effects, as noted in Hannah et al. (2002).

Supported by a wealth of research evidence (Parmesan et al., 2022; Grafakos et al., 2020), adapting to climate change is essential for a wide range of stakeholders, including investors and government bodies. To illustrate the importance of climate adaptation, we provide two examples: one in the context of land investment and the other in electricity storage operations.

In the field of land investment, changes in climate, such as fluctuations in precipitation and temperature levels throughout the year, can either enhance or diminish land productivity (Su et al., 2021). Land investors need to carefully consider appropriate strategies in their investment decisions under changing climate conditions to achieve favorable outcomes. These strategies include making proper investment decisions about when and where to buy and sell land given the investors' limited resources.

In the realm of electricity storage, climate change affects the variability of electricity supply, demand, and prices, presenting challenges for storage systems responsible for maintaining the stability and reliability of the electricity market (Chandramowli and Felder, 2014). Operators of these systems need to develop appropriate operational strategies to adapt to the effects of climate change on the electricity market. These strategies include making proper decisions about when and how much to charge and discharge storage systems given operators' limited resources.

The expense of climate change adaptation is substantial. The cost of adaptation bears significant financial implications, with projected costs rising to an estimated $160–$340 billion by 2030 and $315–$565 billion by 2050, as reported by the United Nations (UN Environmental Programme, 2022). Therefore, from a purely economic perspective, it is important to approach adaptation strategies with optimal efficiency.

The pursuit of such optimization closely aligns with the established purpose of Operations Research (OR). Within this field, individuals employ mathematical models, often referred to as optimization models, to systematically examine the optimal decisions or operational procedures. This analytical approach seeks to either maximize or minimize

specific objectives of interest while conscientiously considering the constraints imposed by the available resources.

The development of an optimization model needs to incorporate input parameters related to climate data to support climate change adaptation. These parameters are a function of climate, such as future prices derived from climate data. To obtain these parameters, we require climate data and the associated function. Given our specific interest in future prices, we refer to this function as a prediction model.

We provide details regarding climate data and prediction models as follows:

- *Climate data*: Prominent climate models, known as General Circulation Models (GCMs), offer climate projections for the next 80 years (Thrasher et al., 2022). GCMs are mathematical and computational models of climate systems used to generate projections based on future concentrations of greenhouse gases (GHG) and socio-economic characteristics (Kriegler et al., 2012). These climate projections can assist stakeholders understand the potential evolution of climate.

- *Prediction models*: The models predict parameters of interest based on climate data for OR models. The specific form of the prediction models depends on the area of interest. However, one popular form is regression models, which are statistical techniques that establish relationships between independent (also called explanatory) variables (e.g., climate variables) and dependent variables (e.g., prices in our applications).

Adaptation to climate change is difficult because of the uncertainty in climate projec-

tions. Successful adaptation relies on accurate climate change forecasts. Climate change forecasts are based on GCMs and future emission scenarios, but these forecasts are subject to uncertainties, including the internal variability of the climate system, response uncertainties of GCMs, and radiative forcing uncertainties (Hawkins and Sutton, 2009; Lehner et al., 2020). These uncertainties lead to an increase in the variance of climate change forecasts over time (Lehner et al., 2020), making GCMs a difficult tool to use if wanting accurate climate predictions.

Climate change adaptation should consider extreme climate scenarios that will become more frequent in the future if no action is taken. This is of paramount importance because the most significant threats to humans will manifest locally through changes in regional extreme weather and climate events (Beniston et al., 2007). For instance, North America has witnessed shifts in the occurrence and severity of extreme events, including floods, droughts, windstorms, heatwaves, and wildfires, in recent years. In North America, only extreme hydrological events, such as floods and droughts, have caused economic losses of about $80 billion between 2010-2014 (Gao et al., 2019). Therefore, it is important for OR models to formulate solutions designed for these extreme cases.

One promising approach to address the uncertainty of climate projections and extreme climate events is to undertake climate change adaptation under the worst-case scenario. The concept involves investigating optimal solutions that are effective across all scenarios introduced by potential climate projections. The underlying rationale is that if a solution proves successful in the worst-case scenario, it will still work in less severe scenarios. This approach has gained attention in previous policy papers, such as those by Hallegatte (2009) and Constantino and Weber (2021), as a promising strategy for long-term climate change

adaptation. In our applications, we generate the worst-case scenario for the next 80 years for each available climate projection.

The adaptation to the worst-case scenario aligns with the domain of robust optimization, an important area of research in the OR field. Robust optimization is a mathematical optimization methodology that represents uncertainty as a set and seeks to determine the optimal policy based on the worst-case scenario, as determined by the characteristics of the uncertainty set.

Furthermore, robust optimization exhibits favorable properties, which enhances its model-solving capabilities. These capabilities make robust optimization tractable over long-term horizons ($\geq$ 50 years), as required to study climate change adaptation (Hannah et al., 2002).

In this context, we integrate three components: climate data (derived from GCM projections), regression models, and robust optimization models to investigate climate change adaptation in the context of worst-case scenarios induced by climate change. In the following sections, we provide more details about these three components.

1.2 Climate Projections

Climate projections provide important climate data for climate change adaptation. Popular climate models, also known as GCMs, are used world-wide to project future climate variables, including temperature, precipitation, wind speed, and solar radiation. We rely on GCM projections to understand climate change and predict future economic parameters

of interest, e.g., prices in our applications.

GCMs partition the Earth's surface into a three-dimensional grid of cells and compute climate values for each cell. For instance, Figure 1.1 illustrates a grid division scheme commonly employed by GCMs. The size of these grid cells determines the model's resolution, with smaller cells offering greater resolution and detail but demanding more computing resources. Detailed explanations regarding how GCMs derive climate values for individual grid cells can be found in National Oceanic and Atmospheric Administration (2023).

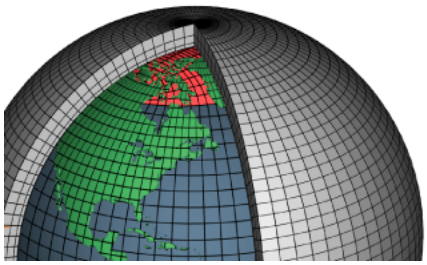


Figure 1.1: A three-dimensional grid of cells used by GCMs (retrieved from National Oceanic and Atmospheric Administration (2023))

GCMs are parameterized by social and emission scenarios, such as Representative Concentration Pathways (RCPs) (Pachauri et al., 2014; Adopted, IPCC, 2014) and Shared Socio-economic Pathways (SSPs) (Riahi et al., 2017).

RCP describes pathways of GHG emissions and atmospheric concentrations leading to different radiative forcing values (Van Vuuren et al., 2011). In contrast, SSP further refine the RCP framework by incorporating the influence of societal choices on radiative forcing, particularly toward the end of the century (Kriegler et al., 2012). In essence, SSP represents an extended and more contemporary iteration of RCP.

Within the RCP framework, two well-known scenarios are RCP4.5, signifying a moderate emission scenario, and RCP8.5, indicative of a high emission scenario (Van Vuuren et al., 2011). The numerical values 4.5 and 8.5 correspond to radiative forcing values in the year 2100, reaching 4.5 W/m^2 and 8.5 W/m^2, respectively. Notably, RCP8.5 represents a more intense warming scenario than RCP4.5.

Within the SSP framework, there are four popular SSPs: SSP1-2.6 (SSP126), SSP2-4.5 (SSP245), SSP3-7.0 (SSP370) and SSP5-8.5 (SSP585). The numerical values from 2.6 to 8.5 denote the (stratospheric-adjusted) radiative forcing in W/m^2 anticipated by the end of the 21st century (Tebaldi et al., 2020). The range from SSP126 to SSP585 spans a sequence reflecting progressively warmer future scenarios.

For each RCP or SSP, a GCM can generate numerical values of climate variables for each grid cell for the next 80 years or even up to 2300 (O'Neill et al., 2016; Taylor et al., 2012). The types of generated climate variables may vary across different GCMs and may include temperature, precipitation, wind speed, and more.

We use RCP in the land investment case because, at the time of studying this case, the climate dataset based on SSP was incomplete. In contrast, we use SSP in our electricity storage system case because, at the time of studying this case, the climate dataset based on SSP was complete.

Multiple GCMs, parameterized by multiple social and emission scenarios, such as RCP4.5 (a mild-emission scenario) and RCP8.5 (a high-emission scenario) (Adopted, IPCC, 2014), generate various climate projections. These climate projections have been collected by the Coupled Model Intercomparison Project (CMIP) (Taylor et al., 2012), administered

by the Intergovernmental Panel on Climate Change (IPCC). Note that there are several CMIPs including the Coupled Model Intercomparison Project Phase 5 (CMIP5) (2010-2014) (Taylor et al., 2012) and the most recent one, the Coupled Model Intercomparison Project Phase 6 (CMIP) began in 2013 (Eyring et al., 2016). We refer to the combination of GCMs and social and emission scenarios as climate scenarios. To gain a comprehensive understanding of the entire spectrum of potential climate impacts on human systems and to study adaptation in worst-case scenarios, we utilize a wide range of climate scenarios.

As the climate variables generated by GCMs may not possess the required temporal or spatial resolution for specific research, some projects offer higher-resolution climate data based on the climate projections from GCMs. These projects include the high-resolution Localized Constructed Analogs (LOCA) CMIP5 Projections (Pierce et al., 2014) and the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6) (Thrasher et al., 2022). We have utilized these high-resolution projections to meet the demands of our research. However, in our electricity storage system case, which necessitates hourly resolutions, we further temporally downscale the data from these projects (e.g., NEX-GDDP-CMIP6).

1.3 Regression Models

Regression models serve as the bridge between climate data and robust optimization by transforming climate values into the parameters required by robust optimization models. The specific type of regression model employed depends on the research questions being addressed. In this dissertation, we utilize linear regression models (Devore, 2011), the most

popular type of regression model, to predict economic parameters, such as land prices and electricity prices, for the next 70-80 years. We will provide the general form of linear regression models here and details about our linear regression models in Chapters 2 and 3.

A linear regression model is a mathematical framework used to study the relationship between output, also referred to as a dependent variable, and the independent variables that influence the dependent variable. Let Y be the dependent variable, and Y_i be the ith (random) observation of Y, where $i = 1, 2, ..., n$ and n is a positive integer. A linear regression model assumes that the mean value of Y_i $(i = 1, 2, ..., n)$ relies on a linear combination of a set of m independent variables $x_{1,i}$ to $x_{m,i}$, associated with a collection of regression parameters β_j (for $j = 0, 1, 2, ..., m$, where m is a positive number). Mathematically, the linear regression model has the following form:

$$\mathbb{E}[Y_i] = \beta_0 + \beta_1 x_{1,i} + ... + \beta_m x_{m,i} \tag{1.1}$$

, where $\mathbb{E}[\cdot]$ denotes the mathematical expectation.

To account for randomness, Model (1.1) assumes that the dependent variable Y is sampled from a specific family of probability distributions, which may include distributions such as the normal, Poisson, or binomial distribution. The normal distribution assumption for Y is a common choice. Even if Y often does not follow a normal distribution in real-world cases, we can transform it to approximate a normal distribution to facilitate statistical tests and estimation. For example, if Y follows a distribution with a long right tail, also known as a right-skewed distribution, we can take the logarithm transform of Y, denoted as $log(Y)$, to make it closely approximate a normal distribution. Such a right-

skewed distribution is often observed in land price data (Schlenker et al., 2006) because a small part of land prices is much higher than the mean of land prices. Additionally, if the volume of observation is large, such as a significant amount of historical land price data in our case, the strict requirement of a normal assumption is not always necessary (Dunn et al., 2018), as the distribution of the data after standardization may closely approximate normality.

The method to solve linear regression models involves estimating the parameters $\beta_0, \beta_1, ..., \beta_m$. Typically, the ordinary least squares (OLS) algorithm is applied to solve linear regression models. Let n be the number of observations, and y_i be the observed sample (data) of Y_i (for $i = 1, 2, ..., n$). The general form of OLS is expressed as follows:

$$\min_{\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_m} \quad \sum_{i=1}^{n} [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + ... + \hat{\beta}_m x_{m,i})]^2 \tag{1.2}$$

, where min denotes *minimize*, and $\hat{\beta}_0, ..., \hat{\beta}_m$ are the estimators of parameters $\beta_0, ..., \beta_m$.

Model (1.2) offers analytical solutions and is solvable with readily available computer software, such as R (R Core Team, 2020) and Python (Van Rossum and Drake, 2009).

The linear regression model assumes that the linear relationship between independent variables and the dependent variable holds when predicting the dependent variable for given independent variables (Dunn et al., 2018). For specific values of the independent variables $x_{1,k}, ..., x_{m,k}$, where k is a positive integer, we can calculate the mean value $\mathbb{E}[Y_k] = \hat{\beta}_0 + \hat{\beta}_1 x_{1,k} + ... + \hat{\beta}_m x_{1,m}$.

In our case, we utilize historical climate data (as independent variables) and historical

price data (as a dependent variable) to solve our linear regression models and obtain the linear relationship between climate variables and prices. Following the previous assumption that the linear relationship holds for predicting, we assume that the linear relationship between climate variables and prices holds for the next 80 years. Climate projection data can be used to predict future prices in the next 80 years using the linear relationship.

Due to observation errors, the estimation of the mean value $\mathbb{E}[Y_k]$, where k is a positive integer, by the OLS algorithm is subject to uncertainty. It is more appropriate to provide a range prediction rather than a single point estimate for $\mathbb{E}[Y_k]$ given the explanatory variables $x_{1,k}, ..., x_{m,k}$ (Dunn et al., 2018).

In the field of statistics, a confidence interval (Dunn et al., 2018) is a statistical tool that computes a range of values from observations or sample data. These intervals are calculated at specified confidence levels, providing a range within which we can reasonably expect the true population parameter to lie with a certain level of confidence. The 95% confidence level is the most common choice. The 95% confidence interval is constructed using sample data in a manner that provides a high likelihood of containing the true mean, e.g., $\mathbb{E}[Y_k]$ in our case, with a probability of approximately 95% when this process is repeated across numerous samples drawn from the same population.

With the given confidence level and values of the explanatory variables $x_{1,k}, ..., x_{m,k}$, where k is a positive number, we can estimate the potential confidence interval of the mean value $\mathbb{E}[Y_k]$. Detailed mathematical procedures for calculating confidence intervals are well-established, and these intervals can be easily computed using computer software, such as R and Python. For further information on solving, validating, and constructing

confidence intervals for linear regression models, interested readers can refer to Dunn et al. (2018).

Considering the inherent uncertainty of climate change, it becomes important to treat the projected economic parameters that are influenced by climate as uncertain. Within the framework of a regression model, we can naturally represent these projected economic parameters as confidence intervals. These intervals then serve as inputs for robust optimization, enabling us to explore adaptation strategies within the context of worst-case scenarios.

1.4 Robust Optimization

Robust optimization serves as the final, yet crucial, step within our research framework. It is the process through which uncertain parameter inputs are mapped into optimal decisions or operations designed to adapt to climate change. In the following paragraphs, we will delve into the specifics of an optimization model and its robust counterpart (i.e., robust optimization).

Let $\boldsymbol{x}$ represent a vector of decision variables, and $\boldsymbol{u}$ represent a vector of parameters that influence the outcomes of our decisions, which may be subject to uncertainty. In our case, $\boldsymbol{u}$ represents an uncertain price trajectory. We define $h(\cdot, \cdot)$ as the objective function, without of loss generality, to be maximized, and $g_j(\cdot, \cdot)$ as the function capturing constraints that may be influenced by $\boldsymbol{u}$, where $j = 1, 2, ..., J$ and J is a positive integer.

Then, we introduce the optimization model as follows:

$$\text{(Nominal problem)} \quad \max_{\boldsymbol{x}} \quad h(\boldsymbol{x}, \boldsymbol{u}) \tag{1.3a}$$

$$\text{s.t.} \quad g_j(\boldsymbol{x}, \boldsymbol{u}) \leq 0 \quad \forall j = 1, ..., J \tag{1.3b}$$

, where *max* represents *maximize* and *s.t.* denotes *subject to*.

The robust optimization framework aims to identify solutions that remain robust against any possible realization of $\boldsymbol{u}$ within a defined uncertainty set $\mathcal{U}$. This is achieved by solving the following robust counterpart:

$$\text{(Robust counterpart)} \quad \max_{\boldsymbol{x}} \quad \min_{\boldsymbol{u} \in \mathcal{U}} h(\boldsymbol{x}, \boldsymbol{u}) \tag{1.4a}$$

$$\text{s.t.} \quad g_j(\boldsymbol{x}, \boldsymbol{u}) \leq 0 \quad \forall \boldsymbol{u} \in \mathcal{U}, \forall j = 1, ..., J \tag{1.4b}$$

Constraints (1.4b) serves the purpose of pinpointing implementable solutions, regardless of the specific realization of $\boldsymbol{u}$. The objective function (1.4a) is employed to identify solutions that maximize the worst-case outcome of the objective function.

To solve Model (1.4), we need to define the uncertainty set $\mathcal{U}$. There are three popular ways to model uncertainty set $\mathcal{U}$ in order to make Model (1.4) tractable. First approach is to represent Constraints (1.4b) as chance constraints (Shapiro et al., 2021). In this context, let $\boldsymbol{u}$ be a random vector, and $\mathcal{P}(\cdot)$ be a probability measure. Additionally, let τ denote the acceptable probability level (also called a tolerant level), where $\tau \in [0, 1]$. The general

form of a chance constraint is as follows:

$$\text{(Chance constraint)} \qquad \mathcal{P}(g_j(\boldsymbol{x}, \boldsymbol{u}) \leq 0) \geq \tau, j \in [1, 2, ..., J] \tag{1.5}$$

Chance constraints have been widely used in various contexts (Shapiro et al., 2021); however, they may not align with our specific requirements in this case. The challenge arises from the difficulty in determining the distribution of uncertain parameters in our climate-related scenarios. GCMs are dependent, and this dependency is impossible to determine (Hallegatte et al., 2012). We cannot model each projection trajectory from GCM independently or ascertain their levels of dependency, which are required to calculate the distribution.

The second approach involves using the distributionally robust optimization (DRO) framework. In this method, $\boldsymbol{u}$ is still modeled as a random vector. However, instead of modeling the uncertainty set $\mathcal{U}$ as a specific distribution, the framework considers it as a family encompassing all potential distributions for $\boldsymbol{u}$, denoted as $\mathcal{D}$. The Constraints (1.4b) must be satisfied under the worst distribution of $\boldsymbol{u}$. The DRO counterpart for Model (1.4) can be expressed as follows:

$$\text{(DRO counterpart)} \quad \max_{\boldsymbol{x}} \quad \min_{F \in \mathcal{D}} \mathbb{E}_{\boldsymbol{u} \sim F}[h(\boldsymbol{x}, \boldsymbol{u})] \tag{1.6a}$$

$$\text{s.t.} \quad \mathbb{E}_{\boldsymbol{u} \sim F}[g_j(\boldsymbol{x}, \boldsymbol{u})] \leq 0 \quad \forall F \in \mathcal{D}, \forall j = 1, ..., J \tag{1.6b}$$

, where $\mathbb{E}_{\boldsymbol{u} \sim F}[\cdot]$ denotes the expectation of $\boldsymbol{u}$ over distribution F ($F \in \mathcal{D}$).

The DRO counterpart may encounter problem-solving issues. Often, the duality trans-

formation is applied to solve the DRO model, but the dual problem of the DRO model may not be convex and can be challenging to solve (Delage, 2021).

The third approach to represent the uncertain set $\mathcal{U}$ is as a box uncertainty set with lower and upper bounds. Let vector $\underline{\mathbf{u}}$ and vector $\bar{\mathbf{u}}$ be the lower and upper bounds of $\boldsymbol{u}$, respectively. Note that each entry of vector $\underline{\mathbf{u}}$ (respectively $\bar{\mathbf{u}}$) is smaller (respectively bigger) than corresponding entry of $\boldsymbol{u}$. We can express the box uncertainty set for $\mathcal{U}$ as $[\underline{\mathbf{u}}, \bar{\mathbf{u}}]$. Consequently, Model (1.4) can be reformulated as follows:

$$\text{(Box-uncertainty-set counterpart)} \quad \max_{\boldsymbol{x}} \quad \min_{\boldsymbol{u}\in[\underline{\mathbf{u}},\bar{\mathbf{u}}]} h(\boldsymbol{x},\boldsymbol{u}) \tag{1.7a}$$

$$\text{s.t.} \quad g_j(\boldsymbol{x},\boldsymbol{u}) \leq 0 \quad \forall \boldsymbol{u} \in [\underline{\mathbf{u}}, \bar{\mathbf{u}}], \forall j = 1, ..., J \tag{1.7b}$$

The box uncertainty set $[\underline{\boldsymbol{u}}, \bar{\boldsymbol{u}}]$ can be derived from the confidence intervals around the predicted mean values that result from regression models. The mean values, such as a mean price trajectory derived from the regression models, are dependent. This dependency arises because the means rely on dependent climate trajectories, which are the inputs to the regression models. Consequently, these dependent means generate corresponding dependent (confidence) intervals.

In our formulation of these intervals, we assume that the realization of uncertainty, such as price in our case, in one interval is independent of the price in other dependent intervals. This assumption allows each component of the vector $\boldsymbol{u}$ to vary across the intervals $[\underline{\boldsymbol{u}}, \bar{\boldsymbol{u}}]$. Each component seeks values to minimize the objective function $h(\boldsymbol{x}, \boldsymbol{u})$, effectively generating a worst-case scenario. As decision variables $\boldsymbol{x}$ seek values to maximize the objective function against the worst-case scenario, Model (1.7) can generate adaptation

decisions to address the scenario.

The robust model based on the box uncertainty set may exhibit linear equivalences that simplify the solving process (Bertsimas and Sim, 2003). In both of our specific applications in land investment and electricity storage systems, Model (1.7) demonstrates linear equivalences, rendering it tractable for long-term time horizon analyses. Therefore, we use the box uncertainty set to model uncertainty parameters for our robust models.

1.5 Conclusion

In summary, this introductory chapter has presented the framework that combines climate data, linear regression models, and robust models to study climate change adaptation. Figure 1.2 illustrates the framework's key components and workflow. First, we use historical data of interest, such as price data in our two applications, and historical weather data to solve our regression model. Second, we temporally or spatially downscale climate projections to achieve the desired resolutions. Third, the downscaled data, along with economic indicators, are input into the best-fitted regression model to predict future price interval trajectories. Fourth, we define the objective functions and constraints for our robust model according to a specific research question. Fifth, we solve the robust model, parameterized by the interval trajectories, to generate valuable insights into climate change adaptation.

Our framework differs from similar work of Garcia-Gonzalo et al. (2016), who study expected outcomes of forest harvest planning under climate change using a linear optimization model and a forest growth model. Our framework aims to study climate change

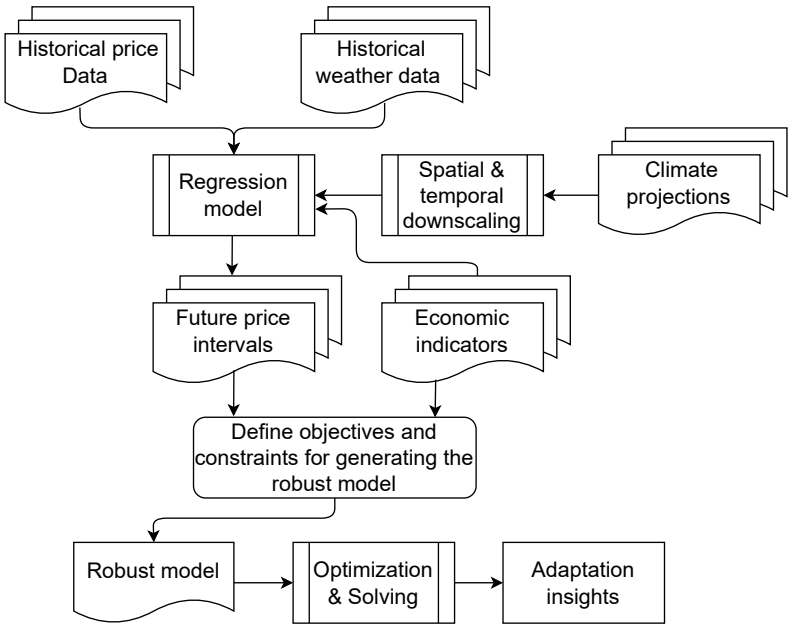


Figure 1.2: Framework and workflow of dissertation methodology

adaptation over worst-case scenarios. Our framework constructs both price prediction models and robust models, with model settings that are entirely different from model settings in Garcia-Gonzalo et al. (2016). We apply our framework to the domains of land investment and electricity markets, which are distinct from the field of forest planning. Additionally, we analyze how different climate scenarios influence optimal decisions, an aspect not mentioned in Garcia-Gonzalo et al. (2016).

To demonstrate the success of the entire framework, we applied it to the study of climate change adaptation in both land investment and electricity storage systems in Chapter 2 and Chapter 3, respectively.

Chapter 2

Robust Optimization for Sequential Investment Problems: An Application to Climate Adaptation in the Mississippi River Basin

This chapter uses our framework and performs case studies assessing how an investor would approach a long sequence of land investment decisions in the Mississippi River Basin (MRB) using popular climate models.

2.1 Introduction

Agriculture contributed about 4% of the global gross domestic product (GDP) in 2018 (The World Bank, 2022) and 0.6% ($134.7 billion) of the U.S. GDP in 2020 (U.S. Department of Agriculture, 2022). Combining with related industries, agriculture accounted for $1.055 trillion in 2020, i.e., 5% of the U.S. GDP (Moraes, 2022). The Mississippi River Basin (MRB) accounted for more than 25% of the total U.S. agriculture GDP in 2020 (Moraes, 2022).

Agricultural production is dependent on climate. Over the last 60 years, climate change has slowed global and U.S. agricultural productivity growth by 21% and 10-15%, respectively (Ortiz-Bobea et al., 2021). The effects of climate change on agricultural output vary by region and while overall the impact has been negative, there are many regions where agricultural productivity has increased (Ortiz-Bobea et al., 2021).

Agricultural land value is heavily determined by its productivity (Mendelsohn et al., 1994; Etwire et al., 2019), so, the land value is expected to change regionally as the climate evolves. The agricultural industry's ability to adapt to climate change will be affected by its ability to predict land value to make efficient and profitable investments.

By changing the levels and distribution of precipitation and temperature over the year, changes in climate may improve or deteriorate land productivity (Su et al., 2021; Ortiz-Bobea et al., 2021; Zhang and Swaminathan, 2020). Climate forecasts (generated by from GCMs), as a result, can be used to predict land productivity and, in turn, land value. In aggregate and on a global level, GCMs predict an increase in global mean temperatures and precipitation variation over the next 70 years (Collins et al., 2013; Adopted, IPCC,

2014).

Climate change will impact land productivity and values and present opportunities and challenges for all stakeholders in agriculture, including investors and producers, governments and policymakers, and the general public who consume the outputs (New frontiers, 2021; Atasu et al., 2020). The costs of climate change in this sector will depend both on the magnitude of these changes and the degree to which investment can be efficiently directed. This chapter focuses on agricultural land investment decisions that adapt to climate change.

Adapting agricultural investment strategies to cope with climate change is difficult for two main reasons. First, climate change projections are uncertain and their variance increases over time (Hawkins and Sutton, 2009; Lehner et al., 2020). Secondly, there is uncertainty about how land value will respond to particular changes in climate since it is difficult to predict factors in the regional economy such as consumer tastes. Consequently, what are the right farmland investment decisions to adapt to uncertain climate change remains unclear, especially considering climate-change's long timeline.

To study investment decisions under climate change, we need to deal with two main problems. The first is to determine farmland value trajectories as a function of climate change projections. The second is to determine investment decisions consistent with investor value and risk preferences over these land value trajectories. These methodological contributions will allow for results linking climate forecasts to trajectories of investment decisions. Understanding the variation of these decisions under different climate scenarios will provide insights into the difficulty of adapting to climate change.

Prior research in this vein has been done linking climate change to land values. In such studies, farmland values are estimated in regression models as price ranges across GCMs and RCPs (Mendelsohn et al., 1994; Deschênes and Greenstone, 2007; Van Passel et al., 2017; Bareille and Chakir, 2023). A separate stream of literature examines land acquisition under uncertain land valuations. They rely on stochastic optimization models to investigate the land acquisition policies over at most 30-year horizons (Schnitkey et al., 1989; Lohano and King, 2009; Spiegel et al., 2020). To the best of our knowledge, no paper examines optimal land acquisition under the specific patterns of valuations and uncertainties associated with climate change. No paper study farmland investment under climate change over a long horizon ($\geq$ 50 years), which as noted in Hannah et al. (2002) is required to assess climate change impacts and adaptations.

As the distribution of climate and price uncertainty is very hard to characterize, stochastic optimization approaches, which require a full characterization of the distribution, are challenging to implement. These algorithms also are typically intractable over the required very long trajectories (Shapiro et al., 2021). In our case, we employ a robust optimization model to study adaptation and investment under worst scenarios over the required horizons.

An important objection to robust optimization is that the resulting policies can be overly pessimistic and often sacrifice too much optimality (Ning and You, 2019). So, we adapt the methods of conservative robust optimization that allows fine tuning of the degree of pessimism (or conservatism levels) while maintaining tractability (Bertsimas and Sim, 2003).

Considering the issues in the study of farmland investment decisions and the merit of robust optimization approach both under climate change, we propose a new robust land investment model to investigate the decisions. We solve this model by converting it into its linear equivalences, which can be quickly solved by standard optimization tools.

Using our proposed land investment model, we conduct an empirical study in the MRB to determine general investment trends over time under climate change in two main steps. First, we fit a regression model and use it to predict the uncertain future farmland value trajectories given 64 climate scenarios (combinations of 32 GCMs and 2 RCPs). These trajectories are reformulated to be the farmland price parameters for the land investment model. Second, based on the price parameters, we solve our land investment model for the 64 climate scenarios. We study the resulting land investment policies in the MRB under different climate scenarios.

The specific contributions of this chapter are as follows:

1. We develop an optimization model for determining periodic investment decisions for uncertain assets over long horizons. Our model considers different conservatism levels through a robust optimization framework. We prove that the model has linear equivalences and show that it can be quickly solved by standard optimization tools.

2. We develop a regression model to predict land price changes based on long-term climate forecasts. Using the data derived from this method, we perform a computational study of investment decisions in the MRB over a 68-year planning horizon for a wide range of forecasted climate scenarios and conservatism levels. The main results include:

(a) Benchmark scenario (Section 2.5.2): Our benchmark scenario considers a investor using the land values based on the mean of all climate models. Under this mean climate, investing in the northernmost area and holding land are the optimal investment policies.

(b) Under different climate scenarios and conservatism levels (Section 2.5.2): We find that the range of latitudes over which investment takes place decreases and varies over climate scenarios and conservatism levels. For high conservatism levels, under RCP4.5 (a mild-range emission scenario), most GCMs find profitable investments, but there are substantial differences in investment patterns depending on the GCM. Under RCP8.5 (a higher-range emission scenario), most GCMs result in no investment.

This chapter is organized as follows. In Section 2.2, we review the literature relevant to 1) the economic impacts of climate change on farmland values and 2) land investment in the face of uncertainty. We then briefly describe our methodology in Section 2.3. In Section 2.4, we introduce a land price prediction model and conduct numerical studies on farmland price prediction in the MRB under climate change scenarios. In Section 2.5, we conduct empirical studies on land investment decisions in the MRB under climate change scenarios. Finally, Section 2.6 concludes this chapter.

2.2 Literature Review

In this section, we discuss two interrelated streams of literature. First, we discuss studies of the economic impacts of climate on land values. Second, we outline the literature studying land investment in the face of uncertainty and position our contributions.

2.2.1 Economic Impacts of Climate Change on Farmland Values

Regression approaches dominate the literature studying the economic impacts of climate change on land values (measured in prices). This approach is a hedonic model used to study the statistical relationship between climate variables and farmland values (Mendelsohn et al., 1994; Bareille and Chakir, 2023). Temperature, precipitation, and other exogenous factors are chosen as explanatory variables for farmland values. The approach has a simple structure and accounts for climate adaptations by farmers, while not needing to model endogenous decision making when measuring the long-term impacts of climate change (Bozzola et al., 2018; Bareille and Chakir, 2023). A range of papers has used this model to study how climate change impacts farmland values in North America (Quaye et al., 2018; Ortiz-Bobea, 2020), Europe (Van Passel et al., 2017; Bareille and Chakir, 2023), and other regions (Chen et al., 2013; Nguyen and Scrimgeour, 2022).

Researchers have used the results of this model to calculate the marginal effects of climate change on land value using forecasted changes in climate variables (Deschênes and Greenstone, 2007; Van Passel et al., 2017; Ortiz-Bobea, 2020; Nguyen and Scrimgeour, 2022). These results are frequently used to estimate regional economic impacts of climate

change.

Until recently most researchers have predicted the range of economic impacts across only one or a small number of climate scenarios. As the number of influential climate models has increased, it's become important to study the variation in land values predicted by different models. The recent paper by Bozzola et al. (2018) studies farmland values in Italy over 16 climate scenarios to explore a broader set of possible climate pathways. Their results imply that using a wide range of climate scenarios helps to understand the range of climate impacts on land values. As a basis for our study of investment policies, we also use a regression-based approach to establish land values and do so over all available 64 high-resolution climate projections (Bracken, 2016). These climate projections are based on 32 popular GCMs and two RCPs used by the CMIP5 (Taylor et al., 2012). While the regression-based approach is not the primary contribution of this chapter, we find comparable results with Mendelsohn et al. (1994) (i.e., farmland values will grow in the North of the MRB and decrease in the South under climate change). By studying land investment over this large set of climate models, we find that, despite similar average results, different climate models lead to quite different land investment decisions.

2.2.2 Land Investment in the Face of Uncertainty

Some literature studies optimal agricultural land investment under uncertainty. Schnitkey et al. (1989) use stochastic programming to determine the optimal decision to purchase or sell farmland for a central Illinois farm in which annual farmland values were linked to stochastic farmland returns. Lohano and King (2009) propose a multi-period farmland

investment portfolio model by considering risk levels in farmland and debt financing for Southwestern Minnesota farms. Di Corato and Zormpas (2022) apply the real options method to examine optimal execution of qualitative farming operations, encompassing active and passive farming strategies, in southern Sweden, while we focus on the dynamics of land acquisition, including quantitative buying and selling decisions. These three methods are applied only to planning horizons of at most 30 years and are unlikely to extend to the broader geographies and longer horizons required to study climate adaptation, while our research considers climate change for a 68-year planning horizon and in the main part of MRB area: Minnesota (MN), Wisconsin (WI), Iowa (IA), Illinois (IL), Missouri (MO), Arkansas (AR), Mississippi (MS), Louisiana (LA), and small parts of other states (see the left subfigure of Figure 2.5).

The relationship between climate and investment policy has also been discussed in the political economy literature. While they do not provide a formal model, Fairbairn et al. (2021) reason that climate uncertainty increases investment risk and impacts farmland investment. They discuss potential arbitrage opportunities which may result from environmental/climate uncertainty.

Other related research is from the forest land management literature. This literature does not consider land acquisition decisions. Rather, these papers consider when and where to harvest lumber. Nonetheless, the research requires similar climate modeling and optimization frameworks. Veliz et al. (2015) and Garcia-Gonzalo et al. (2016) use multistage stochastic programming to analyze optimal harvest planning for the forest industry under environmental and climate uncertainty. Veliz et al. (2015) study the planning of 17 forests in 5 years over about 300 environment scenarios. Their mixed-integer programming model

is difficult to extend to 50 years or longer planning horizon and larger forests because of the high computation requirements of solving the model. Garcia-Gonzalo et al. (2016) study the planning of 24 forest strata in 15 years over 32 climate scenarios by a linear stochastic model which is equivalent to a robust optimization model. They show that for this decision problem, their proposed algorithm may be quickly solved. However, the solutions are overly conservative. In our more complex investment problem, we similarly show that the robust optimization scales well to long-term climate decision problems. In contrast to their algorithm, we introduce the ability to adjust the level of conservatism in the solution. We can explore the characteristics and profitability of land investment decisions for various levels of conservatism.

2.3 Robust Optimization for Sequential Investment

In this section, we introduce a robust land investment model to study farmland investment decisions under climate change. First, we propose a multistage stochastic programming model of the investment problem. Second, we develop a robust version of this model. In service of improving tractability, we show that the robust model has linear equivalences.

2.3.1 Stochastic Programming for Sequential Investment

We begin by constructing a multistage stochastic programming model for an investor to determine farmland acquisition decisions over a long-term planning horizon. The model studies annual buying and selling decisions for land in a set of counties where farmland

prices may vary idiosyncratically over the planning horizon. In each year, the investor has a cash endowment, from land sales and external sources and may decide the number of acres of land to buy and sell in each county. The investor can hold surplus cash after investment for future periods. In addition, the investor needs to consider changes in land prices caused by climate change and variation in their decisions. These price changes vary by county and are uncertain due to the range of potential climate scenarios. The objective of the model is to identify annual land acquisition and sale decisions that maximize expected profits in the terminal period.

We now introduce the key components of our model: planning horizon, external income, county set, price trajectories, decision variables, transaction costs, investor's objective, and constraints.

Planning horizon: We consider a planning horizon of T periods and denote the collection of each planning period as $\mathcal{T} = \{1, 2, ..., T\}$. A farmland investor is at period $t = 0$ (current period), owns no farmland, and makes investment decisions starting at period $t = 1$ and ending at period $T - 1$.

External income: For period $t \in \{1, 2, ..., T-1\}$, the investor gains external cash income B_t.

County set: There are $m - 1$ counties over the focal region. We also denote risk-free assets held by the investor as an artificial county m. These assets allow the investor to transfer surplus cash between periods for future investment. Therefore, we denote the set of counties in the target area as $\mathcal{C} = \{1, 2..., m\}$. Each county $c \in \mathcal{C} \setminus \{m\}$, where $\setminus$ represents set difference and $\mathcal{C} \setminus \{m\} = \{x \in \mathcal{C} : x \notin \{m\}\}$, has available acres of farmland

equal to A_c which remains fixed over the planning horizon. The artificial county m which is treated similarly to land assets has $A_m = M$, where M is a large positive number, to allow holding an arbitrary level of cash.

Price trajectories: We define a positive random vector $\boldsymbol{P}_t = [P_{t1}, P_{t2}, ..., P_{tm}]^\top$ $(t \in \mathcal{T})$ to model land price in each county at period t. Note that $\top$ represents the transpose of the vector. Then, we can model the land price trajectory over the planning horizon in each county as a random process $\boldsymbol{P}_1, \boldsymbol{P}_2, ..., \boldsymbol{P}_T$. For the county m representing cash holdings, we set the price trajectory over the planning horizon $\mathcal{T}$ to reflect annual interest on cash holdings at the inflation rate β $(\beta > 0)$ so that $P_{tm} = (1 + \beta)^{t-1}$ $(\forall t \in \mathcal{T})$. For the remaining counties $c \in \mathcal{C} \setminus \{m\}$, the investor is allowed to realize the real land prices at period t $(\forall t \in \mathcal{T})$ at the beginning of that period, before making land buying and selling decisions. Denote the realization of land prices at period t $(t \in \mathcal{T})$ in each county as $\boldsymbol{p}_t = [p_{t1}, p_{t2}, ..., p_{tm}]^\top$, where p_{tc} is a positive real number and $p_{tm} = P_{tm}$.

Decision variables: The investor can either buy or sell land in each county at the beginning of each planning period before T. The decision variables in each period t $(t \in \mathcal{T} \setminus \{T\})$ are $\boldsymbol{b}_t = [b_{t1}, b_{t2}, ..., b_{tm}]^\top$ and $\boldsymbol{s}_t = [s_{t1}, s_{t2}, ..., s_{tm}]^\top$, which respectively represent the acres of land bought and sold in each county. We set $\boldsymbol{b}_0 = \boldsymbol{0}$ and $\boldsymbol{s}_0 = \boldsymbol{0}$, where $\boldsymbol{0}$ is the m-dimensional zero column vector. The land held by the investor in each county at the beginning of period t $(t \in \mathcal{T})$ is then $\sum_{k=0}^{t-1}(\boldsymbol{b}_k - \boldsymbol{s}_k)$. Finally, we assume that the investor's land acquisition decisions will not affect land prices in the target area and the artificial county.

Transaction cost: There is a transaction cost associated with each buying or selling

decision. We consider transaction costs for buying or selling decisions separately as a percentage of the total value of land that is bought or sold. The transaction cost rate on purchases of land and risk-free assets, i.e., the artificial county, are r_b^l $(0 < r_b^l < 1)$ and r_b^β $(0 \leq r_b^\beta < 1)$, respectively. The transaction cost rate on sales of land and risk-free assets are r_s^l $(0 < r_s^l < 1)$ and r_s^β $(0 \leq r_s^\beta < 1)$.

Investor's objective: The investor's objective is to maximize the expected value of land and risk-free assets at the last period T.

Constraints: The constraints of this model ensure that purchases and sales do not exceed available assets and resources. The constraints also indicate that the investor can either buy or sell farmland. First, total available cash following sales in each period must be sufficient to make all purchases in that period (cash flow constraints): $\sum_{c\in\mathcal{C}\setminus\{m\}}[(1 + r_b^l)p_{tc}b_{tc} - (1 - r_s^l)p_{tc}s_{tc}] + (1 + r_b^\beta)p_{tm}b_{tm} - (1 - r_s^\beta)p_{tm}s_{tm} \leq B_t$ $(\forall t \in \mathcal{T} \setminus \{T\})$. Second, net ownership of farmland in each county and period must not exceed available acres: $0 \leq \sum_{k=1}^{t}(b_{kc} - s_{kc}) \leq A_c$ $(\forall t \in \mathcal{T} \setminus \{T\},\ c \in \mathcal{C})$. Third, no land buying and selling can happen in the same period and the same county (except the artificial county): $b_{tc}s_{tc} = 0$ $(\forall t \in \mathcal{T} \setminus \{T\},\ c \in \mathcal{C} \setminus \{m\})$. Fourth, the transaction in each county and period are non-negative: $b_{tc}, s_{tc} \geq 0$ $(\forall t \in \mathcal{T} \setminus \{T\},\ c \in \mathcal{C})$.

The multistage programming model, foundational for the robust optimization, can be formulated as dynamic programming equations ((2.1a)-(2.1g)); these equations are called Model (2.1). At the last period T, land price $\boldsymbol{p}_T$ is observed, and the value of total owned

land and risk-free assets is:

$$Q_T(\sum_{k=0}^{T-1}(\boldsymbol{b}_k - \boldsymbol{s}_k), \boldsymbol{p}_T) := \big(\sum_{k=0}^{T-1}(\boldsymbol{b}_k - \boldsymbol{s}_k)\big)^\top \boldsymbol{p}_T. \tag{2.1a}$$

At period $t = 1, 2, .., T-1$, the problems are:

$$Q_t(\sum_{k=0}^{t-1}(\boldsymbol{b}_k - \boldsymbol{s}_k), \boldsymbol{p}_t) := max_{\boldsymbol{b}_t, \boldsymbol{s}_t} \mathbb{E}_{\boldsymbol{p}_{t+1} \sim \boldsymbol{P}_{t+1}}[Q_{t+1}(\sum_{k=0}^{t}(\boldsymbol{b}_k - \boldsymbol{s}_k), \boldsymbol{p}_{t+1})] \tag{2.1b}$$

$$s.t. \quad \sum_{c \in \mathcal{C} \setminus \{m\}} [(1 + r_b^l) p_{tc} b_{tc} - (1 - r_s^l) p_{tc} s_{tc}] + (1 + r_b^\beta) p_{tm} b_{tm} - (1 - r_s^\beta) p_{tm} s_{tm} \leq B_t \tag{2.1c}$$

$$0 \leq \sum_{k=1}^{t}(b_{kc} - s_{kc}) \leq A_c \quad \forall c \in \mathcal{C} \tag{2.1d}$$

$$b_{tc} s_{tc} = 0 \quad \forall c \in \mathcal{C} \setminus \{m\} \tag{2.1e}$$

$$b_{tc}, s_{tc} \geq 0 \quad \forall c \in \mathcal{C} \tag{2.1f}$$

The maximum expected value of land and risk free assets at period T can be expressed as:

$$\mathbb{E}_{\boldsymbol{p}_1 \sim \boldsymbol{P}_1}[Q_1(\boldsymbol{0}, \boldsymbol{p}_1)] \tag{2.1g}$$

, where $\mathbb{E}_{p_t \sim \boldsymbol{P}_t}[\cdot]$ $(t \in \mathcal{T})$ denotes the expectation over $\boldsymbol{P}_t$. The optimal value $Q_t(\sum_{k=0}^{t-1}(\boldsymbol{b}_k - \boldsymbol{s}_k), \boldsymbol{p}_t)$ $(\forall t \in \mathcal{T} \setminus \{T\})$ depends on $(\boldsymbol{b}_t - \boldsymbol{s}_t)$ in period t and the distribution of farmland prices $\boldsymbol{P}_{t+1}$.

The method of solving Model (2.1) is to calculate the value-to-go function $Q_t(\sum_{k=1}^{t-1}(\boldsymbol{b}_k -$

$\boldsymbol{s}_k), \boldsymbol{p}_t)$, recursively, going back in time (Shapiro et al., 2021). In the last step, we want to find $\mathbb{E}_{\boldsymbol{p}_1 \sim \boldsymbol{P}_1}[Q_1(\boldsymbol{0}, \boldsymbol{p}_1)]$.

However, solving Model (2.1) is difficult for two reasons. First, when calculating $\mathbb{E}_{\boldsymbol{p}_t \sim \boldsymbol{P}_t}[Q_t(\sum_{k=0}^{t-1}(\boldsymbol{b}_k - \boldsymbol{s}_k), \boldsymbol{p}_t)]$ $(\forall t \in \mathcal{T} \setminus \{T\})$, it is hard to obtain the accurate distribution of $\boldsymbol{P}_t$. The distribution may be not trivial and not standard, which makes solving the model difficult. Second, solving processes of Model (2.1) faces the curse of dimensionality (Shapiro et al., 2021). As Model (2.1) covers a large number of periods at the county level, the computation complexity and requirements for the model instance will render it infeasible for most machines. Therefore, we cannot directly and quickly solve the model.

To make Model (2.1) tractable, we then introduce a robust version of this model to solve. There are two reasons why the robust version works in our case. First, the robust version helps identify decisions that perform well across the range of uncertain future climate scenarios (Bhave et al., 2016). By employing these decisions obtained from the robust version, the investor can have confidence about the potential investment profit in their decisions and understand a broader range of risks that they will face (Hallegatte et al., 2012; Constantino and Weber, 2021). Second, the robust model has linear equivalences that can be directly and quickly solved by using standard optimization tools.

2.3.2 Robust Optimization for Sequential Investment

We transform Model (2.1) into its robust version in three steps. First, we need to model farmland price uncertainty. Second, based on the modeled uncertain prices, we propose a robust version of Model (2.1). In addition, to avoid the overly conservative solutions of the

robust model, we add a conservatism control to the robust model. Third, we transform the robust model with conservatism control into its linear equivalences, which can be solved directly and quickly.

Modeling farmland price uncertainty: Modeling uncertainty (or defining uncertainty sets) plays a critical role in constructing a robust model (Ning and You, 2019; Chen et al., 2023). By following Ben-Tal and Nemirovski (2000) and Bertsimas and Sim (2003), we model uncertainty as intervals (or box uncertainty sets). Specifically, in our case, we model each uncertain farmland price P_{tc} ($t \in \mathcal{T}$, $c \in \mathcal{C} \setminus \{m\}$) as a random variable that is positive, bounded, symmetric, and independent, and its values are restricted within an interval $[\bar{p}_{tc} - \hat{p}_{tc}, \bar{p}_{tc} + \hat{p}_{tc}]$, where $\bar{p}_{tc}$ is the nominal value of P_{tc}, $\hat{p}_{tc}$ is the largest possible deviation of P_{tc} from the nominal value, and $\bar{p}_{tc} - \hat{p}_{tc} > 0$.

Robust model: We formulate a robust investment model where the investor believes that farmland prices will fall within the intervals described in the preceding paragraph **Modeling farmland price uncertainty**. The investor is pessimistic and adversely selects farmland prices depending on whether they are buying (high prices within the interval) or selling (low prices within the interval) in each county. In the worst case, the farmland price is $\bar{p}_{Tc} - \hat{p}_{Tc}$ ($\forall c \in \mathcal{C} \setminus \{m\}$) in terminal period T when we account the value of all owned lands. The investor's objective can be expressed as:

$$\max_{\{b_{tc}, s_{tc}: t \in \mathcal{T} \setminus \{T\}, c \in \mathcal{C}\}} \sum_{c \in \mathcal{C} \setminus \{m\}} [(\bar{p}_{Tc} - \hat{p}_{Tc}) \sum_{t=1}^{T-1} (b_{tc} - s_{tc})] + p_{Tm} \sum_{t=1}^{T-1} (b_{tm} - s_{tm}).$$

The investor also wants to keep non-negative cash flow in the worst case in each period before period T. Then, for $t \in \mathcal{T} \setminus \{T\}$, the cash flow constraints can be expressed as:

$\sum_{c\in\mathcal{C}\setminus\{m\}}[(1+r_b^l)(\bar{p}_{tc}+\hat{p}_{tc})b_{tc}-(1-r_s^l)(\bar{p}_{tc}-\hat{p}_{tc})s_{tc}]+(1+r_b^\beta)p_{tm}b_{tm}-(1-r_s^\beta)p_{tm}s_{tm}\leq B_t.$

Under the pessimistic setting, Model (2.1) becomes a deterministic robust model and can be written as follows:

$$\max_{\{b_{tc},s_{tc}:t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C}\}} \quad \sum_{c\in\mathcal{C}\setminus\{m\}}[(\bar{p}_{Tc}-\hat{p}_{Tc})\sum_{t=1}^{T-1}(b_{tc}-s_{tc})]+p_{Tm}\sum_{t=1}^{T-1}(b_{tm}-s_{tm}) \tag{2.2a}$$

$$s.t. \quad \sum_{c\in\mathcal{C}\setminus\{m\}}[(1+r_b^l)(\bar{p}_{tc}+\hat{p}_{tc})b_{tc}-(1-r_s^l)(\bar{p}_{tc}-\hat{p}_{tc})s_{tc}]+$$

$$(1+r_b^\beta)p_{tm}b_{tm}-(1-r_s^\beta)p_{tm}s_{tm}\leq B_t \quad \forall t\in\mathcal{T}\setminus\{T\} \tag{2.2b}$$

$$0\leq\sum_{t=1}^{k}(b_{tc}-s_{tc})\leq A_c \quad \forall k\in\mathcal{T}\setminus\{T\},c\in\mathcal{C} \tag{2.2c}$$

$$b_{tc}s_{tc}=0 \quad \forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C}\setminus\{m\} \tag{2.2d}$$

$$b_{tc},s_{tc}\geq 0 \quad \forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C} \tag{2.2e}$$

Note that a solution that satisfies a model's constraints is called a feasible solution. When maximizing, an optimal solution is a feasible solution that yields the largest objective function value.

Robust model with conservatism control: Model (2.2) will produce (optimal) solutions that are too conservative in the sense that we lose too much optimality for ensuring robustness (Varas et al., 2014; Ning and You, 2019), because this model forces land prices in all counties to adopt extreme values that adversely influence the solution. For instance, we will see in the empirical investigation in Section 2.5.2, that the robust model can lead to solutions where no land investment decisions are made throughout the planning horizon,

i.e., $b_{tc} = s_{tc} = 0$ ($\forall t \in \mathcal{T} \setminus \{T\}$, $c \in \mathcal{C} \setminus \{m\}$).

However, in reality, it is unlikely for land prices in all counties to simultaneously reach their extreme values (Bertsimas and Sim, 2004). To produce a more realistic solution, a solution should allow some counties to have more favorable prices. To do this in a principled manner, we adapt the conservative robust optimization methods developed in Bertsimas and Sim (2004) to the farmland acquisition problem. The model in Bertsimas and Sim (2004) follows this intuition by allowing a limited set of prices to be more favorable for the investor. The number and degree of such price improvements in each period are limited by a parameter called the *budget of uncertainty*. The particular improvements are selected in an adversarial manner in keeping with robust methods.

To accomplish our objective and allow the model to be adjusted to reflect different levels of conservatism, we introduce per period budgets of uncertainty Γ_t ($\Gamma_t \geq 0, \forall t \in \mathcal{T}$), as defined by Bertsimas and Sim (2003). By choosing the parameter Γ_t ($t \in \mathcal{T}$), decision makers can assess the balance between robustness and the solution's efficacy based on their conservatism levels in each period (Varas et al., 2014). Specifically, we define the budgeted uncertainty set for land prices (Bertsimas and Sim, 2004; Ning and You, 2019) as follows:

$$\{(\boldsymbol{P}_1, \boldsymbol{P}_2, ..., \boldsymbol{P}_T) | P_{tc} = \bar{p}_{tc} + \hat{p}_{tc} \cdot z_{tc}, -1 \leq z_{tc} \leq 1, \sum_{c \in \mathcal{C} \setminus \{m\}} |z_{tc}| \leq \Gamma_t, \forall t \in \mathcal{T}, c \in \mathcal{C} \setminus \{m\}\} \tag{2.3}$$

, where P_{tc} is the c-th component of uncertain (price) vector $\boldsymbol{P}_t$, and z_{tc} denotes the extent and direction of price deviation. These constraints regulate the cumulative deviation of land prices from their expected value within each year to be within the uncertainty budget.

In the robust model with budgets of uncertainty, the objective function maximizes the worst case while ensuring that the budget constraints are satisfied at the given budget of uncertainty Γ_t ($t \in \mathcal{T}$). Moreover, we use $z_{tc} \in [0,1]$ ($\forall t \in \mathcal{T}$, $c \in \mathcal{C} \setminus \{m\}$) because for each county and period the investor either buys or sells at different buy and sell prices. Specifically, the investor buys at price $\bar{p}_{tc} + \hat{p}_{tc} z_{tc}$ and sells at price $\bar{p}_{tc} - \hat{p}_{tc} z_{tc}$. The land price is $\bar{p}_{Tc} - \hat{p}_{Tc} z_{Tc}$ in terminal period T when we calculate the value of all owned lands. Then, the objective function maximizing the worst case at the given budget of constraints can be expressed as:

$$\max_{\{b_{tc}, s_{tc} | \forall t \in \mathcal{T} \setminus T, c \in \mathcal{C}\}} \min_{\{z_{Tc} | c \in \mathcal{C} \setminus \{m\}\}} \Big\{ \sum_{c \in \mathcal{C} \setminus \{m\}} [(\bar{p}_{Tc} - z_{Tc} \hat{p}_{Tc}) \sum_{t \in \mathcal{T} \setminus \{T\}} (b_{tc} - s_{tc})] + p_{Tm} \sum_{t \in \mathcal{T} \setminus \{T\}} (b_{tm} - s_{tm}) \Big\}$$

For $t \in \mathcal{T} \setminus \{T\}$, the cash flow constraints can be expressed as:

$$\max_{\{z_{tc} | c \in \mathcal{C} \setminus m\}} \Big\{ \sum_{c \in \mathcal{C} \setminus \{m\}} [(1 + r_b^l)(\bar{p}_{tc} + z_{tc} \hat{p}_{tc}) b_{tc} - (1 - r_s^l)(\bar{p}_{tc} - z_{tc} \hat{p}_{tc}) s_{tc}] \Big\}$$

$$+ (1 + r_b^\beta) p_{tm} b_{tm} - (1 - r_s^\beta) p_{tm} s_{tm} \leq B_t$$

The robust counterpart of Model (2.1) with the budgets of uncertainty, denoted as Model

(2.4), is as follows:

$$\max_{\{b_{tc},s_{tc}|\forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C}\}} \quad \min_{\{z_{Tc}|c\in\mathcal{C}\setminus\{m\}\}} \Big\{ \sum_{c\in\mathcal{C}\setminus\{m\}} [(\bar{p}_{Tc} - z_{Tc}\hat{p}_{Tc}) \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tc} - s_{tc})] + p_{Tm} \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tm} - s_{tm}) \Big\} \tag{2.4a}$$

$$s.t. \quad \max_{\{z_{tc}|c\in\mathcal{C}\setminus m\}} \Big\{ \sum_{c\in\mathcal{C}\setminus\{m\}} [(1+r_b^l)(\bar{p}_{tc} + z_{tc}\hat{p}_{tc})b_{tc} - (1-r_s^l)(\bar{p}_{tc} - z_{tc}\hat{p}_{tc})s_{tc}] \Big\}$$

$$+ (1+r_b^\beta)p_{tm}b_{tm} - (1-r_s^\beta)p_{tm}s_{tm} \le B_t \quad \forall t \in \mathcal{T}\setminus\{T\} \tag{2.4b}$$

$$0 \le \sum_{t=1}^{k}(b_{tc} - s_{tc}) \le A_c \quad \forall c \in \mathcal{C}, k \in \mathcal{T}\setminus\{T\} \tag{2.4c}$$

$$\sum_{c\in C\setminus m} z_{tc} \le \Gamma_t \quad \forall t \in \mathcal{T} \tag{2.4d}$$

$$b_{tc}s_{tc} = 0 \quad \forall t \in \mathcal{T}\setminus\{T\}, c \in \mathcal{C}\setminus\{m\} \tag{2.4e}$$

$$b_{tc}, s_{tc} \ge 0 \quad \forall t \in \mathcal{T}\setminus\{T\}, c \in \mathcal{C} \tag{2.4f}$$

$$0 \le z_{tc} \le 1 \quad \forall t \in \mathcal{T}, c \in \mathcal{C}\setminus\{m\} \tag{2.4g}$$

Solving by linear transform: We solve Model (2.4) by deriving and then solving its linear equivalence. Because Model (2.4) has max-min operator in its objective function (2.4a), the maximal part in constraints (2.4b), and non-convex constraints (2.4e), it is a non-convex optimization problem (Note that readers can find the definitions of convex sets, convex functions, and convex optimization in Sections 2.1, 3.1, and 4.2, respectively, of Boyd et al. (2004)). This means that solving Model (2.4) using standard optimization tools, algorithms, or heuristics may result in getting stuck in a local optimum rather than

finding the (global) optimal solution (Note that readers can find the definitions of local optimum and global optimum in Section 4.1 of Walton (2022)). To obtain the optimal solution, we transform Model (2.4) into its linear equivalence.

There are two main steps to do the linear transform. First, without considering constraints (2.4e), we show that Model (2.4) has an equivalent linear programming model. We do this step because we can take advantage of strong duality (Bertsimas and Tsitsiklis, 1997) and the existing reference by Bertsimas and Sim (2003). Second, we show that the derived linear programming model is equivalent to Model (2.4), i.e., constraints (2.4e) are always satisfied with any optimal solution of the (derived) equivalent linear programming model. Consequently, we can solve Model (2.4) by solving the linear programming model, which can be quickly and directly solved.

First, without considering constraints (2.4e), we can rely on strong duality and Bertsimas and Sim (2003) to find that Model (2.4) has a linear equivalence as stated in the following theorem:

Theorem 1 *Model* (2.4) *without constraints* (2.4e) *has an equivalent linear programming formulation as follows:*

$$\max_{\{b_{tc},s_{tc}\ (\forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C});h_T,q_{Tc}\ (\forall c\in\mathcal{C}\setminus\{m\})\}}\Big\{\sum_{c\in\mathcal{C}\setminus\{m\}}\bar{p}_{Tc}\sum_{t\in\mathcal{T}\setminus\{T\}}(b_{tc}-s_{tc})$$

$$+p_{Tm}\sum_{t\in\mathcal{T}\setminus\{T\}}(b_{tm}-s_{tm})-\Gamma_T h_T-\sum_{c\in\mathcal{C}\setminus\{m\}}q_{Tc}\Big\} \tag{2.5a}$$

$$s.t.\quad \sum_{c\in\mathcal{C}\setminus\{m\}}[(1+r_b^l)\bar{p}_{tc}b_{tc}-(1-r_s^l)\bar{p}_{tc}s_{tc}]+(1+r_b^\beta)p_{tm}b_{tm}-(1-r_s^\beta)p_{tm}s_{tm}$$

$$+\Gamma_t h_t+\sum_{c\in\mathcal{C}\setminus\{m\}}q_{tc}\le B_t,\quad \forall t\in\mathcal{T}\setminus\{T\} \tag{2.5b}$$

$$h_t+q_{tc}\ge(1+r_b^l)\hat{p}_{tc}b_{tc}+(1-r_s^l)\hat{p}_{tc}s_{tc},\quad \forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C}\setminus\{m\} \tag{2.5c}$$

$$h_T+q_{Tc}\ge\hat{p}_{Tc}\sum_{t\in\mathcal{T}\setminus\{T\}}(b_{tc}-s_{tc}),\quad \forall c\in\mathcal{C}\setminus\{m\} \tag{2.5d}$$

$$0\le\sum_{t=1}^{k}(b_{tc}-s_{tc})\le A_c,\quad \forall c\in\mathcal{C},k\in\mathcal{T}\setminus\{T\} \tag{2.5e}$$

$$b_{tc},s_{tc}\ge 0,\quad \forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C} \tag{2.5f}$$

$$q_{tc}\ge 0,\quad \forall t\in\mathcal{T},c\in\mathcal{C}\setminus\{m\} \tag{2.5g}$$

$$h_t\ge 0,\quad t\in\mathcal{T} \tag{2.5h}$$

We present the proof of Theorem 1 in Appendix A.2.

We explain the newly introduced notations and the objective function in Model (2.5). For the new notations, h_t and q_{tc} ($t\in\mathcal{T},c\in\mathcal{C}\setminus\{m\}$) are the dual variables of z_{tc}. In the objective function (2.5a), the first item $\sum_{c\in\mathcal{C}\setminus\{m\}}\bar{p}_{Tc}\sum_{t\in\mathcal{T}\setminus\{T\}}(b_{tc}-s_{tc})$ is the value of owned lands in period T based on the nominal land prices $\bar{p}_{Tc}$ ($c\in\mathcal{C}$). The second item $p_{Tm}\sum_{t\in\mathcal{T}\setminus\{T\}}(b_{tm}-s_{tm})$ is the value of risk-free assets in the artificial county m. The third item $-\Gamma_T h_T-\sum_{c\in C\setminus m}q_{Tc}$ represents the cost of dealing with worst-case price deviation in

period T.

Second, we show that the linear Model (2.5) is equivalent to Model (2.4). To do this, we need to prove that any optimal solution of Model (2.5) ensures constraints (2.4e). We found that the optimal solution of Model (2.5) has the property in the following lemma:

Lemma 1 *Let* $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ *be the optimal solution of Model* (2.5)*, where* $\boldsymbol{b}^* = (\boldsymbol{b}_1^*, \boldsymbol{b}_2^*, ... , \boldsymbol{b}_{T-1}^*)$ *and* $\boldsymbol{s}^* = (\boldsymbol{s}_1^*, \boldsymbol{s}_2^*, ..., \boldsymbol{s}_{T-1}^*)$*, at least one of* b_{tc}^* *and* s_{tc}^* $(\forall t \in \mathcal{T} \setminus \{T\}, c \in \mathcal{C} \setminus \{m\})$ *is 0.*

Further, if $r_x^\beta + r_s^\beta > 0$*, i.e., the transaction cost of investing in risk-free assets exists, at least one of* b_{tm}^* *and* s_{tm}^* $(\forall t \in \mathcal{T} \setminus \{T\})$ *is 0.*

We present the proof of Lemma 1 in Appendix A.3.

Lemma 1 reflects no farmland buying and selling in the same period and in the same county in the optimal solution of Model (2.4). In other words, buying and selling land at the same county and in the same period leads to a reduction in profit, because 1) there are transaction fees for either buying or selling, and 2) buying prices are not lower than selling prices.

As the objective function (2.5a) maximizes its value, the optimal solution will only choose to buy or sell (or neither buy nor sell) land at one county in a period. In this scenario, the optimal solution of Model (2.5) automatically ensures constraints (2.4e). Consequently, we can solve non-convex robust Model (2.4) by solving linear robust programming Model (2.5).

2.4 Study I: Farmland Value in the MRB under Climate Change Scenarios

In this section, we study farmland value in the MRB under different climate scenarios. Specifically, we predict future farmland prices from 2023 to 2090 in counties whose centroid centers are located between the longitudes of 94.5° W and 89.0° W and between the longitudes of 29.4° N and 48.2° N. To facilitate understanding, we define the Northern MRB as the area with latitude greater than 39.3° N, which is the middle of the latitude range [29.4, 48.2]. The remaining region within the MRB is denoted as the Southern MRB. The scenario-based case study comprises three key steps. First, we introduce a regression model to link farmland values and climate change. Second, in order to fit the regression model, we introduce the data sources and data prepossessing. The climate scenarios forecast temperature and precipitation changes which vary by season and county. Third, we fit the regression model, forecast farmland values, and present farmland value results.

2.4.1 Regression Model

We rely on a regression approach to analyze the impacts of climate change on land values (Mendelsohn et al., 1994; Deschênes and Greenstone, 2007; Van Passel et al., 2017). Specifically, we introduce a log-linear quadratic four-season regression model at the county level to give the relation between seasonal temperature/precipitation and land values. This model is reasonable for four reasons. First, farmland values, in the U.S. as in other

countries, are highly skewed with a long right tail so that a logarithm model is preferable (Schlenker et al., 2006). Second, the relationship between climate and land values is non-linear and often assumed to be quadratic (Mendelsohn et al., 1994; Deschênes and Greenstone, 2007; Fezzi and Bateman, 2015; Van Passel et al., 2017; Bozzola et al., 2018). Third, a four-season model is appropriate for the U.S. because its climate allows different crops to grow throughout the year (Massetti et al., 2016). Finally, we rely on the county as a basic spatial level, because of the extensive county-level data available on farm inputs and outputs, and its popularity in the literature (Mendelsohn et al., 1994; Deschênes and Greenstone, 2007; Mendelsohn and Reinsborough, 2007; Quaye et al., 2018).

We do not include interaction terms of temperature and precipitation in the regression model, as the model without these terms produces satisfactory fitting results and avoids potential multicollinearity.

We introduce the settings and notations used in the regression model: seasonal temperature and precipitation (independent variables), dummy variables, farmland values (dependent variable), and coefficients.

Seasonal temperature and precipitation: We use standard meteorological seasons. We set that March, April, and May are spring; June, July, and August are summer; September, October, and November are fall; December, January, and February are winter. We index the four seasons from spring to winter with $i \in \mathcal{I} = \{1, 2, 3, 4\}$. Then, we denote the seasonal temperature (value) at county c ($c \in \mathcal{C} \setminus \{m\}$) in season i ($i \in \mathcal{I}$) in year t as $\tilde{T}_{c,i,t}$. Similarly, we let the seasonal precipitation (value) be $\tilde{R}_{c,i,t}$.

Dummy variables: In this study, we consider two types of dummy variables (also known

as indicators) to capture yearly differences in the dependent variable and absorb unobserved county-specific variances, respectively.

Let $\mathcal{Y} = \{y_1, y_2, y_3, ...\}$ be a collection of the finite years utilized for fitting the regression model, with $|\mathcal{Y}|$ representing the total number of years in this collection. We let $I_y(t)$ $(y \in \mathcal{Y} \setminus \{Y_1\}, t \in \mathcal{Y})$ represent year indicators employed to capture yearly differences, including those caused by technology changes, in the dependent variable. $I_y(t)$ takes a value of one if year t is equal to year y and zero otherwise. We exclude y_1 from the set $\mathcal{Y}$ because $|\mathcal{Y}| - 1$ year indicators are sufficient to identify all $|\mathcal{Y}|$ years.

The notation $I_{\bar{c}}(c)$ $(\bar{c} \in \mathcal{C} \setminus \{1, m\}, c \in \mathcal{C} \setminus \{m\})$ represents fixed county indicators used to absorb all unobserved county-specific variances arising from factors such as state agriculture policies and subsidies or other unobserved characteristics. $I_{\bar{c}}(c)$ takes a value of one when county c is the same as county $\bar{c}$ and zero otherwise. Note that the county set $\mathcal{C} = \{1, 2, ..., m\}$, where county m is the artificial county employed to represent risk-free assets. $m - 2$ $(= |\mathcal{C} \setminus \{1, m\}|)$ county indicators are sufficient to identify $m - 1$ (real) counties in the MRB.

Farmland values: We let $V_{c,t}$ $(c \in \mathcal{C} \setminus \{m\})$ be the discounted farmland value (price) (\$/acre) at county c in year t and $lnV_{c,t}$ be the logarithm of $V_{c,t}$.

Coefficients: We let $\beta_{\tilde{T},i}$ and $\beta_{\tilde{R},i}$ $(i \in \mathcal{I})$ be the coefficients of $\tilde{T}_{c,i,t}$ and $\tilde{R}_{c,i,t}$, respectively. Similarly, we let $\gamma_{\tilde{T},i}$ and $\gamma_{\tilde{R},i}$ be the coefficients of squared seasonal temperature $\tilde{T}^2_{c,i,t}$ and squared seasonal precipitation $\tilde{R}^2_{c,i,t}$. Let β_y $(y \in \mathcal{Y} \setminus \{Y_1\})$ be the coefficient of $I_y(t)$ and $\beta_{\bar{c}}$ $(\bar{c} \in \mathcal{C} \setminus \{1, m\})$ be the coefficient of $I_{\bar{c}}(c)$.

In the regression model, we let a be a constant. $\mu_{c,t}$ $(c \in \mathcal{C} \setminus \{m\})$ is a stochastic

error term that is assumed not to be correlated with climate. Specifically, we propose the regression model as follows:

$$\begin{aligned} lnV_{c,t} = a + \sum_{i=1}^{4}(\beta_{\tilde{T},i}\tilde{T}_{c,i,t} + \gamma_{\tilde{T},i}\tilde{T}^2_{c,i,t} + \beta_{\tilde{R},i}\tilde{R}_{c,i,t} + \gamma_{\tilde{R},i}\tilde{R}^2_{c,i,t}) + \sum_{y\in\mathcal{Y}\setminus\{Y_1\}} \beta_y I_y(t) \\ + \sum_{\bar{c}\in\mathcal{C}\setminus\{1,m\}} \beta_{\bar{c}} I_{\bar{c}}(c) + \mu_{c,t} \end{aligned} \tag{2.6}$$

Based on Model (2.6), we can estimate the percentage change in land values corresponding to a minor change in seasonal temperature and precipitation. The percentage change is called the marginal impact in percentage (Van Passel et al., 2017). We calculate the marginal impacts in percentage of seasonal temperature, denoted as $F(\tilde{T}_{c,i,t})$, by taking the partial derivative of $V_{c,t}$ at $\tilde{T}_{c,i,t}$ and dividing this value by $V_{c,t}$. Similarly, we can calculate the marginal impacts in percentage of seasonal precipitation, denoted as $F(\tilde{R}_{c,i,t})$. Consequently, we have the following two equations (Van Passel et al., 2017):

$$F(\tilde{T}_{c,i,t}) := \frac{\partial V_{c,t}}{\partial \tilde{T}_{c,i,t}}/V_{c,t} = \beta_{\tilde{T},i} + 2\gamma_{\tilde{T},i}\tilde{T}_{c,i,t} \quad \forall c \in \mathcal{C}\setminus\{m\}, i \in \mathcal{I}, t \in \mathcal{T} \tag{2.7}$$

$$F(\tilde{R}_{c,i,t}) := \frac{\partial V_{c,t}}{\partial \tilde{R}_{c,i,t}}/V_{c,t} = \beta_{\tilde{R},i} + 2\gamma_{\tilde{R},i}\tilde{R}_{c,i,t} \quad \forall c \in \mathcal{C}\setminus\{m\}, i \in \mathcal{I}, t \in \mathcal{T} \tag{2.8}$$

These two equations will be used to understand the marginal impacts of seasonal temperatures and precipitations on farmland values under climate change in Section 2.4.3.

2.4.2 Data Source and Data Prepossessing

To fit Model (2.6) and predict future farmland values, we collected the most proper and detailed data available on historical farmland values, temperature, precipitation, future climate scenarios. We also present the patterns of the forecasted climate data for each season in the MRB.

Farmland data: The farmland value is the total asset value including agriculture land and farm buildings. The farmland value data come from the United States Census of Agriculture (USDA) and are available every five years, specifically in 1997, 2002, 2007, 2012, and 2017 (National Agricultural Statistics Service, United States Department of Agriculture, 2017). The limitation of selecting data every fifth year is that we can only observe price trends at each of these five-year intervals in the dataset, whereas we need the trends at an annual level.

These data are observed at the finest county level, and the farmland value is measured in $/acre. As some counties missed these data in some census years, we drop the counties without complete data of all census years, i.e., 1997, 2002, 2007, 2012, and 2017, leaving 545 counties. All farmland values are converted into U.S. dollars in 2017 by multiplying the sum of one and the cumulative inflation rate between 2017 and the targeted years. Specifically, the cumulative inflation rates from 1997, 2002, 2007, and 2012 to 2017 are 0.527, 0.363, 0.182, and 0.068, respectively (CoinNews Media Group, 2022).

Histogram plots in Appendix A.1 reveal that the farmland data are highly right-skewed with a long right tail, while the log-transformed farmland data are less skewed and more symmetric. These patterns indicate a preference for a logarithmic regression model in our

case.

The longitude and latitude of the centroids of the counties are obtained from the *housingData* R package (v0.3.0; Hafen (2016)). Farmland size data in each U.S. county is derived from the United States Census Bureau (The United States Census Bureau, 2012).

Historical seasonal temperature and precipitation data: We rely on the county-level data from the National Oceanic and Atmospheric Administration's Climate Divisional Database (nClimDiv) (Vose et al., 2014a) for the contiguous United States. These data are bias-corrected and provide monthly average (surface air) temperature and monthly precipitation values for each county, spanning from 1895 to the present (Vose et al., 2014b). We use these data to develop the seasonal measures of average temperature (Fahrenheit) and total precipitation (inch/month) for the farmland in each county (in the MRB) for the five census years (i.e., 1997, 2002, 2007, 2012, and 2017), see Appendix A.4 for the details.

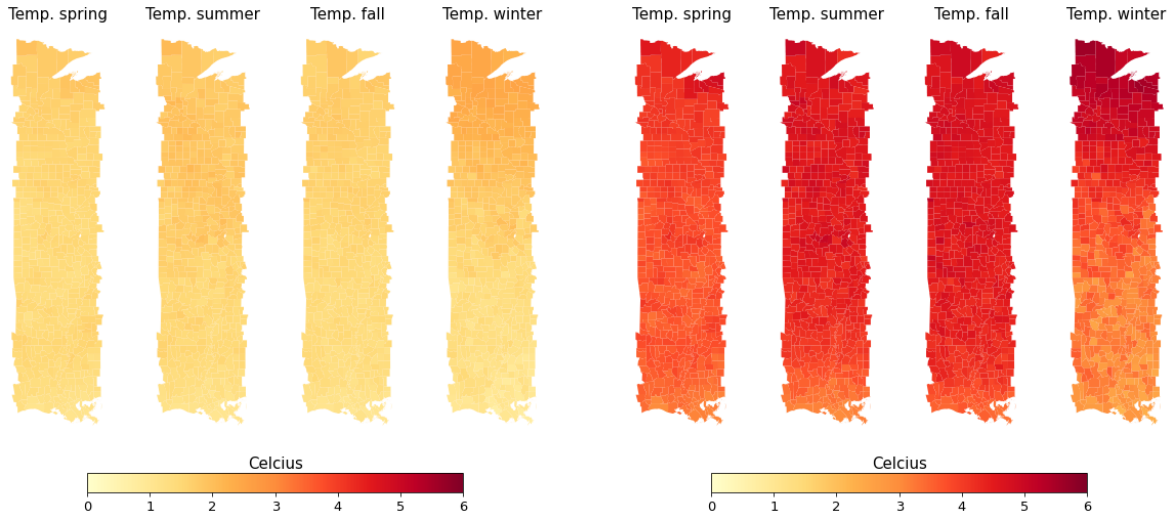
Projections for future seasonal temperature and precipitation: We generate these projections based on the high-resolution Localized Constructed Analogs (LOCA) CMIP5 Projections (Pierce et al., 2014). There are 64 LOCA climate projections (or climate scenarios) generated by 32 climate models under RCP4.5 and RCP8.5, respectively. Each LOCA climate projection provides the bias-corrected daily precipitation, maximum surface air temperature, and minimum surface air temperature at about 6 km $\times$ 6 km kilometer grid cells for the contiguous United States from 2006 to 2099 (or 2100). The data on those climate projections are used to develop month-by-year measures of precipitation and average surface air temperature covering the farmlands in each county in the MRB during the 2023-2099 period. To obtain the monthly average surface air temperature projections, we

compute the mean of the daily maximum surface air temperature and minimum surface air temperature obtained from the LOCA dataset for each month. Similarly, to obtain the monthly precipitation projections, we sum up the daily precipitation values obtained from the LOCA dataset for each month.

To obtain county-level climate projections, we downscale the monthly average surface air temperature and precipitation projections to match each county. We perform the downscaling using the closest downscaling method, where for each county a grid location that includes the county's centroid center point is identified and the climate projections for that location are used as the projections for that specific county. The raster R package (version 3.6.14, Hijmans (2023)) is utilized for implementing this downscaling procedure.

We use these downscaled month-by-year data to develop the seasonal measures of average temperature and precipitation (inch/month) for the farmland in each county (in the MRB) for the 2021-2090 period. Refer to Appendix A.4 for more details. The derived seasonal data serve as inputs of Model (2.6) for forecasting future land prices,

Forecasted climate in the MRB: We use the obtained county-level seasonal temperature and precipitation projections to present climate change in the MRB over the study horizon. To avoid numerous figures, we use the average temperature and precipitation over all 32 GCMs to show climate change between 2023 and 2090 at county level. Figures 2.1 and 2.2 respectively present net differences in temperature and precipitation by county between the years of 2023 and 2090. In Figure 2.1, the mean temperature in 2090 increases by about 2 °C under RCP4.5 and by about 5 °C under RCP8.5 over all four seasons compared to 2023. The figure also shows that temperature will increase slightly more in the North



(a) RCP4.5 (b) RCP8.5

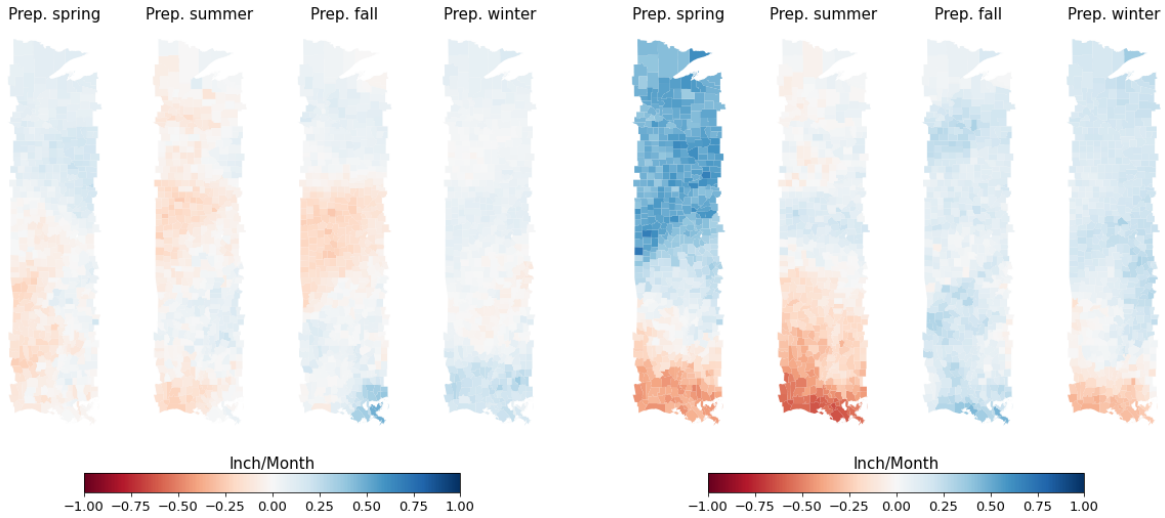
Figure 2.1: Seasonal temperature change between 2023 and 2090 in the MRB

of the MRB than in the South in all four seasons under both scenarios. In contrast to temperature, changes in precipitation, shown in Figure 2.2, are more varied. The mean precipitation in 2090 increases in the North over the Spring, Fall, and Winter under both RCPs, while precipitation either increases or decreases in the remaining area compared to 2023. Unlike the other seasons, the figure shows that summer precipitation change varies even at similar latitudes.

2.4.3 Farmland Value Results

Based on the data described in Section 2.4.2, we fit Model (2.6) and forecast farmland values.

Fitting regression model: Utilizing the *Farmland data* and *Historical Seasonal Tem-*



(a) RCP4.5 (b) RCP8.5

Figure 2.2: Seasonal precipitation change between 2023 and 2090 in the MRB

perature and Precipitation data in years $\mathcal{Y} = \{1997, 2002, 2007, 2012, 2017\}$, we employed the ordinary least square method (Dunn et al., 2018) via R Statistical Software (v3.6.3; R Core Team (2020)) to fit Model (2.6) for the MRB case.

We present the results of the best-fitted Model 2.6 in the MRB case in Table 2.1. In this table, the fall temperature, winter temperature, fall precipitation, and winter precipitation are 95% statistically significant. Moreover, if the local fall temperature (respectively (resp.) precipitation) rises, the farmland values in the MRB will continuously increase before reaching their turning point, i.e., 59.8 °F or 15.4 °C (resp. 7.05 inch or 17.9 cm). Because slightly warmer and wetter falls extend the growing season, farms produce greater yields and increase their revenue (Liang et al., 2017). However, the rise of winter temperature (resp. precipitation) will decrease farmland values before it reaches the turning point of 31.85 °C (resp. 14.7 cm). This decrease may occur because a warmer and wetter winter

may encourage diseases and insect pests that reduce farmland yields and revenue (Deressa, 2007). Overall, the rise of temperature and precipitation in fall and winter have opposite effects on farmland values in the MRB.

Further, we show that the marginal impacts of seasonal temperature and precipitation on land values vary over the MRB. By Equations (2.7) and (2.8), we calculated the marginal impact in percentage of land values for the mean climate scenario in 2023. In Figure 2.3, we present the seasonal temperature and precipitation marginal impacts in percentage across the MRB at the county level under RCP4.5. We do not need to show the marginal impacts in percentage under RCP8.5 in 2023 as the climate is similar under these two RCPs in 2023. The marginal impacts in percentage under RCP4.5 and RCP8.5 in 2090 can be found in Appendix A.7. In Figure 2.3, farmland values decrease with warmer spring and winter temperatures across the MRB and decrease with warmer fall temperatures in the South of the MRB. In the Northern MRB, farmland values increase with warmer summer and fall temperatures. On the other hand, farmland values across the MRB decrease with wetter springs, summers, and winters. However, farmland values across the MRB increase with a wetter fall. Overall, examining all four seasons, warmer temperatures may benefit farmland values in the Northern MRB, and more precipitation may harm farmland values across the MRB.

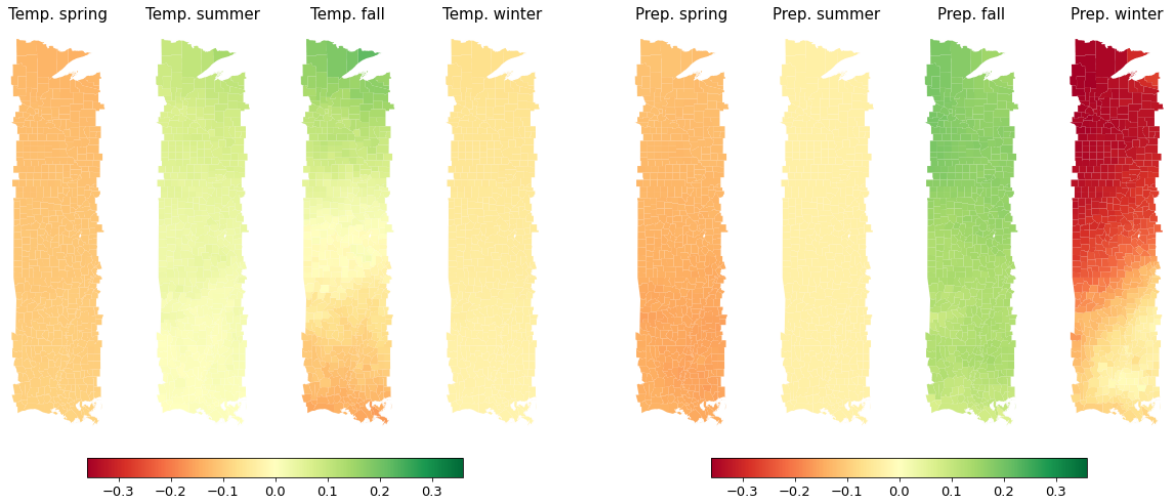
Farmland value forecasts from climate data: To forecast future farmland values and corresponding confidence intervals for each climate scenario and each county within the MRB, we incorporated the *Projections for Future Seasonal Temperature and Precipitation* data into the best-fitted Model (2.6). We assume that in future years, farmland in the MRB would have the same annual conditions as in the year 2017, except for climate conditions.

Table 2.1: Summary of the best-fitted Model (2.6)

Independent variables	Coeff.	Std. Error	P-value	95% CI
Spring temperature	-0.1892	0.1216	0.12	[-0.4277, 0.0493]
Spring temperature squared	0.0007	0.0011	0.54	[-0.0015, 0.0028]
Spring precipitation	-0.0876	0.0480	0.07	[-0.1817, 0.0066]
Spring precipitation squared	0.0059	0.0048	0.22	[-0.0036, 0.0154]
Summer temperature	0.4494	0.2388	0.06	[-0.019, 0.9178]
Summer temperature squared	-0.0027	0.0016	0.09	[-0.0058, 0.0004]
Summer precipitation	-0.0388	0.0305	0.20	[-0.0985, 0.0210]
Summer precipitation squared	0.0005	0.0027	0.86	[-0.0049, 0.0059]
Fall temperature	0.7041	0.1025	0.00*	[0.5031, 0.9050]
Fall temperature squared	-0.0059	0.0009	0.00*	[-0.0077, -0.0041]
Fall precipitation	0.2857	0.0540	0.00*	[0.1799, 0.3915]
Fall precipitation squared	-0.0203	0.0069	0.00*	[-0.0338, -0.0067]
Winter temperature	-0.0843	0.0129	0.00*	[-0.1097, -0.0589]
Winter temperature squared	0.0005	0.0002	0.03*	[0.0000, 0.0009]
Winter precipitation	-0.4261	0.0445	0.00*	[-0.5133, -0.3389]
Winter precipitation squared	0.0368	0.0041	0.00*	[0.0287, 0.0449]
Number of observations: 2725				
Residual standard error: 0.1465 on 2160 degrees of freedom				
Multiple R-squared: 0.923; Adjusted R-squared: 0.9029				
F-statistic: 45.91 on 564 and 2160 DF, p-value: $< 2.2e^{-16}$				
*: 95% statistical significance; Coeff.: coefficient; Std.: standard; CI: confidence interval				

Mathematically, year indicators for the future are identical to the year indicator in 2017 when predicting farmland values, denoted by $\sum_{y \in \mathcal{Y} \setminus \{1997\}} \beta_y I_y(t) = \beta_{2017}$ $(\forall t \in \mathcal{T})$.

For convenience's sake, we measure future farmland prices in 2023 USD. We use a 0.02 constant yearly inflation rate to discount future farmland values to 2023 USD because the Federal Open Market Committee targets a 0.02 inflation rate as a long-term goal (Board of Governors of the Federal Reserve System, 2021).



(a) Temperature (Temp.)

(b) Precipitation (Prep.)

Figure 2.3: Marginal impact in percentage of farmland values of mean climate in 2023 under RCP4.5

2.4.4 Discussion of Farmland Value Forecasts

Here we discuss how the specifics of a climate trajectory translate into differences in land valuations. Table 2.1 shows the coefficients of the regression model used to predict farmland values from a particular climate forecast. These relationships are established from historical data linking county-level seasonal climate values to prices and show that season dramatically changes how temperature and precipitation impact the value of agricultural land. In particular, Figure 2.3 shows that, for RCP4.5 and over the full horizon, the average climate scenario forecasts warmer spring and winter temperatures harm farmland values across the MRB, while warmer summer and fall temperatures benefit farmland values in the Northern MRB. More precipitation in springs, summers, and winters harm farmland values across the MRB, while wetter fall seasons benefit valuations. These seasonal and

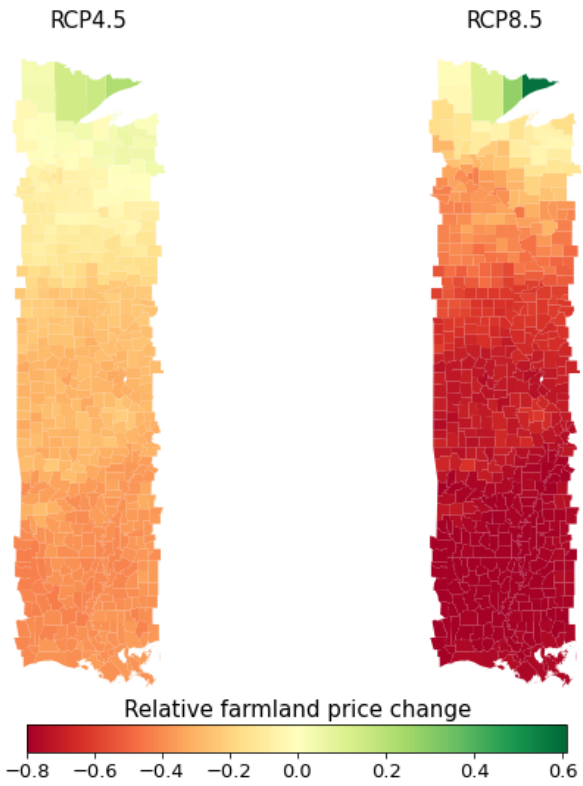


Figure 2.4: Relative farmland price changes between mean of 2081-2090 and mean of 2023-2032

locational differences imply that caution needs to be applied when considering aggregating climate values whether the aggregation is annual, over a large geographic footprint, or between climate models. Two climate forecasts may have similar average precipitation and temperature while having very different land value trajectories because of differences in how precipitation and temperature are distributed.

2.5 Study II: Investment Decisions in the MRB under Climate Change Scenarios

In this section, we conduct an empirical study employing linear robust Model (2.5) to investigate land investment decisions made by a farmland investor in the MRB under different climate scenarios. We set 2022 as period 0 (current period) for the investor. We study the investor's farmland investment from 2023 to 2090 in counties whose centroid centers are located between the longitudes of 94.5° W and 89.0° W. The empirical study consists of two key steps. First, we introduce the methodology for the empirical study of land investment decisions. Second, we present the farmland investment outcomes for both the benchmark case and the robust case.

2.5.1 Methodology for the Empirical Study of Investment Decisions

The methodology for the empirical study of investment decisions has three steps. First, based on the best-fitted regression model (2.6), we formulate farmland price uncertainty as intervals for each climate scenario. These intervals give price parameters for solving Model (2.5) under different climate scenarios. Second, we model the budgets of uncertainty for Model (2.5). Third, for the same 68-year period, we solve Model (2.5) multiple times under different scenarios to investigate farmland investment decisions made in the MRB case.

Modeling farmland value uncertainty: Based on the best-fitted Model (2.6), we model future farmland price uncertainty as confidence interval trajectories with two steps. This

modeling process involves generating confidence interval trajectories, offering lower and upper bounds of the uncertain land prices for each county. Intuitively, these confidence intervals represent the investors' comprehension of the potential maximal variation, or the worst-case scenario, in land prices for each county. Wider confidence intervals signify greater land price fluctuations and the potential occurrence of extreme prices in each county, while narrower intervals indicate reduced variations and smaller extreme prices.

First, for each climate scenario generated by the combination of GCMs and RCPs, we predicted the mean of $lnV_{t,c}$ ($\forall c \in \mathcal{C} \setminus \{m\}$, $t \in \mathcal{T}$) and its corresponding α% confidence interval (CIα), where α% is the confidence level (Devore, 2011), for each county and each year. We denote the mean as $ln\bar{V}_{t,c}$ and the range of the CIα as $2d_{t,c}(\alpha)$, where $d_{t,c}(\alpha)$ equals half of the range. Consequently, we can express the CIα for $ln\bar{V}_{t,c}$ as $[ln\bar{V}_{t,c} - d_{t,c}(\alpha), ln\bar{V}_{t,c} + d_{t,c}(\alpha)]$. In our case, we employ both 68% confidence intervals (CI68) and 95% confidence intervals (CI95). The CI95 permits a broader range of price variation compared to the CI68.

Second, we transform the CIα for $ln\bar{V}_{t,c}$ into the interval for P_{tc}. We let e be the exponential constant, $\bar{p}_{tc} = \frac{1}{2}(e^{ln\bar{V}_{t,c}-d_{t,c}(\alpha)} + e^{ln\bar{V}_{t,c}+d_{t,c}(\alpha)})$, and $\hat{p}_{tc} = \bar{p}_{tc} - e^{ln\bar{V}_{t,c}-d_{t,c}(\alpha)}$. We model uncertain P_{tc} ($\forall c \in \mathcal{C} \setminus \{m\}$ and $t \in \mathcal{T}$) as a random variable that is bounded, symmetric, and independent, and its values are confined within $[\bar{p}_{tc} - \hat{p}_{tc}, \bar{p}_{tc} + \hat{p}_{tc}]$ $(= [e^{ln\bar{V}_{t,c}-d_{t,c}(\alpha)}, e^{ln\bar{V}_{t,c}+d_{t,c}(\alpha)}])$.

Note that when $\Gamma_t = 0$ ($\forall t \in \mathcal{T}$) the investor ignores price variance and uses mean prices. In this case, we let $\bar{p}_{tc} = \bar{V}_{t,c}$ and $\hat{p}_{tc} = 0$ ($\forall t \in \mathcal{T}$, $c \in \mathcal{C} \setminus \{m\}$). In essence, the investor assumes no uncertainty in pricing and considers only the mean values.

Modeling budgets of uncertainty: In our case, we determine the budgets of uncertainty for each year as percentages of the total number of counties: $\Gamma_1 = \Gamma_2 = ... = \Gamma_T = \Gamma := \lambda(m-1)$, where $\lambda \in [0,1]$, for two main reasons. First, this way allows constant conservatism over time and does not rely on extra parameters. This allows our model to be simpler and easier to solve. Second, this way also enables us to study the non-conservative case ($\Gamma = 0$) and the most conservative case ($\Gamma = m-1$). Setting $\Gamma = 0$ leads to decision makers ignoring price variance and using only the value $\bar{p}_{tc}$ ($c \in \mathcal{C} \setminus \{m\}$). If $\Gamma = m-1$, decision makers are at their most pessimistic, considering the farmland prices within the confidence interval which minimize the objective value. Note that we keep the price deterministic at the artificial county m.

Solving the Robust model at Different Scenarios: We solve Model (2.5) multiple times under different scenarios. These scenarios come from the combinations of climate scenarios, confidence levels, and conservatism levels. We form the scenarios using a two step process, after which, we solve the model under these scenarios.

First, we formed the combinations of 64 climate scenarios and two confidence levels to generate price interval trajectories. We used 64 climate scenarios, which were generated from 32 GCMs and 2 RCPs as documented in (Bracken, 2016). With two confidence intervals, CI68 and CI95, we have 128 farmland price interval trajectories for each county in the MRB from 2023 to 2090. We used these interval trajectories as the price parameters of Model (2.5) when solving the model.

Besides the price parameters, we set the remaining (economic) parameters in Model (2.5) as follows: inflation rate $\beta = 0.02$, external income $B_t = 60,000$ ($t \in \mathcal{T} \setminus T$),

transaction cost rate on buying land $r_b^l = 0.01$, transaction cost rate on selling land $r_s^l = 0.06$, transaction cost rate on buying risk-free asserts $r_b^\beta = 0$, and transaction cost rate on selling risk-free asserts $r_s^\beta = 0.01$. For more detailed information about these economic parameters, please refer to Appendix A.5.

Second, we further formed the combinations of these 128 interval trajectories with different conservatism levels. For conservatism levels, we chose 16 different levels: $\{0, 0.01(m-1), 0.02(m-1), 0.03(m-1), 0.04(m-1), 0.05(m-1), 0.1(m-1), 0.2(m-1), ..., 0.9(m-1), (m-1)\}$. For notational simplicity, we used only decimals to express conservatism levels, denoted as $\{0, 0.01, 0.02, ..., 0.05, 0.1, 0.2, ..., 0.9, 1\}$. Therefore, we need to solve the model for 2048 ($= 128 \times 16$) scenarios.

Finally, we conducted the solution process for the linear robust Model (2.5) encompassing these 2048 scenarios, leading to 2048 optimal solutions. For each scenario, the model has 110,292 variables and 220,583 constraints, and the detailed cardinality specifying the numbers of variables and constraints can be found in Appendix A.6. The solving process utilized CPLEX (20.1.0) and ran on a 24-core Linux machine with 128 GB of RAM. The total time required to solve the 2048 scenarios was 40 hours, with an average of 70 seconds needed to address each instance.

2.5.2 Farmland Investment Results

This section presents numerical results illustrating investment decisions in the MRB in response to the 64 climate scenarios. The section is divided into two parts. As a benchmark, the first part studies investment decisions in the MRB using the estimated mean farmland

price. The second part studies robust investment decisions for the 64 climate scenarios.

Benchmark farmland investment decisions

In this part, we study optimal farmland investment decisions without price uncertainty under climate change. Specifically, we investigate investment decisions under the average climate scenario as a benchmark case for this research.

We averaged county-level temperature and precipitation projections equally over all 32 GCMs for RCP4.5 and RCP8.5. For each RCP, we input the average projections into the regression model (2.6) to generate mean farmland price trajectories at the county level from 2023 to 2090. Then, we input the trajectories into the robust land investment model (2.5) (with $\Gamma_t = 0$, $\forall t \in \mathcal{T}$) to get the optimal land buying and selling decisions for each RCP. These decisions tell investors to buy and sell farmland, in which counties, at which period, and of how many acres in the MRB.

We present the geographical range of the optimal farmland investment decisions in the MRB. In Figure 2.5, the right subfigure shows a map of the MRB. In the left and middle subfigures, the x-axis represents years and the y-axis represents scaled latitudes. We use blue (resp. violet) dots to represent the geographical latitudes of centroids of counties with farmland buying (resp. selling) decisions. The latitudes are scaled by minus 29.43° N, which is the minimal latitude of a county's centroid in the MRB. We use black-dashed lines to represent the upper bound (=18.8) and lower bound (=0.0) of these scaled latitudes. For convenience purposes, we let the scaled latitudes of centroids of counties without farmland buying and selling decisions, i.e., with no acquisition, be -1. In addition, we use red-

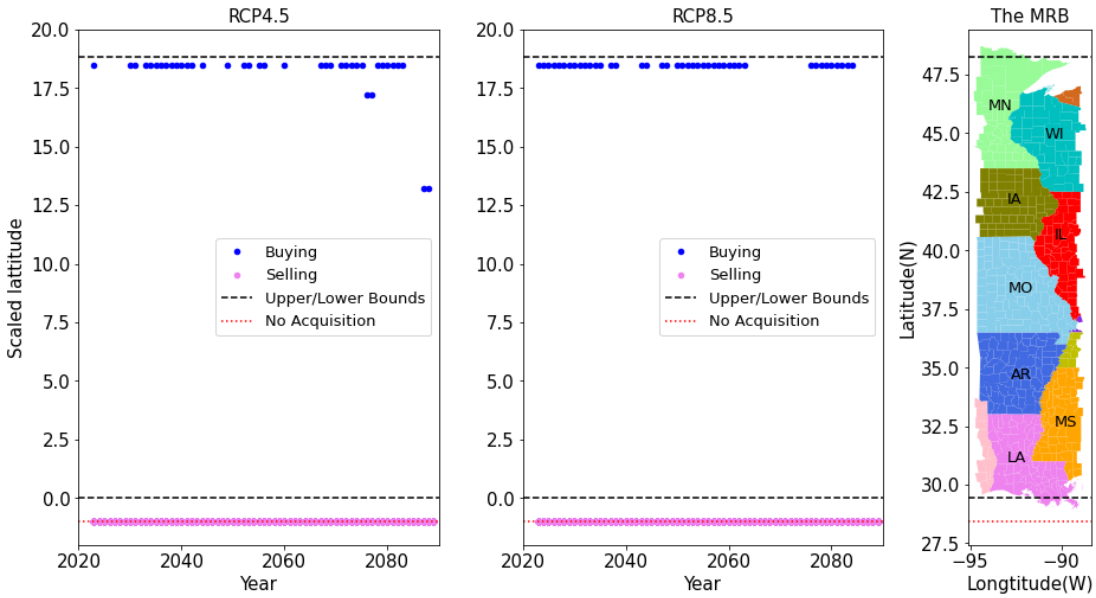


Figure 2.5: The farmland investment latitude in the MRB under the average climate scenario

dashed lines to represent this no-acquisition decision. The middle of the scaled latitude range [0,18.8] is 9.4.

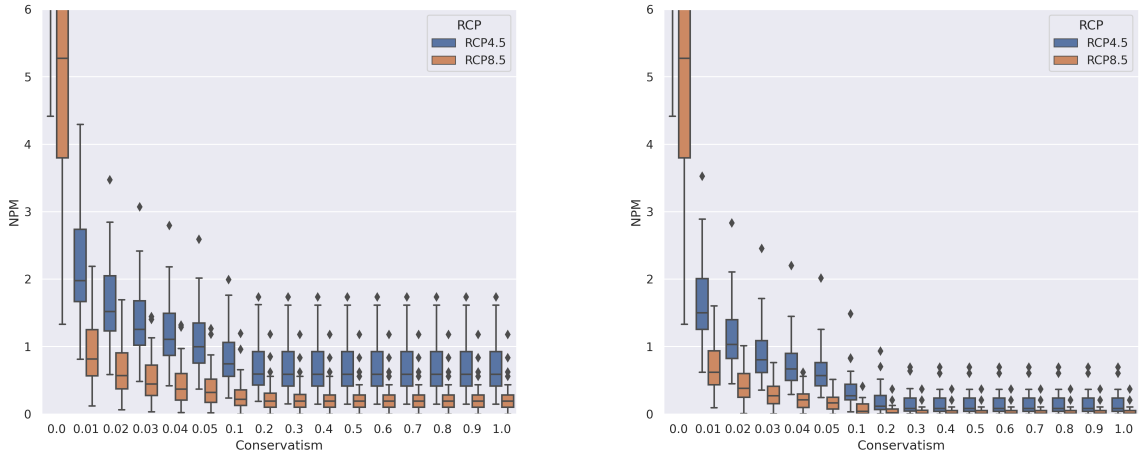
In Figure 2.5, generally, the scaled latitudes for buying under both RCPs are close to 18.5 over the planning horizon, which always suggests investment in the northernmost MRB (i.e., in MN). However, there are some exceptions in certain years under RCP4.5. The scaled latitudes for buying equals about 17.2 in 2076-2077 and 13.2 in 2087-2088, suggesting buying farmland in WI (according to the optimal investment decisions). On the other hand, for selling, the scaled latitudes equal -1, indicating that the investor retains their farmland investments under both RCPs. Overall, under both RCPs, buying and holding farmland in the Northern MRB is the optimal investment decision based on the mean temperature and precipitation of all 32 GCMs.

Robust investment decisions for forecasted climate scenarios

We present robust farmland investment decisions under multiple scenarios from three perspectives: investor profitability for all scenarios and conservatism levels, examples of investment details for specific climate scenarios, and summarised investment strategies for all scenarios and conservatism levels.

The optimal investment profits: Based on the optimal investment decisions, we investigated the relationship between the optimal investment profits (the optimal objective values of Model (2.5)) and different climate and conservatism scenarios. For presentation purposes, instead of showing these profits, we show the net profit margins (NPMs) that have been calculated from these profits. Specifically, the NPM is calculated by (the optimal objective value - cost)/cost, where the cost equals the discounted cumulative external incomes during 2023-2089 (i.e., $\sum_{t=1}^{67} \frac{1}{(1+\beta)^{t-1}} B_t$). For comparison purposes, we measure the optimal objective values, the cost, and NPMs, all in 2023 USD.

Figures 2.6a and 2.6b present Box-and-whisker plots for NPMs over different RCPs and conservatism scenarios under CI68 and CI95, respectively. Each plot represents the interquartile range of 32 NPMs from 32 GCMs in a box that has the 25th percentile, median, and 75th percentile displayed within. Besides the box, the diamond-shaped points above the whiskers are outliers: more precisely they are extreme values in the 32 NPMs. In both figures, all the NPMs decrease and then converge as the conservatism level increases. As the conservatism level passes 0.3, all the NPMs in both figures converge to a value that varies among RCPs and CIs. In addition, in both figures, NPMs under RCP4.5 converge to a bigger value than NPMs under RCP8.5 do. The medians of NPMs under RCP4.5



(a) CI68 (b) CI95

Figure 2.6: The box-and-whisker plots for net profit marginals (NPMs)
Each boxplot displays the variation of the NVMs of 32 GCMs at a certain conservatism.

converge to 0.59 (for CI68) and 0.08 (for CI95), while the medians of NPMs under RCP8.5 converge to 0.19 (for CI68) and 0.01 (for CI95), respectively. The convergence makes sense as a higher conservatism level makes the land price uncertainty range wider, then buying at a high price and selling at a low price generate less money. In addition, the investors can expect more profit under RCP4.5 than under RCP8.5, again due to larger decrease in land prices under RCP8.5.

Investments for the HadGEM2-ES GCM: To understand the details of how investments take place in a specific climate scenario, we examined the individual purchases and sales actions in response to scenarios derived from the HadGEM2-ES GCM (Bellouin et al., 2011). We chose this climate model because it is widely used to generate high-resolution regional climate projections in North America (Mearns et al., 2017), and it contains the

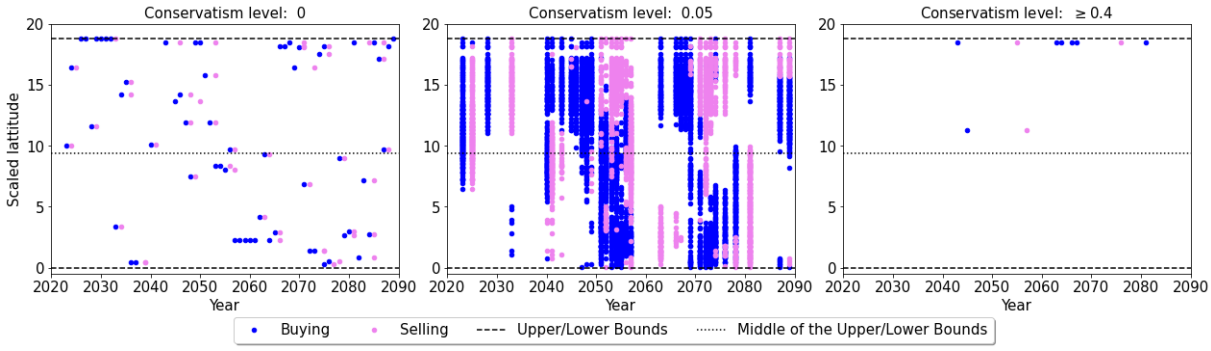
most complete climate processes among the HadGEM2 family (Bellouin et al., 2011).

Because land valuations tend to vary in the North-South axis, we focus on presenting the latitude of farmland buying decisions in the MRB during 2023-2089 under the HadGEM2-ES. Figure 2.7 and Figure 2.8 provide visual representations of the scaled latitudinal buying/selling decisions (represented by blue/violet dots) across various levels of conservatism (0, 0.05, 0.4, and 1). In both figures, the x-axis represents years from 2023 to 2089, the y-axis represents the scaled latitude from 0 to 18.8, and the three subfigures show the buying/selling latitudes at levels of conservatism: 0, 0.05, $\geq$0.4.

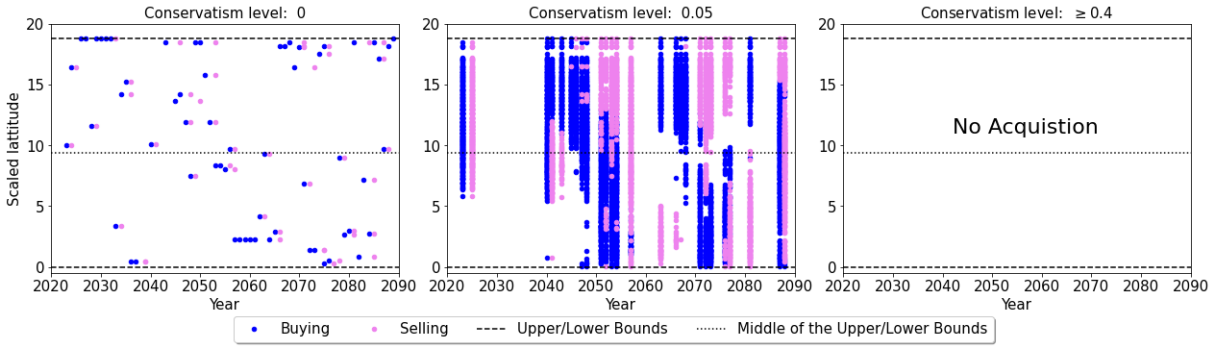
In Figure 2.7a, corresponding to condition RCP4.5 and CI68, both buying and selling occur over the entire MRB area at conservatism levels 0 and 0.05. We note that when conservatism level increases from 0 to 0.05, in periods where investments are made, they are spread over a larger set of latitudes to reduce exposure to worst-case outcomes. At conservatism levels greater than 0.4, acquisitions and sales occur exclusively in the Northern MRB. This pattern indicates that investment decisions encompass the entire MRB at lower conservatism levels but shift towards the Northern MRB as conservatism increases beyond 0.4. At the larger confidence interval for land price valuations, CI95, worst-case outcomes are more severe. Under RCP4.5 and CI95, no investment occurs when conservatism levels surpass 0.4 (Figure 2.7b).

Figure 2.8 shows investments for RCP8.5. The results are similar to the results for RCP4.5 except that in all cases there is a reduction in the number of land acquisitions. We observe that for CI68, at lower levels of conservatism there remain some opportunities over the entire set of latitudes. At CI95, there are no profitable investments for even modest

levels of conservatism.



(a) CI68

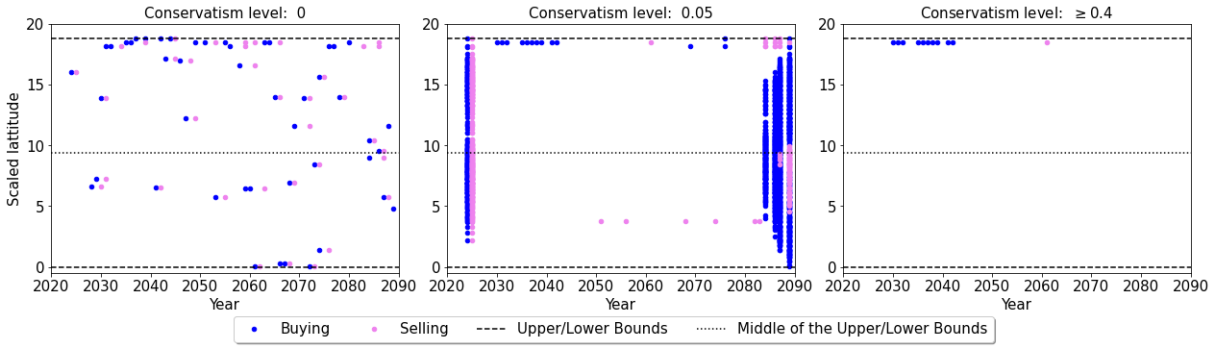


(b) CI95

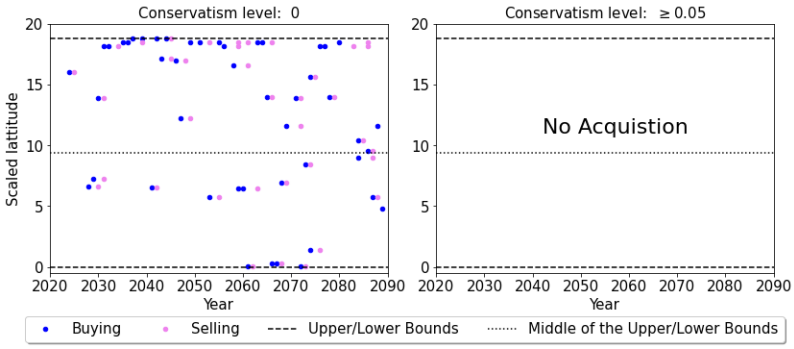
Figure 2.7: Farmland investment latitude in the MRB based on the HadGEM2-ES GCM under RCP4.5

Investment strategies for all climate scenarios: To understand how different climate scenarios affect farmland investment, we studied the investments for all 32 GCMs under both RCP4.5 and RCP8.5 at a range of conservatism levels.

Table 2.2 and 2.3 present the ranges of scaled buying latitudes in the 2080s under CI68 for RCP4.5 and RCP8.5, respectively. In these tables, the first column corresponds to the GCM, the first row represents conservatism levels, and NI denotes the abbreviation



(a) CI68



(b) CI95

Figure 2.8: Farmland investment latitude in the MRB based on the HadGEM2-ES GCM under RCP8.5

for no investment. In Table 2.2, as the conservatism level increases, the ranges shrink to cover just the Northern MRB under nine of the 32 GCMs (e.g. CESM1-CAM5 (Neale et al., 2010) and inmcm4 (Volodin et al., 2010)). However, under the remaining 23 GCMs, the farmland investment portfolio suggests different investment strategies. 11 of the 23 GCMs suggest investment in the South, 10 of 23 GCMs suggest investment across the North and South, and two of 23 GCMs suggest no investment. These patterns from the 32 GCMs show that the portfolio of investment decisions varies over GCMs. In addition, the patterns are different from the previous conclusions for the mean climate scenario (the benchmark case), which suggests always investing in the North. Similarly, in Table 2.3, under RCP8.5, the portfolio of the investment decisions again varies among GCMs. eight of the 32 GCMs lead to investing only in the North, while four GCMs lead to investing only in the South. However, investors end up with no investment decision for 20 of the 32 GCMs as the conservatism level increases.

Similarly, Table 2.4 and 2.5 show the ranges of the scaled buying latitudes in the 2080s under CI95. At high conservatism levels, under conditions RCP4.5 and CI95, three of the 32 GCMs suggest investing in the North, five GCMs suggest investing in the South, and 24 GCMs suggest no investment. Under RCP8.5 and CI95, 30 of the 32 GCMs suggest no investment. When comparing CI68-based and CI95-based ranges of the scaled buying latitudes, during the same period (the 2080s), fewer GCMs suggest investing in the MRB. This phenomenon occurs because CI95 indicates higher price variances than CI68 and increases the cost of dealing with worst-case price deviation.

2.5.3 Discussion of Forecasted Farmland Investments

The empirical study reveals how granular details of temperature and precipitation trajectories associated with different climate models may impact investment decisions in the MRB. These results provide insights into how challenging to quantify disagreements between forecasts translate into different investment strategies.

The challenge associated with considering aggregated forecasts is evident when we compare the investments associated with the benchmark model with the results from climate forecasts associated with individual climate models. The benchmark uses average climate values for each season/county. For both RCPs, the benchmark climate leads to farmland values growing in the northernmost latitudes of the MRB but decreasing in other areas. The investment strategies associated with the benchmark model are the expected: the most profitable strategy is to invest at Northern latitudes. When investment strategies associated with individual climate models are examined, the results are very different (see Tables 2.2-2.5). At higher conservatism levels where less robust investment opportunities are filtered out, investment latitudes are highly contingent on the forecast. For example, at RCP4.5, CI68, and a conservatism level of 0.5, multiple forecasts lead to investment specifically in Northern (e.g. CESM1-CAM5, CMCC-CMS), Southern (e.g. ACCESS1-3, CCSM4) latitudes and over both Northern and Southern latitudes (e.g. GISS-E2-R, MIROC-ESM-CHEM). Of note, for this case, only 28% of the climate scenarios find that the benchmark strategy of investing at extreme Northern latitudes is robust. At RCP8.5, differences in the climate scenarios also lead to a range of strategies. However, at this higher emissions pathway, 63% of the scenarios do not have robust profitable investment

opportunities (conservatism level of 0.5, CI68).

These results point to the need to consider individual climate forecasts independently when planning or evaluating adaptation strategies. It is important that the granular details of individual climate paths and the dynamics of the particular investment or decision problem are considered when determining a strategy and evaluating which communities are likely to enjoy economic benefits or detriments.

2.6 Conclusion

This chapter has two main sets of contributions: (1) The development of a scalable methodology for assessing investment decisions on a horizon and scale suitable for climate adaptation research. (2) The application of these methods to a case study of how a farmland investor's decisions are contingent on valuations formed from different climate forecasts.

The solution methods extend the work of Bertsimas and Sim (2003) and Bertsimas and Sim (2004) on robust optimization with varied conservatism to a sequential investment problem where holdings may be held indefinitely. We show that this problem can be linearized and scales appropriate both in terms of the number of assets and the investment horizons to be useful for the study of climate adaptation.

The empirical study uses these robust optimization methods to study investment in over 500 counties in the Mississippi River Basin from 2023 to 2090. The study reveals variations among climate scenarios that will impact the profitability and strategy of robust investment decisions. Specifically, we determined the optimal robust investment decisions

by applying the robust optimization methods developed for sequential investment decisions to trajectories of land values implied by a family of climate models used by the IPCC, considering both low and high emission pathways.

For the low emission pathway, RCP4.5, the benchmark climate scenario, determined by averaging climate variables over the climate models, leads to investment consistent with "conventional wisdom": invest at Northern latitudes of the MRB. However, the investment patterns associated with individual climate scenarios are highly varied. Depending on the particular scenario, economic beneficiaries may be at many different locations over the MRB.

At the higher emissions RCP8.5 pathway, investment associated with the benchmark scenario remains profitable at Northern latitudes. However, for the individual climate scenarios, the level of profitability decreases and there are few robust investment opportunities. This does not imply that the land is no longer productive, but rather that the value of the land does not reliably increase over the investment horizon.

The primary takeaway from these findings reinforces the importance of considering a range of granular climate scenarios with detailed seasonal and locational data when evaluating climate adaptation decisions.

Table 2.2: The range of scaled buying latitude under RCP4.5 and CI68 in the 2080s

GCM \ Conservatism	0	0.05	0.2	0.5	1
ACCESS1-0	[1.8, 18.5]	[0.0, 18.8]	NI	NI	NI
ACCESS1-3	[4.8, 16.0]	[3.9, 18.8]	[0.0, 18.8]	[7.2, 8.4]	[7.2, 8.4]
CCSM4	[1.3, 15.3]	[0.0, 17.2]	[0.0, 18.5]	[0.8, 1.7]	[0.8, 1.7]
CESM1-BGC	[0.3, 18.5]	[0.0, 18.8]	[4.0, 18.5]	[8.3, 18.5]	[8.3, 18.5]
CESM1-CAM5	[9.3, 18.8]	[9.2, 18.8]	[0.0, 18.8]	[12.3, 18.8]	[12.3, 18.8]
CMCC-CMS	[3.3, 18.5]	[0.0, 18.8]	[0.0, 18.8]	[14.0, 14.9]	[14.0, 14.9]
CNRM-CM5	[1.4, 18.5]	[0.0, 18.8]	[0.0, 17.5]	8.4	8.4
EC-EARTH	[3.3, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[1.8, 2.2]	[1.8, 2.2]
FGOALS-g2	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[12.6, 14.2]	[12.6, 14.2]
GFDL-CM3	[0.5, 18.1]	[0.0, 18.8]	[0.0, 14.3]	[9.2, 9.6]	[9.2, 9.6]
GFDL-ESM2G	[0.8, 17.2]	[0.0, 18.8]	1.3	1.3	1.3
GFDL-ESM2M	[0.5, 18.5]	[0.1, 18.8]	[2.8, 18.8]	[6.8, 12.6]	[6.8, 12.6]
GISS-E2-H	[0.0, 18.8]	[0.3, 18.8]	[0.0, 18.8]	3.8	3.8
GISS-E2-R	[1.4, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[7.9, 15.2]	[7.9, 15.2]
HadGEM2-AO	[1.0, 14.9]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[1.0, 1.9]
HadGEM2-CC	[0.0, 18.5]	[0.0, 18.8]	NI	NI	NI
HadGEM2-ES	[0.8, 18.8]	[0.0, 18.8]	18.5	18.5	18.5
IPSL-CM5A-LR	[0.1, 18.8]	[0.0, 18.8]	6.5	6.5	6.5
IPSL-CM5A-MR	[0.8, 18.8]	[0.0, 18.8]	[0.3, 18.8]	18.5	18.5
MIROC-ESM	[8.1, 18.8]	[0.0, 16.5]	[0.5, 18.5]	[7.2, 8.8]	[7.2, 8.8]
MIROC-ESM-CHEM	[0.5, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[7.9, 14.0]
MPI-ESM-LR	[2.5, 18.2]	[0.0, 18.8]	[0.0, 18.8]	[7.9, 14.0]	[7.9, 14.0]
MPI-ESM-MR	[0.0, 17.2]	[0.0, 18.8]	[0.0, 18.5]	[0.1, 5.1]	[0.1, 5.1]
MRI-CGCM3	[0.5, 18.8]	[0.0, 18.8]	[0.3, 18.8]	[0.5, 14.0]	[0.5, 14.0]
NorESM1-M	[0.5, 18.5]	[0.1, 18.8]	[0.0, 18.8]	[0.4, 18.5]	[0.4, 18.5]
bcc-csm1-1	[1.4, 16.6]	[0.0, 18.8]	[1.1, 15.2]	[11.0, 11.8]	[11.0, 11.8]
bcc-csm1-1-m	[0.0, 16.6]	[0.0, 18.8]	[0.0, 17.1]	0.3	0.3
inmcm4	[0.5, 18.8]	[0.1, 18.8]	[0.0, 18.8]	18.2	18.2
MIROC5	[2.4, 16.5]	[0.1, 18.8]	[0.0, 18.8]	[11.6, 14.0]	[11.6, 14.0]
CSIRO-Mk3-6-0	[3.7, 18.5]	[0.3, 18.8]	[0.0, 18.5]	16.0	16.0
CMCC-CM	[0.3, 18.5]	[0.1, 18.8]	[0.0, 18.8]	[2.4, 18.5]	[2.4, 18.5]
CanESM2	[0.5, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[8.0, 12.6]	[8.0, 12.6]

NI: no investment

Table 2.3: The range of scaled buying latitude under RCP8.5 and CI68 in the 2080s

GCM \ Conservatism	0	0.05	0.2	0.5	1
ACCESS1-0	[4.9, 17.2]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[12.6, 15.2]
ACCESS1-3	[0.8, 16.5]	[2.8, 18.5]	[0.0, 18.8]	14.0	14.0
CCSM4	[0.8, 18.5]	[0.0, 18.8]	NI	NI	NI
CESM1-BGC	[6.5, 18.8]	NI	NI	NI	NI
CESM1-CAM5	[2.0, 18.5]	[0.0, 18.2]	[0.0, 18.8]	NI	NI
CMCC-CMS	[9.6, 18.5]	[1.0, 18.8]	NI	NI	NI
CNRM-CM5	[2.6, 18.8]	[0.0, 18.8]	[0.0, 18.8]	8.0	8.0
EC-EARTH	[0.7, 18.2]	[0.0, 18.8]	[0.0, 18.8]	18.5	18.5
FGOALS-g2	[2.2, 18.5]	[0.0, 17.2]	[0.0, 18.1]	NI	NI
GFDL-CM3	[0.3, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[12.6, 13.0]
GFDL-ESM2G	[2.6, 18.8]	[1.0, 17.2]	11.8	11.8	11.8
GFDL-ESM2M	[1.8, 18.2]	[5.4, 18.8]	[7.2, 8.0]	[7.2, 8.0]	[7.2, 8.0]
GISS-E2-H	[2.5, 18.5]	[0.0, 18.8]	[0.1, 18.8]	NI	NI
GISS-E2-R	[4.1, 18.2]	[0.0, 18.8]	NI	NI	NI
HadGEM2-AO	[1.0, 18.2]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
HadGEM2-CC	[3.3, 18.2]	[1.6, 18.8]	NI	NI	NI
HadGEM2-ES	[4.8, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
IPSL-CM5A-LR	[1.8, 18.5]	[0.0, 18.8]	[0.0, 18.8]	16.7	16.7
IPSL-CM5A-MR	[4.2, 18.5]	NI	NI	NI	NI
MIROC-ESM	[0.5, 16.5]	[0.0, 18.8]	NI	NI	NI
MIROC-ESM-CHEM	[4.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	17.5
MPI-ESM-LR	[1.9, 18.2]	[0.0, 15.8]	[0.0, 18.8]	NI	NI
MPI-ESM-MR	[1.7, 18.5]	[0.1, 18.8]	[0.0, 18.8]	NI	NI
MRI-CGCM3	[2.8, 18.2]	[0.1, 18.8]	[0.0, 18.8]	5.4	5.4
NorESM1-M	[6.6, 18.5]	[0.3, 18.8]	[0.1, 18.8]	NI	NI
bcc-csm1-1	[0.5, 16.6]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
bcc-csm1-1-m	[0.3, 18.8]	[0.0, 18.2]	[0.0, 18.8]	0.5	0.5
inmcm4	[0.5, 17.2]	[0.4, 18.5]	[0.0, 18.8]	[12.3, 13.0]	[12.3, 13.0]
MIROC5	[10.0, 18.5]	[0.3, 18.8]	NI	NI	NI
CSIRO-Mk3-6-0	[2.3, 11.3]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI
CMCC-CM	[0.7, 18.2]	[0.0, 18.8]	NI	NI	NI
CanESM2	[5.2, 18.8]	[0.9, 16.1]	[0.0, 18.8]	NI	NI

NI: no investment

Table 2.4: The range of scaled buying latitude under RCP4.5 and CI95 in the 2080s

GCM \ Conservatism	0	0.05	0.2	0.5	1
ACCESS1-0	[1.8, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
ACCESS1-3	[4.8, 16.0]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
CCSM4	[1.3, 15.3]	[0.0, 18.5]	[0.0, 18.2]	NI	NI
CESM1-BGC	[0.3, 18.5]	[0.0, 18.8]	[0.8, 18.8]	18.5	18.5
CESM1-CAM5	[9.3, 18.8]	[7.8, 18.8]	[10.4, 18.8]	NI	NI
CMCC-CMS	[3.3, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
CNRM-CM5	[1.4, 18.5]	[0.0, 18.8]	[0.0, 18.8]	8.0	8.0
EC-EARTH	[3.3, 18.8]	[0.0, 18.5]	[0.0, 13.0]	NI	NI
FGOALS-g2	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	14.0	14.0
GFDL-CM3	[0.5, 18.1]	[0.0, 14.6]	[0.0, 16.0]	[0.0, 16.0]	9.3
GFDL-ESM2G	[0.8, 17.2]	[0.0, 18.8]	[0.5, 18.8]	1.3	1.3
GFDL-ESM2M	[0.5, 18.5]	[0.5, 18.8]	[0.0, 18.8]	[7.2, 7.3]	[7.2, 7.3]
GISS-E2-H	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
GISS-E2-R	[1.4, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
HadGEM2-AO	[1.0, 14.9]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
HadGEM2-CC	[0.0, 18.5]	[0.0, 18.8]	NI	NI	NI
HadGEM2-ES	[0.8, 18.8]	[0.0, 18.8]	NI	NI	NI
IPSL-CM5A-LR	[0.1, 18.8]	[0.3, 18.8]	[0.0, 18.8]	NI	NI
IPSL-CM5A-MR	[0.8, 18.8]	[0.0, 18.8]	[0.0, 18.5]	NI	NI
MIROC-ESM	[8.1, 18.8]	[0.4, 14.4]	[0.7, 15.8]	NI	NI
MIROC-ESM-CHEM	[0.5, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI
MPI-ESM-LR	[2.5, 18.2]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
MPI-ESM-MR	[0.0, 17.2]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
MRI-CGCM3	[0.5, 18.8]	[0.0, 18.8]	[6.1, 15.6]	NI	NI
NorESM1-M	[0.5, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
bcc-csm1-1	[1.4, 16.6]	[1.2, 18.8]	[1.2, 18.8]	[11.8, 12.6]	[11.8, 12.6]
bcc-csm1-1-m	[0.0, 16.6]	[0.0, 18.8]	0.3	0.3	0.3
inmcm4	[0.5, 18.8]	[0.0, 18.8]	[0.5, 17.2]	NI	NI
MIROC5	[2.4, 16.5]	[0.0, 18.8]	[0.0, 13.0]	NI	NI
CSIRO-Mk3-6-0	[3.7, 18.5]	[0.3, 18.8]	[0.8, 18.1]	NI	NI
CMCC-CM	[0.3, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
CanESM2	[0.5, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI	NI

NI: no investment

Table 2.5: The range of scaled buying latitude under RCP8.5 and CI95 in the 2080s

GCM \ Conservatism	0	0.05	0.2	0.5	1
ACCESS1-0	[4.9, 17.2]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI
ACCESS1-3	[0.8, 16.5]	[0.8, 18.5]	NI	NI	NI
CCSM4	[0.8, 18.5]	[0.0, 16.6]	NI	NI	NI
CESM1-BGC	[6.5, 18.8]	[0.5, 16.6]	NI	NI	NI
CESM1-CAM5	[2.0, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
CMCC-CMS	[9.6, 18.5]	[1.0, 18.2]	NI	NI	NI
CNRM-CM5	[2.6, 18.8]	[0.0, 18.8]	[0.0, 17.5]	[0.0, 17.5]	NI
EC-EARTH	[0.7, 18.2]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
FGOALS-g2	[2.2, 18.5]	[0.0, 17.2]	NI	NI	NI
GFDL-CM3	[0.3, 18.8]	[0.0, 18.8]	NI	NI	NI
GFDL-ESM2G	[2.6, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
GFDL-ESM2M	[1.8, 18.2]	[4.0, 18.8]	[0.5, 18.8]	8.0	8.0
GISS-E2-H	[2.5, 18.5]	[1.0, 18.8]	[0.0, 18.8]	NI	NI
GISS-E2-R	[4.1, 18.2]	[0.0, 17.2]	NI	NI	NI
HadGEM2-AO	[1.0, 18.2]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
HadGEM2-CC	[3.3, 18.2]	[5.1, 18.8]	NI	NI	NI
HadGEM2-ES	[4.8, 18.5]	NI	NI	NI	NI
IPSL-CM5A-LR	[1.8, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
IPSL-CM5A-MR	[4.2, 18.5]	[0.0, 18.8]	NI	NI	NI
MIROC-ESM	[0.5, 16.5]	NI	NI	NI	NI
MIROC-ESM-CHEM	[4.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	14.0
MPI-ESM-LR	[1.9, 18.2]	[0.0, 16.1]	[0.0, 18.8]	NI	NI
MPI-ESM-MR	[1.7, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
MRI-CGCM3	[2.8, 18.2]	[0.0, 18.8]	[0.0, 18.5]	NI	NI
NorESM1-M	[6.6, 18.5]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
bcc-csm1-1	[0.5, 16.6]	[0.0, 18.8]	[0.0, 18.8]	NI	NI
bcc-csm1-1-m	[0.3, 18.8]	[0.0, 18.8]	[0.0, 14.7]	NI	NI
inmcm4	[0.5, 17.2]	[0.4, 18.5]	[6.1, 17.2]	NI	NI
MIROC5	[10.0, 18.5]	[0.3, 18.8]	[0.0, 18.1]	NI	NI
CSIRO-Mk3-6-0	[2.3, 11.3]	[0.0, 18.8]	[0.0, 18.8]	[0.0, 18.8]	NI
CMCC-CM	[0.7, 18.2]	[0.0, 18.8]	NI	NI	NI
CanESM2	[5.2, 18.8]	[0.0, 15.5]	NI	NI	NI

NI: no investment

Chapter 3

Robust Optimization of Battery Energy Storage System Arbitrage: An Application to Climate Adaptation in the MISO

This chapter investigates the operation of a battery energy storage system (BESS) in a North American deregulated electricity market, adapting to climate change. We propose a climate downscaling method to downscale climate variables to the desired resolution for electricity price prediction. We introduce a robust optimization model to determine the BESS's operating strategy over an 80-year period (from 2021 to 2100) under various climate projections. We conduct two empirical studies to investigate how climate change

affects electricity prices and the optimal BESS operations in the context of arbitrage within a U.S. deregulated electricity market.

3.1 Introduction

In 2022, the U.S. generates electricity for more than 4.3 trillion kilowatt-hours (kWh), being the second largest electricity market in the world after China (Energy Information Administration, 2023a). In 2021, the revenue of the electricity industry in the U.S. is about 424 billion USD (Statista, 2023).

Within the U.S., Midcontinent Independent System Operator (MISO) emerged as one of the world's major wholesale electricity markets, facilitating annual transactions valued at $40 billion USD for a population of 45 million (MISO, 2023). MISO, which operates across all or parts of 15 U.S. states (Alevin NRDCINC1, 2021), is the largest Independent System Operator in the U.S. in terms of geographic coverage. This broad spatial pattern is advantageous for studying the spatial impacts of climate change on electricity markets.

Renewable energy generates about 22% of all U.S. electricity, with 10.2% from wind and 3.4% from solar (Energy Information Administration, 2023b). In the fourth quarter of 2022, renewables constituted approximately 20% of MISO's energy mix, with wind contributing 14% (S&P Global, 2023).

Both the U.S. and MISO are positioned for substantial growth in renewable energy adoption over the next two decades. The U.S. has set ambitious targets, aiming for wind and solar energy to make up 60% to 80% of its electricity generation by 2035 (Office of En-

ergy Efficiency and Renewable Energy , 2022). Similarly, MISO is planning for an energy system with an 80% annual renewable energy penetration rate before 2043, with an emphasis on increasing wind and solar energy generation (Natural Resources Defense Council, 2023). However, the inherent variability and unpredictability of renewable energy sources, such as solar and wind power, pose significant challenges in managing these intermittent energy supplies, leading to potential instability in electricity markets (Chandramowli and Felder, 2014).

Energy Storage Systems (ESSs) play a crucial role in effectively managing the intermittency from renewable energy by shifting energy supply. ESSs function by accumulating and storing surplus energy during periods of renewable energy abundance and subsequently discharging this stored energy into the grid when demand necessitates it. These functionalities, complemented by ancillary services such as voltage stabilization and frequency regulation, substantively enhance grid stability and efficiency of electricity energy management (Kumar and Palanisamy, 2020; Rana et al., 2023).

ESSs have three primary categories: thermal energy storage systems, mechanical energy storage systems, and battery energy storage systems (BESSs) (Koohi-Fayegh and Rosen, 2020).

1. Thermal energy storage systems store excess input energy as thermal energy in a storage medium, like water, aquifer, and rock, and then release the stored energy to generate electricity.

2. Mechanical energy storage systems, which include pumped hydro power and flywheels, capture input energy in the form of gravitational potential or kinetic energy.

3. BESSs, particularly lithium-ion batteries, store external energy as chemical energy, offering a readily convertible source of electricity.

When compared to the other two storage system categories, BESSs offer distinct advantages, including their ability to provide short-term control functions such as load leveling (which reduces electricity demand fluctuations within seconds) and load following (adjusting to electricity fluctuations throughout the day) (Poullikkas, 2013).

BESSs also have the benefit of engaging in minutely or hourly energy arbitrage for revenue generation as they can effectively switch charge/discharge operations within less than a minute (Cheng and Powell, 2016). This energy arbitrage is one of the main applications of ESS (Peñaranda et al., 2021). The arbitrage entails the strategic accumulation of energy during off-peak periods, with subsequent discharge and sale during peak energy demand periods (Santos et al., 2023).

BESSs play a central role in renewable energy storage markets, with approximately 65% of time-shift projects for renewable energy utilizing BESSs (Rahman et al., 2020). Additionally, the market for BESS is set for continuous growth, while BESS prices are anticipated to decrease. For example, the lithium-ion battery market, a prominent subset of BESSs, is projected to expand from $22 billion in 2019 to $116 billion in 2030, while the price of lithium-ion battery packs dropped from 1,100 $/kWh in 2010 to 156 $/kWh in 2019 (Hannan et al., 2021). Considering the critical role BESSs play in the current and future renewable energy storage markets, this chapter primarily focuses on the operation and implications of BESSs.

Climate change will have multiple impacts on energy markets, including BESSs, by

changing demand and supply patterns. These implications include indirectly increasing renewable penetration and changing the dynamics of the supply from renewable sources (Gernaat et al., 2021), such as when and where it is windy or cloudy. Since electricity demand and supply strongly influence electricity prices, particularly in deregulated electricity markets, like MSIO, climate change will substantially influence pricing dynamics (Mulder and Scholtens, 2013). These implications on electricity demand, supply, and price dynamics will affect the operations and revenues of energy markets, including those related to BESS arbitrage and ancillary services (McConnell et al., 2015; MacDonald et al., 2012). Stakeholders including BESS operators will need to adapt to climate change challenges.

However, adapting BESS operation strategies to climate change is challenging. These strategies rely on electricity price predictions in the context of climate change. Nevertheless, there is currently a lack of a comprehensive understanding of the relationship between electricity prices and climate variables. Predicting how electricity values will evolve over time in response to climate change poses a substantial challenge. Consequently, making appropriate BESS operational decisions to adapt to uncertain climate change remains a difficult task, especially given the long timeline of climate change

This chapter aims to understand the role BESSs will play in deregulated North American electricity markets facing climate change scenarios. Specifically, under climate change, what will be opportunity for BESSs to be profitable and how certain is the profit? To make progress toward this broad question, we need to build a model of how the operations of BESSs are likely to behave under climate change. We will study income-maximizing arbitrage decisions of a BESS in the MISO electricity market under future climate change scenarios. The study provides lower bounds for the income of the BESS since the BESS

can generate additional income from other functions, such as ancillary services. We hope this study will provide insights into incentives for investments and policies for BESSs' development.

To study BESS operations under climate change, we need to address two main problems. The first is to determine electricity prices as a function of climate change projections. The second is to determine operational decisions over these electricity price trajectories. These methodological contributions will facilitate the derivation of results linking climate forecasts to operation decision trajectories. Understanding the variation in these decisions under different GCMs and SSPs will provide insights into the challenges of adapting to climate change.

Prior research in this vein has been done linking climate change to electricity prices. In such studies, electricity prices are estimated in regression models as price ranges mostly across a short time period (several days and months). Another area of research has examined BESS operations under uncertain electricity valuations using stochastic optimization approaches to investigate BESS operation policies over several weeks or months (Cheng and Powell, 2016; Zhou et al., 2019). However, stochastic optimization approaches are challenging to implement in characterizing the distributions of climate and price uncertainty, which can be difficult to fully describe. Moreover, such algorithms are typically intractable over the long trajectories required to study the impact of climate change (Shapiro et al., 2021).

Given the challenges in studying BESS operations and the benefits of using robust optimization under climate change, we introduce a new robust BESS operation model to

examine operations under climate change. We show that the model has a linear equivalent form, which can be efficiently solved using standard optimization techniques. With the proposed model, we conduct empirical studies in MISO to investigate the optimal BESS operations and operating incomes over an 80-year period under 108 potential climate scenarios.

To support our 80-year-long period study, we assume that the electricity system will remain unchanged during this period. While increased renewable penetration can lead to reductions in electricity prices for the current electricity mix, the variance in electricity prices may still exhibit consistency with climate patterns, regardless of changes in the system. This is because electricity supply, demand, and price variations are more significantly influenced by climate variability than by changes in the magnitude of electricity prices (Hill et al., 2021).

The specific contributions of this chapter are as follows:

1. We develop a robust model, with linear equivalences, for studying the BESS's arbitrage operations under uncertainty.

2. We develop a regression model to predict electricity price changes based on long-term climate forecasts. Using the data derived from this approach, we conduct a numerical study of the operation decisions for a BESS in MISO over an 80-year planning horizon, considering a wide range of forecasted climate scenarios and spatial areas. The main findings include:

   (a) Climate change may reduce electricity price variance but increase operating incomes through arbitrage in Northern MISO.

(b) Under climate change scenarios, North-central MISO exhibits higher operating incomes for BESS than Southern MISO.

This chapter is organized as follows. In Section 3.2, we review the literature related to our research. We then briefly describe our methodology in Section 3.3. In Section 3.4, we conduct an empirical study to predict electricity prices within MISO under climate change scenarios. In Section 3.5, we do an empirical study on BESS operations in MISO under climate change scenarios. Section 3.6 concludes this chapter.

3.2 Literature Review

There has been a strong interest within the academic community in predicting electricity prices and modelling the operations of BESSs under climate change. We discuss some of the most important relevant literature from two interrelated streams: 1) predictive models in predicting electricity prices and 2) BESS operation models in the face of uncertainty.

3.2.1 Predictive models in predicting electricity prices

In our study, we employ a regression model for predicting electricity prices under the influence of climate change. Although various methods exist for electricity price prediction, including machine learning (ML) models (Fan et al., 2007; Yang et al., 2022; Tschora et al., 2022), time series (TS) models (Yang et al., 2017), and regression models (Mulder and Scholtens, 2013; Zhou et al., 2019), the choice of a regression model corresponds to our specific research needs, as discussed below.

ML models work by learning information directly from data without relying on a predetermined equation as a model. However, ML models commonly focus on point forecasts. Ideally, electricity price prediction should include uncertainty estimates because price uncertainty has the highest impact on market participants (Ziel and Steinert, 2018).

TS models represent dynamic systems tailored to fit sequential data, but a single TS model may not perform well in predicting multi-periodic and non-stationary data, especially over long-term horizons. Electricity prices exhibit multiple periodicities (e.g., hourly, weekly, monthly, and yearly) and are highly dynamic, with sudden and unpredictable changes, making them non-stationary (Weron, 2014). Researchers have explored hybrid TS models that combine other approaches, such as ML models (Yang et al., 2017). However, these hybrid models may be constrained by high computational complexity or limited prediction horizons, typically spanning only several days or months (Yang et al., 2017; González et al., 2017). In contrast, our research necessitates the prediction of electricity prices at hourly intervals over an 80-year horizon.

Regression models, a statistical technique, establish relationships between dependent variables (e.g., climate variables) and independent variables (e.g., electricity prices). These models offer a straightforward structure that facilitates the inclusion of multidimensional climate inputs. They have been commonly used for linking climate variables to electricity prices in European markets (Mulder and Scholtens, 2013) and U.S. electricity markets (Zhou et al., 2019). Additionally, the uncertainty of electricity prices can be simulated through the use of confidence intervals within the regression model.

Instead of modeling electricity supply and demand separately to generate electricity

prices, we directly map climate variables to electricity prices. The rationale behind this mapping is that both electricity demand and supply are influenced by climate variables. Since electricity prices are determined by both supply and demand, they are essentially affected by climate variables (Mulder and Scholtens, 2013).

3.2.2 BESS operation model in the face of uncertainty

Several studies have explored BESS operation models under various forms of uncertainty. Cheng and Powell (2016) utilize a Markov Decision Process (MDP) approach to analyze the optimal operation of a battery storage system for energy arbitrage and frequency regulation in a electricity bus under electricity demand and price uncertainty. However, their model's complexity, driven by a large state space and planning periods, may not be suitable for long-term hourly planning horizons, such as the 8,760 periods (i.e., the number of hours in a year) considered in our study. Similarly, Zhou et al. (2019) apply an MDP model to investigate operating policies for a storage system integrated with a wind farm, considering uncertain wind speeds and electricity prices. Again, their models faced challenges due to the expansive state space, particularly in the context of long-term planning. Liu et al. (2022) study the optimal policy of energy storage and a wind plant under wind generation uncertainty. While they focused on qualitative policy, our study provides quantitative results on when and how much to charge/discharge power into/out of a BESS. Unlike the three mentioned research approaches, our robust model offers linear equivalence and scalability, making it well-suited for addressing long-term issues stemming from climate change. Furthermore, our model considers the intricate impacts of electricity grids within

an electricity market by incorporating climate variables for (all or part of) 15 U.S. states within the MISO market, as shown in Figure 3.1.

Some studies have focused on deterministic models for assessing the size and operation of energy storage systems. Peñaranda et al. (2021) adopt mixed-integer linear programming (MILP) to study BESS's optimal operation. Arbabzadeh et al. (2019) use a deterministic linear model to study the optimal size and operation of an energy storage system. Nguyen et al. (2017) utilize a deterministic model to study the maximal revenue of energy storage system from participating in arbitrage and frequency regulation in MISO. In our study, we use a robust model to account for electricity price uncertainty induced by climate change, making it more effective in generating adaptive decisions under climate change. The robust model also provides a lower bound on the viability of a BESS in the face of climate change.

Another stream of work centers on the profitability of energy storage systems. Although Zafirakis et al. (2016) do not give the optimal operations of energy storage system, they pointed out the financial value of arbitrage for energy storage. Additionally, Sioshansi et al. (2009) highlight that the inherent value of energy storage can vary significantly with the energy market structure. Our research complements this body of work by presenting BESS operating incomes, which exhibit variations across climate scenarios and geographic areas.

3.3 Robust Optimization for Sequential Operations

In this section, we introduce a robust BESS model to study the optimal operation of a BESS under climate change in two steps. First, we propose a multistage stochastic programming model of the operation problem. Second, we develop a robust version of this model. In service of improving tractability, we show that the robust model has linear equivalences under certain conditions.

3.3.1 Stochastic Programming for BESS Operations

We study the operation model of BESS arbitrage by an operator at an MISO bus node. It's important to note that in the context of an electricity market, a bus represents a specific location or node within the network where electrical power is generated, consumed, or transmitted (Von Meier, 2006). The model studies hourly charge and discharge decisions for the BESS in a bus where electricity prices vary idiosyncratically over the planning horizon. Additionally, the operator needs to consider electricity price changes caused by climate variables in their decisions. These price changes are uncertain because of the range of climate scenarios. The objective of the model is to investigate hourly BESS charge and discharge decisions that maximize the total discounted expected operating incomes over the planning horizon. These operating incomes only consider the charging costs and discharge incomes, without including other costs or incomes, such as those related to ancillary services.

We introduce the model settings for the BESS in an MISO bus under climate change,

including the planning horizon, BESS specifications, electricity prices, decision variables, constraints, state variables, and the objective.

Planning horizon: The operator makes arbitrage decisions periodically over a finite planning horizon $\mathcal{T} := \{1, 2, ..., T\}$. Period t ($t \in \mathcal{T}$) is defined as the time interval $[t, t+1)$. In this chapter, each period equals one hour. The operation decisions of the BESS occur at the beginning of each period in $\mathcal{T}$.

Battery energy storage system: We consider a BESS that provides only energy arbitrage in each period t ($t \in \mathcal{T}$). We assume that the BESS has finite energy capacity and power capacity. We also assume that the BESS is near the bus so that there is no transmission loss between the BESS and the bus. In each hour, the operator decides to either charge or discharge energy. If they decide to charge (respectively (resp.) discharge) energy, then the BESS will be charged (resp. discharged) for one hour. For the BESS, we use the following parameters:

- C^S (MWh): Energy capacity of the BESS; $C^S > 0$.
- C^C, C^D (MWh): Charging, discharging (energy) capacity in an hour (the amount of energy that can be charged or discharged in an hour); $C^C, C^D > 0$
- η^C, η^D: Charging, discharging efficiency of the BESS; η^C, $\eta^D \in (0, 1)$
- δ: One-period risk-free discount rate; $\rho \in (0, 1]$.
- x_t (MWh): BESS storage level by the end of period t ($t \in \mathcal{T}$); $x_t > 0$.

Despite the potential significance of the self-discharge rate of the BESS, we chose to exclude it from our model based on several factors. Firstly, the self-discharge rate of

lithium batteries, which are commonly used in BESS, is relatively low, typically ranging from 2% to 5% per month (Liao et al., 2022). Our model takes into account the discharge and charge frequency at an hourly resolution, during which the self-discharge amount is minimal. Additionally, the self-discharge rate is influenced by environmental temperature and storage level, rendering it a non-linear function (Liao et al., 2022). Accurately modeling the self-discharge rate in this context poses significant challenges. Given the complexity involved in modeling the self-discharge rate and the focus of our model on hourly short-term charging and discharging frequencies, we exclude self-discharge in our model.

Moreover, our model does not incorporate battery deficiency, which refers to a decrease in battery capacity over time. Instead, we assume that the battery capacity remains constant throughout the planning horizon of the model. This decision is motivated by our emphasis on the operating income aspect of the BESS, prioritizing financial gains from charging and discharging operations over the costs associated with battery capacity degradation over time.

Electricity prices: We define a random vector P_t to model spot energy market prices in the bus at period t ($t \in \mathcal{T}$) in \$/MWh. These prices remain stable for each period (hour). Hourly-stable prices like these are common in deregulated electricity markets, such as the day-ahead electricity market (ISO New England, 2023), and will be discussed in detail in Section 3.4.1. We model the electricity prices over the planning horizon as a time series $\{P_t\}_{t\in\mathcal{T}}$. These prices are not identically distributed, and their distributions depend on the specific climate scenario. We let p_t ($t \in \mathcal{T}$) be the observation of P_t at the beginning of period t.

Decision variables: At the beginning of each period $t \in \mathcal{T}$, the operator observes the storage level x_t and electricity price p_t and decides how much energy is charged to the BESS or discharged from the BESS. At each period (hour) t ($t \in \mathcal{T}$), we denote the amount of electrical energy (MWh) charged to the battery and discharged from the battery as q_t^C and q_t^D, respectively, where q_t^C, $q_t^D \geq 0$. We disable illegal wash trades of the BESS by setting $q_t^C q_t^D = 0$. We assume that the operator is a price taker and their arbitrage decisions do not affect electricity prices. At the end of each period $t \in \mathcal{T}$, the BESS's storage level becomes $x_{t+1} = x_t + \eta^C q_t^C - q_t^D$. The decision variables should satisfy energy capacity limitation, charging capacity limitation, and discharging capacity limitation: $0 \leq x_t + \eta^C q_t^C - q_t^D \leq C^S$, $0 \leq q_t^C \leq C^C$, and $0 \leq q_t^D \leq C^D$ ($\forall t \in T$).

State variables: Let state S_t be (x_t, p_t), where x_t is an element of its feasible set space $\mathcal{X}$, and p_t is an element of its feasible set space $\mathcal{P}$. The given initial state is $S_0 := (x_0, p_0)$. The state space for each period is defined as $\mathcal{S} := \mathcal{X} \times \mathcal{P}$, where $\times$ denotes the Cartesian product of the two sets.

Operating income function: We define an operating income function that maps (S_t, q_t^C, q_t^D) to an operating income. The operating income earned during period $t \in \mathcal{T}$ can be expressed as:

$$R(S_t, q_t^C, q_t^D) = p_t(\eta^D q_t^D - q_t^C) \tag{3.1}$$

, where $0 \leq x_t + \eta^C q_t^C - q_t^D \leq C^S$ and $q_t^C q_t^D = 0$.

Operator's objective: We let a feasible policy π be a sequence of decision rules $(A_t^\pi(S_t))_{t \in \mathcal{T}}$ that outputs a feasible decision (q_t^C, q_t^D) given the state S_t in period t. The operator's goal is to identify a feasible policy within the feasible policy space Π that maximizes the sum-

mation of two factors: (1) the total discounted expected operating incomes and (2) the expected value of remaining power at the end of period T:

$$\text{Max}_{\pi\in\Pi} \sum_{t\in\mathcal{T}} \delta^t \mathbb{E}_{p_t\sim P_t}[R(S_t^\pi, A_t^\pi(S_t^\pi))|S_0] + \delta^T \mathbb{E}_{p_T\sim P_T}[p_T S_T^\pi] \tag{3.2}$$

, where $\mathbb{E}_{p_t\sim P_t}[\cdot]$ $(t \in \mathcal{T})$ denotes the expectation over random price P_t. Note that the state S_t^π is dependent on the policy π and p_t is one of the inputs of the state S_t^π.

While Model (3.2) can be addressed through the Backward Algorithm (Shapiro et al., 2021), solving this state-based stochastic programming for BESS poses significant challenges due to its high computational complexity and extensive state size, stemming from accounting for the long-term effects of climate change. Furthermore, obtaining an accurate distribution of P_t $(t \in \mathcal{T})$, which is related to the distribution of climate projections, during the model-solving process is a non-trivial task. The distribution may be not trivial and not standard, adding to the complexity of solving the model.

To enhance the tractability of Model (3.2), we introduce a robust variant of this model for resolution. There are two primary reasons why the robust version is effective in our case. Firstly, the robust version aids in identifying decisions that perform well across a range of uncertain future climate scenarios (Bhave et al., 2016). By employing these decisions derived from the robust version, the BESS operator can gain confidence in the potential profitability of their operation decisions and develop a broader understanding of the risks they will face (Constantino and Weber, 2021; Hallegatte et al., 2012). Secondly, the robust model features linear equivalences that can be resolved directly and efficiently using standard optimization tools.

3.3.2 Robust Model

To help solve Model (3.2), we convert the BESS model into its robust version in two steps. First, we model uncertain electricity prices as independent intervals over time. Second, based on these intervals, we introduce the robust version of the model.

Modeling electricity price uncertainty: In our context, we model the electricity price P_t ($\forall t \in \mathcal{T}$) in the bus as a symmetric, independent, and bounded random variable, which falls within the interval $[\bar{p}_t - \hat{p}_t, \bar{p}_t + \hat{p}_t]$, where $\bar{p}_t$ represents the mean value of P_t, and $\hat{p}_t$ represents the maximum potential deviation of P_t from its mean value. Note that we do not require $\bar{p}_t - \hat{p}_t > 0$ as electricity prices can be negative (Zhou et al., 2019).

Robust model: We develop a robust investment model in which the operator anticipates electricity prices to fall within the interval $[\bar{p}_t - \hat{p}_t, \bar{p}_t + \hat{p}_t]$ ($t \in \mathcal{T}$). The operator adopts a pessimistic stance and adversely selects electricity prices based on whether they are charging (considering high prices within the interval) or discharging (considering low prices within the interval). The operator's objective is to maximize operating incomes of the BESS under the worst variance of electricity prices, i.e., charging at a high price and discharging at a low price. The operating incomes are the summation of two factors: (1) the total discounted operating incomes and (2) the discounted value of the remaining power at the end of period T, both under the worst variance of electricity prices. Within this pessimistic framework, we can describe the robust version of BESS Model (3.2) as follows, denoted as

Model (3.3):

$$\text{Max}_{\{q_t^C, q_t^D : \forall t \in \mathcal{T}\}} \quad \sum_{t \in \mathcal{T}} \delta^t [(\bar{p}_t - \hat{p}_t)\eta^D q_t^D - (\bar{p}_t + \hat{p}_t) q_t^C] + \delta^T (\bar{p}_T - \hat{p}_T)[x_0 + \sum_{t \in \mathcal{T}} (\eta^C q_t^C - q_t^D)] \tag{3.3a}$$

$$\text{s.t.} \quad 0 \leq x_0 + \sum_{t=0}^{k} (\eta^C q_t^C - q_t^D) \leq C^S \quad \forall k \in \mathcal{T} \tag{3.3b}$$

$$q_t^C q_t^D = 0 \quad \forall t \in \mathcal{T} \tag{3.3c}$$

$$0 \leq q_t^C \leq C^C \quad \forall t \in \mathcal{T} \tag{3.3d}$$

$$0 \leq q_t^D \leq C^D \quad \forall t \in \mathcal{T} \tag{3.3e}$$

Constraints (3.3a), (3.3d), and (3.3e) ensure energy capacity limitation, charging capacity limitation, and discharging capacity limitation in each planning period, respectively. Constraints (3.3c) prevent illegal wash trades of the BESS in each planning period.

Solving by linear transform: We address Model (3.3) by demonstrating its linear equivalence, allowing for rapid resolution using standard optimization tools. Due to the presence of non-convex constraints (3.3c), Model (3.3) constitutes a non-convex optimization problem. Consequently, solving Model (3.3) using standard optimization tools, algorithms, or heuristics may not generate the (global) optimal solution. To overcome this challenge, we show that constraints (3.3c) are inactive and can be eliminated from Model (3.3), as demonstrated in the following Lemma 2.

Lemma 2 *If* $\bar{p}_t + \hat{p}_t > \eta_C \eta_D (\bar{p}_t - \hat{p}_t)$ $(\forall t \in \mathcal{T})$, *without constraints* (3.3c), *the optimal charge and discharge profiles* $\{q_t^{C\star}, q_t^{D\star}, \forall t \in \mathcal{T}\}$ *of Model* (3.3) *satisfy that at least one of*

q_t^C, q_t^D is 0 at any time t. In other words, constraints (3.3c) are redundant given $\bar{p}_t + \hat{p}_t > (\bar{p}_t - \hat{p}_t)\eta_C\eta_D$ $(\forall t \in \mathcal{T})$.

We present the proof of Lemma 2 in Appendix B.1.

Lemma 2 indicates that, at any given period $t \in \mathcal{T}$, if the upper bound of the electricity price (i.e., $\bar{p}_t + \hat{p}_t$) is greater than the product of the lower bound of the electricity price and both the charge and discharge efficiency of the BESS (i.e., $\eta_C\eta_D(\bar{p}_t - \hat{p}_t)$), charging and discharging the BESS in the same period will result in an operating-income loss for the BESS.

This loss occurs due to two reasons: 1) both charging and discharging result in electricity loss, or, in other words, financial loss, as the charging and discharging efficiency (i.e., η_C and η_D) is less than 0, and 2) charging prices are not lower than discharging prices, and charging at high prices and discharge at low prices may result in a financial loss. As the objective function (3.3a) maximizes its value, the optimal solution will choose to either charge or discharge (or neither) in each period.

If the upper bound of the electricity price $\bar{p}_t + \hat{p}_t$ is consistently positive for all $t \in \mathcal{T}$, then the inequality $\bar{p}_t + \hat{p}_t > \eta_C\eta_D(\bar{p}_t - \hat{p}_t)$ holds for all $t \in \mathcal{T}$. This implies that constraints (3.3c) are redundant when the upper bound of the electricity price is always positive for all $t \in \mathcal{T}$ by Lemma 2. Therefore, Model (3.3) is a linear programming model that can be solved in polynomial time (Karmarkar, 1984) given $\bar{p}_t + \hat{p}_t > 0$ $(\forall t \in \mathcal{T})$.

Furthermore, if any of the upper bounds of the electricity prices is negative, a minor adjustment can be applied by rounding it to 0, transforming Model (3.3) into a linear model.

3.4 Study I: Day-ahead LMP in MISO under Climate Change Scenarios

In this section, we predict future electricity prices, which offers price bounds for the robust BESS Model (3.3), under different climate scenarios. Specifically, we predict hourly electricity prices for selected (bus) nodes in MISO from 2021 up to 2100. There are five main steps to do the scenario-based case study. Firstly, we offer an introduction to the MISO market, providing essential context and insights about our research background, as detailed in Section 3.4.1. Second, we introduce a regression model to link electricity prices and climate scenarios in section 3.4.2. Third, we present the data sources and data processing procedures utilized for fitting the regression model in Section 3.4.3. We proceed to fit the regression model and predict the day-ahead LMPs within MISO in Section 3.4.4 Finally, we conduct a thorough analysis of the prediction outcomes in Section 3.4.5.

3.4.1 Midcontinent Independent System Operator

Midcontinent Independent System Operator (MISO) — an Independent System Operator and Regional Transmission Organization, Midcontinent Independent System — is responsible for the operation of the deregulated wholesale electricity markets in the Midwest U.S. and Manitoba, Canada. In this research, we only focus on the U.S. part of MISO. Figure 3.1 shows the map of the MISO that covers all or part of 15 U.S. states: Arkansas (AR), Illinois (IL), Indiana (IN), Iowa (IA), Kentucky (KY), Louisiana (LA), Michigan (MI), Minnesota (MN), Mississippi (MS), Missouri (MO), Montana (MT), North Dakota (ND),

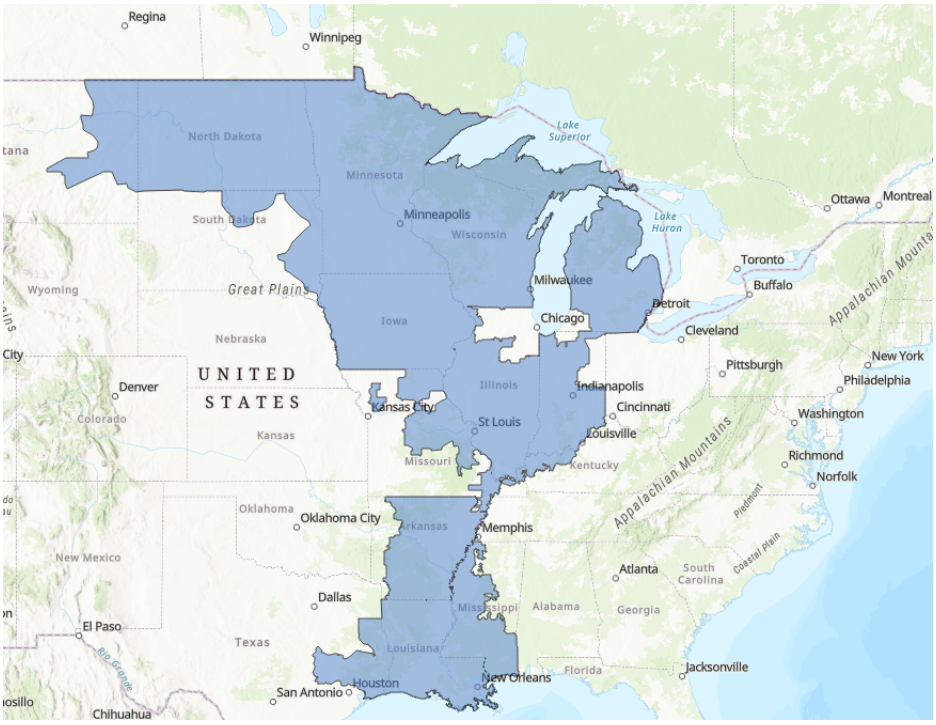


Figure 3.1: The map of MISO (retrieved from Alevin NRDCINC1 (2021))

South Dakota (SD), Texas (TX), and Wisconsin (WI).

MISO operates through a sequence of markets: the day-ahead market enables strategic planning for the efficient utilization of generating units over various hours, while the real-time market is designed to respond swiftly to unpredictable changes in demand (Midcontinent, 2016). Specifically, first, the day-ahead market establishes prices and plans hourly production and delivery for the subsequent operating day. Second, a real-time market accepts generation bids 30 minutes before operation and employs them to resolve last-minute demand changes every five minutes.

Within the deregulated wholesale electricity markets, such as MISO, electricity prices

vary due to constraints imposed by limitations in grid transportation capacity. To account for these variations in value, the MISO employs a pricing methodology known as locational marginal pricing (LMP) (Birge et al., 2018). Under the LMP framework, the price attributed to each specific node or location corresponds to the marginal cost associated with supplying an additional MWh of electricity at that node.

MISO provides comprehensive data encompassing two fundamental pricing metrics: hourly day-ahead LMP data relevant to the day-ahead market and real-time LMP data pertaining to the real-time market. In the context of our research, we rely on day-ahead hourly LMP data. This preference is motivated by the greater predictability associated with day-ahead LMPs (Birge et al., 2018), aligning seamlessly with the extensive temporal scope of our long-term study.

MISO provides a favorable environment for the implementation of battery electricity storage systems, as evidenced by the opening of its energy storage market with over 13 gigawatts (GW) in interconnection queue (Walton, 2022). The operator's board of directors of MISO also approved a $10.3 billion transmission plan to support electricity battery for storing renewable energy (Walton, 2022). Given these favorable conditions, we chose MISO as the primary focus for our case study on BESS operations.

In our case, we assume that the BESS functions as a price taker within the MISO market. This assumption is predicated on the BESS's arbitrage capacity of 10 MWh per hour, as used in our case study. This capacity represents less than 0.01% of the total hourly electricity demand in MISO, which ranges between 118.2 GW and 125.2 GW (Brown, 2022). Consequently, the hourly arbitrage amount of the BESS constitutes only a small fraction of

the total electricity demand per hour, and as a result, arbitrage decisions have a negligible impact on the electricity market and electricity price.

3.4.2 Regression Model

To obtain the day-ahead LMP bounds as parameters for the robust Model (3.3), we employ an electricity price model to estimate the trajectories of the bounds from 2021 to 2100. References (Zhou et al., 2019; Mulder and Scholtens, 2013) point out that electricity prices, such as day-ahead LMPs, are complex time series with mean reversion, spikes, seasonality, climate-condition sensitivity, and negative values. In light of these two references, we form an electricity price model (i.e., a regression model) as a combination of an inverse hyperbolic sine function (to handle positive-negative values) and a seasonality function with climate variables (to fit seasonality and climate impacts). We have opted not to include an autoregression (AR) model in our electricity price model to account for mean reversion, because an AR model may not significantly improve the electricity price model. Furthermore, we have not incorporated a framework for addressing electricity price spikes, as our focus is on the overall change in electricity prices over an extended (decades-long) time horizon.

We introduce the main components of the electricity price model: electricity prices, inverse hyperbolic sine transform, and seasonality function with climate variables.

Electricity prices: We denote day-ahead LMPs as hourly-level time series $\{P_t\}_{t\in\mathcal{T}}$ during the planning horizon $\mathcal{T}$.

Inverse hyperbolic sine transform: The electricity prices are commonly transferred to

obtain more stable variances before model fitting (Contreras et al., 2003). We use the inverse hyperbolic sine function to transform electricity prices because this transform is analogous to the natural logarithm but can model negative prices (Zhou et al., 2019). Let $\sinh^{-1}(x) = log(x + \sqrt{x^2 + 1})$ be the inverse hyperbolic sine function. We denote this transformed electricity prices as $\{\sinh^{-1}(P_t)\}_{t\in\mathcal{T}}$. We can get electricity prices P_t from this transformed electricity prices by calculating $\sinh(\sinh^{-1}(P_t))$, where $\sinh(x) := (e^x - e^{-x})/2$ is the hyperbolic sine function.

Seasonality function with climate variables: This function includes four types of climate variables that affect electricity prices and four types of time dummy variables. As a node's local climate and climate in other areas among the MISO may both affect the node's LMPs, we consider climate data in the whole MISO to study the LMPs of a MISO's node. Specifically, we use the climate data of all counties in MISO. Let a be a constant and $\mathcal{C}$ be the collection of counties covering the whole MISO. In period $t \in \mathcal{T}$, for each county $c \in \mathcal{C}$, we employ four types of hourly climate variables including (2-meter) air temperature (denoted as $x_{1,t,c}$), precipitation ($x_{2,t,c}$), (10-meter) wind speeds ($x_{3,t,c}$), and surface downwelling shortwave radiation ($x_{4,t,c}$). We choose these climate variables because temperature is the main driver of electricity demand (Fan et al., 2019), precipitation impacts both electricity demand and supply (Mideksa and Kallbekken, 2010), and wind speed and shortwave radiation influence renewable energy generation from wind and solar sources, respectively (Carta et al., 2009; The Office of Energy Efficiency and Renewable Energy, 2023).

We let $\beta_{i,c}$ ($\forall c \in \mathcal{C}$ and $i = 1, 2, 3, 4$) be $x_{i,t,c}$'s coefficient. Let $I_y(t)$, $I_m(t)$, $I_d(t)$, and $I_h(t)$ be the four types of dummy variables that equal one if period t is in year y, month

m, weekday d, and hour h, respectively, and zero otherwise. Note that we fit electricity price data from year Y_1 to year Y_2 ($Y_2 > Y_1$), so we use $|Y_2 - Y_1|$ dummy variables $I_y(t)$ to distinguish these $|Y_2 - Y_1 + 1|$ years. Similarly, we employ 11 dummy variables for months, six for weekdays, and 23 for hours. Let β_y, β_m, β_d, and β_h be the coefficients for these four types of dummy variables, respectively. Then, we model the seasonality function with climate variables as:

$$f(t) = a + \sum_{c \in \mathcal{C}} \sum_{i=1}^{4} \beta_{i,c} x_{i,t,c} + \sum_{y=Y_1}^{Y_2 - Y_1} \beta_y I_y(t) + \sum_{m=1}^{11} \beta_m I_m(t) + \sum_{d=1}^{6} \beta_d I_d(t) + \sum_{h=1}^{23} \beta_h I_h(t) \quad (3.4)$$

We express the electricity price model as follows:

$$sinh^{-1}(P_t) = f(t) + \epsilon_t \quad (3.5)$$

, where we assume that ϵ_t ($t \in \mathcal{T}$) are independent and identical random errors.

In Figure 3.2, we show a framework to solve regression Model (3.5), predict future day-ahead LMPs, and use the predicted values to solve robust Model (3.3) under climate change scenarios. The upper part (Section 3.4) of Figure 3.2 shows the framework of the data sources and data processing (in Section 3.4.3) for solving Model (3.5), predicting future day-ahead LMPs (in Section 3.4.4), and day-ahead LMP results (in Section 3.4.5).

In Section 3.4.3, we first collect historical hourly electricity price data and historical hourly weather data to solve regression Model (3.5). Second, as future climate data is at a daily level, we propose a downscaling method to convert these daily data into hourly resolutions. We conduct the downscaling for each climate scenario, generating 108 hourly

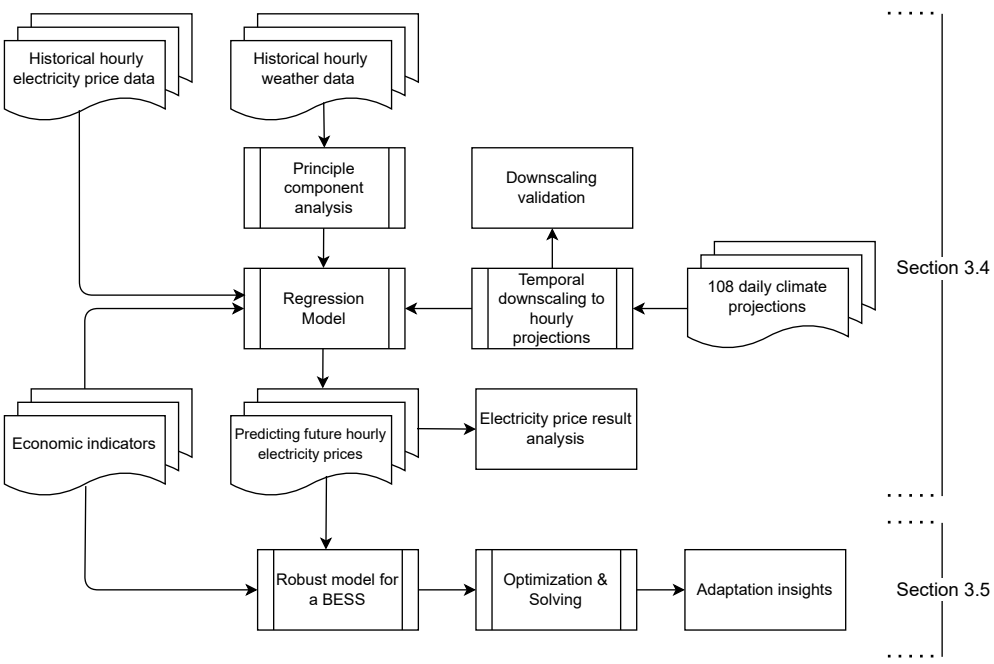


Figure 3.2: Methodology framework of Chapter 3

climate trajectories from 2021 to 2100. We also validate the downscaling method by measuring the downscaling errors. Third, we analyze the results of the downscaled hourly climate data.

In Section 3.4.4, we first use the Principal Component Analysis (PCA) method to address multicollinearity issues among historical hourly climate data. Second, we predict future electricity prices and provide an example of the resulting predicted electricity prices. Note that the PCA is also applied to the (downscaled) future hourly climate projections before predicting electricity prices.

In Section 3.4.5, we present the results of electricity price predictions.

3.4.3 Data Sources and Data Processing

To estimate the parameters of Model (3.5) and subsequently generate forecasts of LMP trajectories from 2021 to 2100, we need to acquire historical day-ahead LMP data, historical climate data, and future climate projections data, all of which should be available at an hourly level. Firstly, we introduce available data sources for Model (3.5). Secondly, as future climate projections are only available at daily level, we downscale them to hourly level. Finally, we present some key results about downscaled climate.

Data source and data prepossessing

We introduce the day-ahead LMPs and climate datasets for Model (3.5). While historical day-ahead LMP and historical climate data are available at the required level, hourly climate projections for the future are unavailable. Instead, we introduce available daily climate projections for the future, which can be downscaled to the required hourly level.

Historical day-ahead LMP data: We derive the hourly day-ahead LMPs in 2016, 2017, 2018, and 2019 from the MISO Real-Time and Market Data source (MISO Energy, 2023). We do not use day-ahead LMPs in 2020, as the year is affected by pandemics and presents different demand/supply patterns for day-ahead LMPs compared to 2016-2019 (Santiago et al., 2021). These four-year day-ahead LMP values represent the hourly price of electricity at specific nodes within MISO system and are measured in \$/MWh. Specifically, we focused on gathering day-ahead LMP data for generation nodes located in Zones 2, 4, 8, and 9 of MISO, as detailed in Table 3.5 and illustrated in Figure 3.12. We select these four Zones because they span vertically from north to south, allowing us to assess the impact

of climate variations across different latitudes. We drop the nodes within the four Zones without complete hourly data of all the four years, i.e., 2016, 2017, 2018, and 2019, resulting in a final dataset of 358 nodes.

Due to the enormous number of selected nodes (i.e., 358 nodes) in MISO, executing our electricity price model and BESS operation model for each individual node is time-consuming. To simplify the process for our research, we only consider nodes that are not highly correlated with other nodes within the MISO network. Nodes exhibiting correlation values ranging from 0.7 to 1.0 or -0.7 to -1.0 are considered highly correlated (Mukaka, 2012).

Specifically, first, we calculate the similarity between the (day-ahead) LMP time series from January 1, 2019, 00 : 00 : 00, to December 31, 2019, 23 : 00 : 00 for each pair of the 358 nodes using the cosine similarity. Given two n-dimensional vectors, $\mathbf{A}$ and $\mathbf{B}$, the cosine similarity between the two vectors is defined as $\mathbf{A} \cdot \mathbf{B} / (||\mathbf{A}||\ ||\mathbf{B}||)$, where $\cdot$ represents the dot product and $||\cdot||$ is the Euclidean norm. Second, we arrange the nodes by descending order of the sum of their cosine similarity values with other nodes. Third, for each node in this order, we calculate its correlation with other nodes and exclude the node displaying a correlation exceeding 0.7 or falling below -0.7 with other nodes. Consequently, we select a total of 13 nodes from Zone 2, Zone 4, Zone 8, and Zone 9, which exhibit not-high correlation: ALTE.SHEEPSIN1, WPS.CRANECREK, and WPS.CUSTEGEN1 in Zone 2; AMIL.HEN_AMT_P, AMIL.TRA_ATRAE, MEC.NEALS_4, AMIL.STWF and in Zone 4; EAI.ANO2, EAI.INDEPEND2, EAI.INDEPEND1, and EAI.AECCHYDRO2 in Zone 8; LEPA.HOUMA_G16 and LEPA.MURRAY in Zone 9 (see Table 3.5).

To incorporate regression Model (3.5), we adjust the historical hourly day-ahead LMPs from 2016 to 2019 to the USD value of the last hour of 2019. We carry out this adjustment using an annual inflation rate of 0.02, which is set as the long-term inflation goal by the Federal Open Market Committee (Board of Governors of the Federal Reserve System, 2021). Assuming that each hour experiences the same inflation rate, we can calculate the inflation rate for each hour as approximately 2.3×10^{-6} using the formula $(1.02)^{1/(365\times24)} \approx (1.02)^{1/(366\times24)} \approx 1 + 2.3 \times 10^{-6}$, where 365×24 (respectively 366×24) is the total number of hours in a non-leap (respectively leap) year.

Hourly historical climate data: We obtain the hourly historical climate data from the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) dataset (Copernicus Climate Change Service, 2023) for the required four types of climate variables. The ERA5 dataset offers hourly global climate and weather data ranging from 1940 to the present day. The dataset is gridded to a regular latitude-longitude grid of 0.25 degrees. The ERA5 directly offers the values of the three (out of four) required climate variables 2-meter air temperature, total precipitation, and surface shortwave radiation (or called solar radiation) downward. The ERA5 dataset does not provide the 10-meter (10m) wind speed ($w10$) data directly. Rather, the dataset offers the 10m v-component (or called northward component) of wind ($u10$) and 10m u-component (or called eastward component) of wind ($v10$) data. To obtain the $w10$ data, we calculate the magnitude of the $u10$ wind and $v10$ using the formula: $w10 = [(u10)^2 + (v10)^2]^{1/2}$ (Smith et al., 1999). Table 3.1 presents the details about these five climate variables. For our MISO case, we use the five climate variables in Table 3.1 from 2016 to 2019 for the latitude-longitude grids covering the Continental United States from the ERA5 dataset.

Table 3.1: Historical hourly climate variables in the ERA5 data

Variable	**symbol (unit)**	**Notes**
1. 2-meter (air) temperature	*t2m* (K)	K = °C + 273.15
2. Total precipitation (m/hour)	*tp* (m/hour)	-
3. 10m u-component of wind	*u10* (m/s)	using the magnitude of the
4. 10m v-component of wind	*v10* (m/s)	*u10* and *v10* to get *w10*
5. Surface solar radiation downwards	*ssrd* (J/m^2)	$J/m^2 = 3{,}600 \times W/m^2$ (only for this data)*

*: Copernicus Climate Change Service (2023)

Hourly climate projections for future (not available): Future hourly climate projections from 2021 to 2100 for the four required climate variables are not available for the MISO area. Although the North American Coordinated Regional Climate Downscaling experiment (Mearns et al., 2017) and the CMIP6 do provide hourly air temperature and precipitation projections for the North America, wind speed and shortwave radiation projections at an hourly level are not available.

The most related and complete climate projection data for those four climate variables are at the daily level. In fact, we can downscale daily climate projections onto hourly climate projections using some popular temporal-downscaling methods. The nearest neighbour method (Lee and Singh, 2018) is widely used to downscale climate projections to finer resolutions. We use this method to downscale daily climate projections to hourly levels. Before conducting the downscaling process, we first introduce the available daily climate projections used to generate hourly climate projections for future as follows.

Daily climate projections for future: The climate projections are derived from the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6) (Thrasher

et al., 2022). NEX-GDDP-CMIP6 assumes that the relative spatial patterns observed from the reference period (1960-2014) will remain constant under future climate change (Thrasher et al., 2022). There are about 140 climate projections generated by 35 GCMs, with each projection parameterized based on four distinct Sustainable Development Scenarios (SSP): SSP126, SSP245, SSP370, and SSP585. These SSPs describe global socioeconomic changes and greenhouse gas emission trajectories up to 2100. Progressing from SSP126 to SSP585, the socio-economic development trajectory relies more on fossil fuels, consequently leading to more greenhouse gas emissions.

The NEX-GDDP-CMIP6 GCMs provides some bias-corrected daily climate variables, including near-surface air temperature, precipitation, wind speed, or surface downwelling shortwave radiation, at 0.25-degree horizontal resolution from 2015 to 2100. Table 3.2 shows the four variables from the NEX-GDDP-CMIP6. We only utilize the GCMs that provide data for all four required climate variables across all four SSPs, resulting in 27 GCMs. Further details about the 27 GCMs can be found in Table 3.3. The daily data simulated by the 27 GCMs are used to develop hourly measures of climate variables covering MISO during the 2021-2100 period.

Table 3.2: Daily climate variables in the NEX-GDDP-CMIP6 climate change projection

Variable	symbol (unit)	Notes
1. (2-meter) Near-surface air temperature	*tas* (K)	-
2. Precipitation	*pr* (kg/m^2/s)	1kg/m^2/s = 3.6m/hour
3. (10-meter) Surface wind speed	*sfcWind* (m/s)	-
4. Surface downwelling shortwave radiation	*rsds* (W/m^2)	*rsds* and *ssrd* are often used interchangeably (Stephens et al., 2012; AMS, 2012)

Table 3.3: A list of NEX-GDDP-CMIP6 GCMs used in the analysis and their reporting institutions

Number	Model Name	Institution
1	ACCESS-CM2	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology
2	ACCESS-ESM1-5	
3	BCC-CSM2-MR	Beijing Climate Center
4	CanESM5	Canadian Centre for Climate Modelling and Analysis
5	CESM2	Community Earth System Model Contributors
6	CMCC-CM2-SR5	Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici
7	CMCC-ESM2	
8	CNRM-CM6-1	Centre National de Recherches Météorologiques–Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
9	CNRM-ESM2-1	
10	EC-Earth3	EC–EARTH consortium
11	EC-Earth3-Veg-LR	
12	FGOALS-g3	Chinese Academy of Sciences
13	GFDL-ESM4	NOAA Geophysical Fluid Dynamics Laboratory
14	GISS-E2-1-G	Goddard Institute for Space Studies
15	INM-CM4-8	Institute for Numerical Mathematics
16	INM-CM5-0	
17	IPSL-CM6A-LR	L′Institut Pierre–Simon Laplace
18	KACE-1-0-G	National Institute of Meteorological Sciences /Korea Meteorological Administration
19	MIROC-ES2L	Japan Agency for Marine–Earth Science and Technology, Atmosphere and Ocean Research Institute, The University of Tokyo, National Institute for Environmental Studies, and RIKEN Center for Computational Science
20	MIROC6	
21	MPI-ESM1-2-HR	Max Planck Institute for Meteorology
22	MPI-ESM1-2-LR	
23	MRI-ESM2-0	Meteorological Research Institute
24	NorESM2-LM	Norwegian Climate Centre
25	NorESM2-MM	
26	TaiESM1	Research Center for Environmental Changes, Academia Sinica
27	UKESM1-0-LL	Met Office Hadley Centre

Sources: Thrasher et al. (2022); WDC Climate (2023)

Note that to facilitate a comparison of climate variables between the ERA and NEX-GDDP-CMIP6 datasets, we standardize the measurement units and adopt consistent symbols, i.e., *w10*, *t2m*, *tp*, and *ssrd*, for both datasets. Additionally, we convert both datasets' time zones to the Eastern Standard Time, aligning them with the time zone used by the historical day-ahead LMP data.

Climate temporal downscaling: the nearest neighbor

In this part, we downscale the daily climate data simulated by the 27 GCMs to an hourly level. The k-nearest neighbors method (k-NN) is a statistical technique that is commonly used to downscale daily climate projections to an hourly resolution (Lee and Singh, 2018). The main assumption of the method is that if two days have similar daily mean climate values, they are considered to have similar climate conditions at hourly intervals.

In our study, we specifically utilize the 1-nearest neighbor (1-NN) approach. It is important to note that our approach does not depend on considering multiple neighbors to generate uncertainty in climate projections. Instead, we incorporate multiple GCMs and SSPs to account for the uncertainty in climate projections. Additionally, to ensure continuity of climate between consecutive hours in downscaling, we will modify the distance measure in the original 1-NN method.

1-NN method: The original 1-NN method involves the following steps, which are illustrated here for a county c in the MISO region on a future day under one of the combinations of GCMs and SSPs:

1. Establishment of sample pools: From a GCM under a SSP of NEX-GDDP-CMIP6

datasets, we collect daily (mean) climate values for the four climate variables, denoted as $\mathcal{I} := \{w10, t2m, tp, ssrd\}$, for the future days between 2021 to 2100 (although some days are in the past, we still call them future days in this research). Let $y_{i,d_f,c}$ be the value of the daily climate variable i ($i \in \mathcal{I}$) at county $c \in \mathcal{C}$ on a future day d_f in $\mathcal{D}_f$, where $\mathcal{D}_f$ is the collection of days between 2021 to 2100.

From the ERA5 dataset, we collect hourly observations $x_{i,d_o,c} = \{x_{i,d_o,c,h}\}_{h\in\{0,1,...,23\}}$ for climate variable i ($i \in \mathcal{I}$) at county $c \in \mathcal{C}$ at the h-th ($h \in \{0, 1, ..., 23\}$) hour on a historical day $d_o \in \mathcal{D}_o$, where $\mathcal{D}_o$ is the collection of days between 1991 to 2020.

2. Estimation of distance: we estimate the distance between the daily climate projection $y_{i,d_f,c}$ and daily observations $\bar{y}_{i,d_o,c_o} := \frac{1}{24}\sum_{h=0}^{23} x_{i,d_o,c_o,h}$ using the Euclidean distance, given by $D_{d_o,c_o} = \sum_{i\in\mathcal{I}}(y_{i,d_f,c} - \bar{y}_{i,d_o,c_o})^2$ for each $d_o \in \mathcal{D}_o$ and $c_o \in \mathcal{C}$.

3. Estimation of the nearest sample: we arrange the estimated distances D_{d_o,c_o} ($d_o \in \mathcal{D}_o$ and $c_o \in \mathcal{C}_o$) from step (2) in ascending order and reserve the indices (d_o and c_o) of the smallest distance. Denote the reserved indices as d^* and c^*.

4. Estimation of hourly projection: we estimate the downscaled hourly projection $\hat{x}_{i,d_f,c} = \{\hat{x}_{i,d_f,c,h}\}_{h\in\{0,1,...,23\}}$ for $y_{i,d_f,c}$, where each hourly observation $x_{i,d^*,c^*,h}$ is assigned to $\hat{x}_{i,d_f,c,h}$ for each $i \in \mathcal{I}$ and $h \in \{0, 1, ..., 23\}$.

Modifying the distance measure in the 1-NN: In climate downscaling, ensuring continuity between consecutive hours is essential. Specifically, it is important to ensure that the climate conditions in the last hour of a given day are consistent with the climate conditions in the first hour of the subsequent day. To achieve this, we modify the distance measure

D_{d_o,c_o} in the step 2 (of the *1-NN method*) by including the time-based distance between the climate values of the last hour of a day and the climate values of the first hour of the following day. This adjustment is applied to 10-meter wind speed (*w10*) and 2-meter air temperature (*t2m*). However, for solar surface radiation downward (*ssrd*) and precipitation (*tp*), which exhibit discontinuities over hourly intervals, we do not incorporate them in the time-based distance. This is because solar surface radiation downward is zero during nighttime hours and precipitation can vary greatly from one hour to the next, including cases where it may rain in one hour and not rain in the subsequent hour. The revised distance measure becomes $D_{d_o,c_o} = \sum_{i\in\mathcal{I}}(y_{i,d_f,c} - \bar{y}_{i,d_o,c_o})^2 + \sum_{i\in\setminus\{w10,t2m\}}(x_{i,d_o,c_o,0} - \hat{x}_{i,d_f-1,c,23})^2$. Additionally, we initialize $\hat{x}_{i,d_f-1,c,23}$ with the values of the climate variables from the last hour of the last historical day (i.e., 23:00:00 December 31, 2020). This ensures a consistent starting point for the downscaling process.

To ensure that a future rainy (respectively non-rainy) day is matched with a historical rainy (respectively non-rainy) day, we exclusively select rainy (respectively non-rainy) days from the historical data as the source for the future rainy (respectively non-rainy) day in the modified 1-NN method. This selection process helps to maintain the consistency of precipitation patterns in the nearest neighbor matching for future climate projections.

Additionally, to maintain consistency between future and historical daylight lengths, we select days from the historical dataset that correspond to the same season as the future days. In defining the seasons, we adhere to the Meteorological Seasons convention, wherein Spring begins on March 1, Summer begins on June 1, Autumn begins on September 1, and Winter begins on December 1.

Downscaling process of the modified 1-NN: The downscaling process of the modified 1-NN for climate projections involves two key steps: data normalization and the establishment of criteria to evaluate the quality of the downscaling results. We perform the data normalization process to bring the climate variables involved in the downscaling process to a range of [0,1]. This step ensures that the input data is comparable and eliminates any potential biases or disparities caused by differences in units or magnitudes. Regarding the evaluation criteria, we use the absolute and relative differences, commonly referred to as errors, to quantitatively assess the goodness of the downscaling procedure. We discuss the two steps and provide examples for our downscaling results as follows.

Step 1: data normalization As we conducted our study at a county level within MISO, we spatially downscaled all the climate data from both ERA5 and NEX-GDDP-CMIP6 datasets, which are represented in grid format, to 883 counties within MISO. This downscaling was accomplished using the *remapbil* operator of the Climate Data Operators (CDO) (v1.9.9rc1; Schulzweida (2022)). To ensure uniformity, we normalized all the resulting climate data values to fall within the interval [0, 1]. This was achieved by dividing climate values by the widths of their corresponding validation bounds, as specified in Table 3.4. However, it is important to highlight that hourly *ssrd* values may exceed its daily bounds shown in Table 3.4.

Due to the nature of *ssrd* being zero during the night for 11 hours, and relatively weak during the early morning and early night hours, the range of daily (mean) *ssrd* values becomes smaller compared to the hourly *ssrd* values. To address this issue, we have adjusted the bounds for hourly *ssrd* by multiplying them by 24/11. Here, the value 11 $(= 24 - 11 - 2)$ represents the number of hours with non-negligible *ssrd* within a day.

Consequently, the adjusted bound for hourly *ssrd* are set as [0, 12000/11]. The maximum value of the normalized hourly based on the adjusted bound *ssrd* is 0.98, falling within the range of [0,1].

The histograms presented in Figures 3.3c, 3.3a, 3.3d, and 3.3b illustrate the numerical distribution of the four normalized variables for the historical daily data, which represents the normalized daily mean of hourly ERA5 data. These distributions are depicted with purple bars. The normalized values for wind speed (*w10*) and precipitation (*tp*) are highly skewed and primarily fall within the ranges of [0, 0.2] (see Figure 3.3c) and [0, 0.1] (see Figure 3.3d), respectively, while the values of the remaining two variables temperature (*t2m*) and radiation (*ssrd*) exhibit a much wider spread.

To ensure a balanced consideration of the importance of each variable within the modified 1-NN approach, we need to further process the wind speed (*w10*) and precipitation (*tp*) variables. In line with the recommendation by Fu et al. (2010) in dealing with skewed daily wind and precipitation data, we apply the root (power) transformation to deal with skewed daily precipitation data. The nth-root transformation is defined as $\sqrt[n]{\cdot}$, where we choose n as a positive integer in our study.

Upon analysis, we find that the 4th root ($\sqrt[4]{\cdot}$) of the normalized historical wind speed (*w10*) data offers a more centralized distribution around 0.5 compared to lower roots (e.g., 3rd root) or higher roots (e.g., 5th root), as depicted in Figure 3.3c. Therefore, we select the 4th root transformation for the values of the normalized daily wind speed.

Similarly, for the normalized historical precipitation (*tp*) data, we find that the 10th root of the normalized positive *tp* (i.e., rainy day) data present a median 0.49, which is

closer to the center of the normalization interval [0, 1], when compared to lower roots (e.g., 9th root with a median 0.46) or higher roots (e.g., 11th root with a median 0.53), as shown in Figure 3.3d. Therefore, we choose the 10th root transformation for the values of the normalized daily precipitation data.

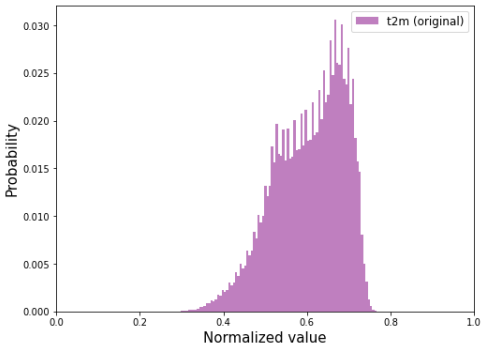
Note that, to accommodate the operational wind range (4-25 m/s) (or [4/50, 25/50] after normalization) recommended for wind turbines (Association, 2023), we truncate the wind speed (*w10*) data after normalization and before the root transformation. Any normalized wind speed values outside the range [4/50, 25/50] are set to zero.

We run the modified 1-NN over the normalized and downscaled climate data for all 883 MISO counties, 27 GCMs, and 4 SSPs. This process results 108 (=27×4) hourly climate trajectories for each MISO county between 2021 and 2100.

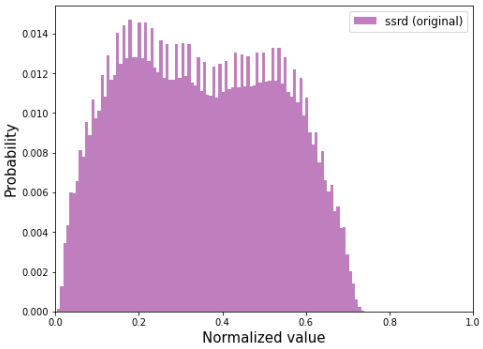
Table 3.4: Bounds for validation of climate data values (retrieved from Thrasher et al. (2022))

Variables	**Lower bound**	**Upper bound**
w10	0 m/s	50 m/s
t2m	200 K	340 K
tp	0 kg/m^2/s	0.012 kg/m^2/s
ssrd	0 W/m^2	500 W/m^2

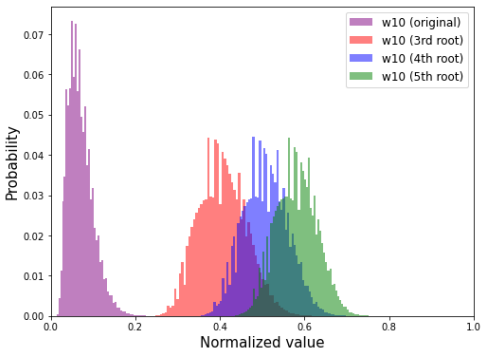
Step 2: criteria To evaluate the accuracy of the climate downscaling, it is necessary to define a measure of goodness for the temporally downscaled (nearest-day) climate. One possible approach, as proposed by Michel et al. (2021), is to compute the absolute and relative differences between the future projections and the temporally downscaled time series at a seasonal level. Let $\bar{y}_{i,d,c}$ denote the value of projected future daily (of GCMs) for the climate variables i ($i \in \mathcal{I}$) on day d in county c ($c \in \mathcal{C}$). Similarly, let $\hat{y}_{i,d,c}$ denote



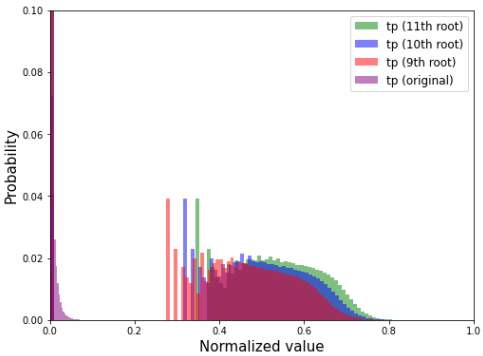
(a) 2-meter air temperature



(b) Surface solar radiation downwards



(c) Wind speed



(d) (Total) precipitation

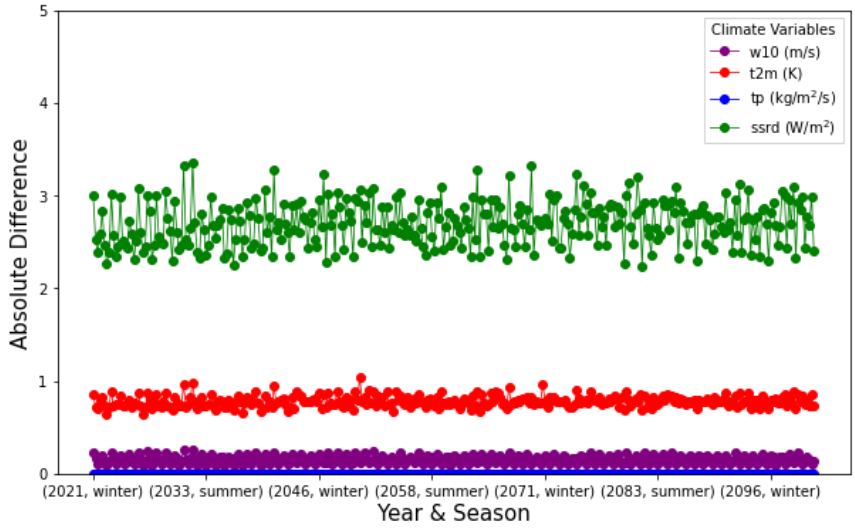
Figure 3.3: The histogram of normalized historical daily climate data

the value of the temporally downscaled future daily for the climate variables i on day d in county c. We calculate the daily absolute difference as $\frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}}|\hat{y}_{i,d,c}-\bar{y}_{i,d,c}|$ and the daily relative difference as $\frac{1}{|\mathcal{C}|}\sum_{c\in\mathcal{C}}|(\hat{y}_{i,d,c}-\bar{y}_{i,d,c})/\bar{y}_{i,d,c}|$, where $|\mathcal{C}|$ equals the number of counties in the county set $|\mathcal{C}|$. To obtain the seasonal absolute (resp. relative) difference for the climate variables i in each year, we sum the daily absolute differences (resp. relative) within that season and year and then divide the summation by the number of days in the season and the year. We calculate the seasonal absolute and relative differences for the four seasons over the period from 2021 to 2100 for the four climate variables.

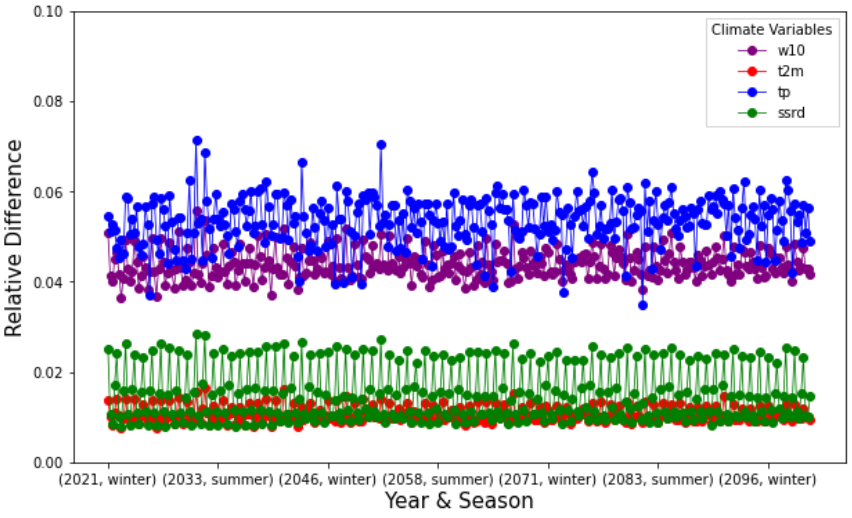
An example: downscaling differences For example, Figure 3.4a illustrates the absolute differences between the temporally downscaled (nearest-day) climate data and the corresponding future climate data obtained from ACCESS-CM under SSP126. The figure presents that the 1-NN method yields absolute differences below 5.2 W/m^2 for *ssrd*, below 1.2 K for *t2m*, below 0.3 m/s for *w10*, and below 0.1 kg/m^2/s for *tp*.

Figure 3.4b presents the relative differences between the temporally downscaled (nearest-day) climate data and the corresponding future climate data obtained from the ACCESS-CM GCM under SSP126. The figure illustrates that the 1-NN method yields relative differences below 0.02 for *t2m*, below 0.04 for *ssrd*, below 0.06 for *w10*, and below 0.1 for *tp*, suggesting that the modified 1-NN method has a low relative bias.

An example: temporal climate continuity in consecutive days For the temporally downscaled climate, we also examined the consistency of climate differences between the initial hour of one day and the final hour of the previous day. Our analysis compared these differences with historical data, demonstrating favorable continuity. Figures 3.5a and 3.5b



(a) The absolute difference



(b) The relative difference

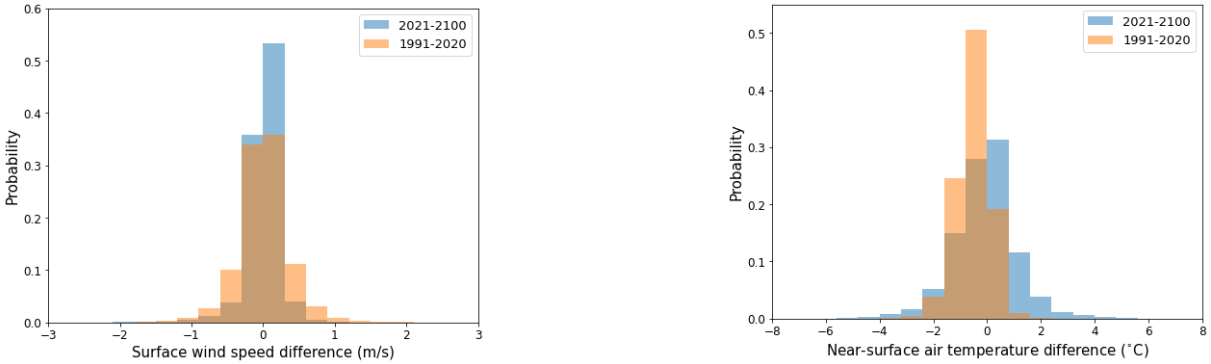
Figure 3.4: Example: the absolute/relative difference of the (modified) 1-NN method in the ACCESS-CM GCM & SSP126 case

present the climate differences between the two time periods (the initial hour of one day and the final hour of the previous day) for *w10* and *t2m*, respectively, under the ACCESS-CM GCM & SSP126. We calculate the differences for each county for both the historical period (1991-2020) and the downscaled future period (2021-2100). Yellow bars represent the distribution of climate differences during the historical period, while blue bars depict the distribution for the downscaled future period under the ACCESS-CM GCM & SSP126. The yellow bars indicate that differences can be positive or negative, such as slightly higher temperatures (or wind speed) at 00:00 on a day compared to 23:00 on the previous day. For *w10*, the blue bars closely matched the yellow bars, indicating strong (temporal) continuity. For *t2m*, the blue bars exhibited a wider range than the yellow bars. However, the probability for the wider parts that are bigger than 2.4 °C or smaller than -4 °C is less than 0.05 and rare.

An example: spatial climate continuity While the modified 1-NN method does not explicitly incorporate spatial climate correlations among counties within MISO, it does exhibit a partial preservation of such correlations. This is because neighboring counties with similar climatic characteristics tend to share similar nearest days during the downscaling process, leading to comparable downscaled climate outcomes. For example, we use the the modified 1-NN method to find the nearest historical day for a future climate projection on May 31, 2100, (2100-May-31) under the ACCESS-CM GCM & SSP126 for each MISO county. Figures 3.6, 3.7, 3.8, and 3.9 illustrate the comparison between the future daily climate and its nearest historical daily climate (or called downscaled daily climate). Overall, the temporal downscaling by the modified 1-NN exhibited favorable spatial consistency for all four climate variables. However, there is a bias observed for the downscaled *t2m* variable

in the southernmost MISO, characterized by a lower value. Similarly, the downscaled *ssrd* variable also demonstrates a bias with lower values in the middle MISO. These observations suggest the presence of potential inaccuracies or limitations in the downscaling process for *t2m* and *ssrd* variables in the southernmost and middle MISO regions, respectively.

In summary, the temporal downscaling method presents satisfactory results in downscaling accuracy (low relative downscaling difference), temporal climate continuity in consecutive days, and spatial climate continuity.



(a) *w10* (b) *t2m*

Figure 3.5: The histograms of climate difference between 00:00 on a day and 23:00 on the previous day under the ACCESS-CM GCM & SSP126

Results of Temporally Downscaled Future Climate

We present key results concerning the temporal downscaling of climate. Specifically, we analyze the evolving standard variances of climate variables, as these variances may have notable implications for the variability in day-ahead LMPs, a critical revenue source in electrical arbitrage.

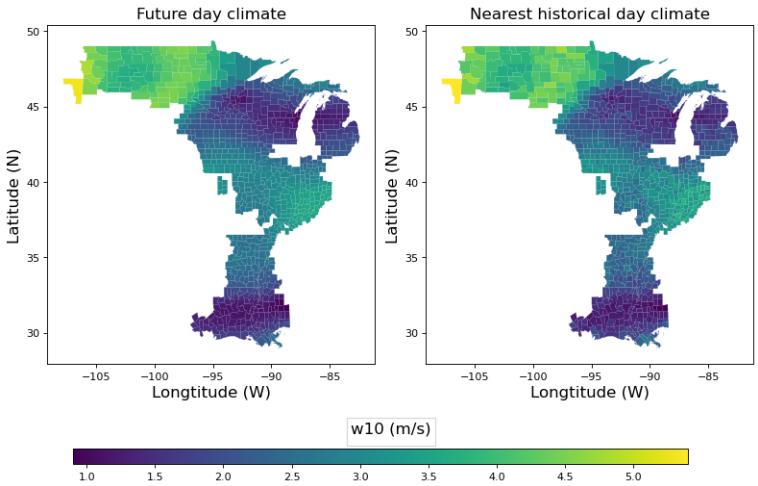


Figure 3.6: Wind speed downscaling on 2100-May-31 under the ACCESS-CM GCM & SSP126

To conduct our analysis, we compute the cumulative sum of annual standard variances over MISO counties using the downscaled county-level daily climate in each year. Note that the daily climate data have been scaled to fit within the interval [0,1] for the purpose of making comparisons between different climate variables. This analysis is performed for each GCM, SSP, climate variable, and year among 2021-2100 periods.

We present the cumulative sum of annual standard variances across the 27 GCMs for each SSP and climate variables in Figure 3.10. In the figure, the solid lines represent the estimated annual means of the cumulative sums derived from the 27 GCMs for each SSP, and the shaded regions represent the ranges of the means for corresponding SSP. While standard variances exhibit annual fluctuations, distinct trends emerge for each climate variable and SSP. Specifically, we observe a decreasing trend in the standard variances of wind over time. In contrast, the standard variances of precipitation tend to increase over the years. Regarding temperature and shortwave solar radiation, they do not have

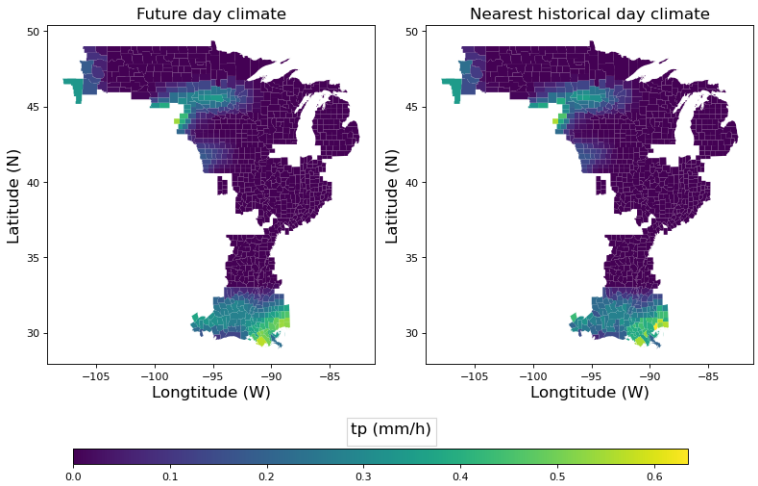


Figure 3.7: Total precipitation downscaling on 2100-May-31 under the ACCESS-CM GCM & SSP126

obvious increasing or decreasing trends. However, as emission scenarios grow from SSP126 to SSP585, there is a noticeable rise in the increasing trends in precipitation over the years. Among the four SSP scenarios, the SSP585 scenario consistently displays the most standard variance increase in precipitation as time advances.

These findings in the cumulative climate-variance changes carry significance for the electricity markets, particularly in the context of wind and precipitation, as higher (lower) variances in these two climate variables have the potential to introduce greater (lower) variability into day-ahead LMPs. Wind variances influence wind supply and, subsequently, electricity prices over time. Precipitation variances impact water supply, which can, in turn, affect electrical energy storage and power plant cooling, ultimately influencing electricity prices.

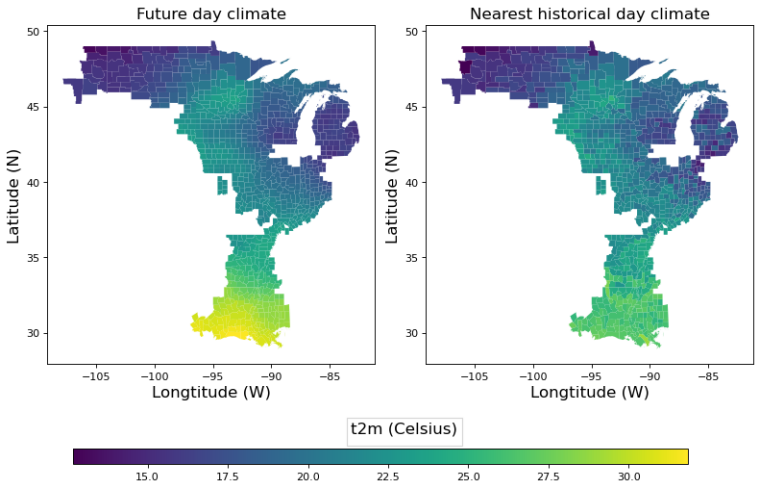


Figure 3.8: 2-meter temperature downscaling on 2100-May-31 under the ACCESS-CM GCM & SSP126

3.4.4 Day-ahead LMP Prediction

After getting the downscaled hourly climate data for the future, we are ready to train the electricity price model and further forecast future day-ahead LMPs. We use the ERA5 data and the MISO day-ahead LMP data to fit regression Model (3.5) for each selected node. However, Model (3.5) may encounter multicollinearity due to the correlation among climate variables. We apply the Principal Component Analysis (PCA) to solve the multicollinearity in Model (3.5). Then, we employ the ordinary squared least method to solve Model (3.5) and generate the best-fitted Model (3.5). After getting the best-fitted model, we input the temporally downscaled hourly climate trajectories to estimate future day-ahead LMPs trajectories. This process generates 108 price trajectories at an hourly resolution for each of the 13 selected MISO nodes, spanning the years from 2021 to 2100. In total, we obtain 1,404 (=108 * 13) such price trajectories, enabling comprehensive analysis of variations in

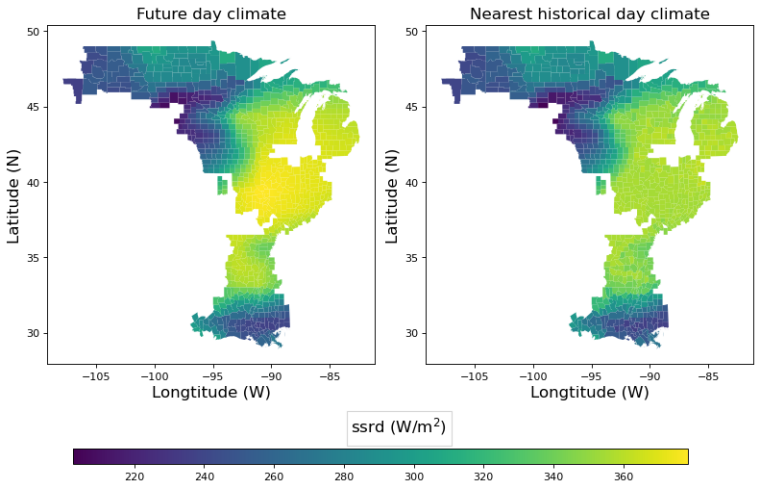
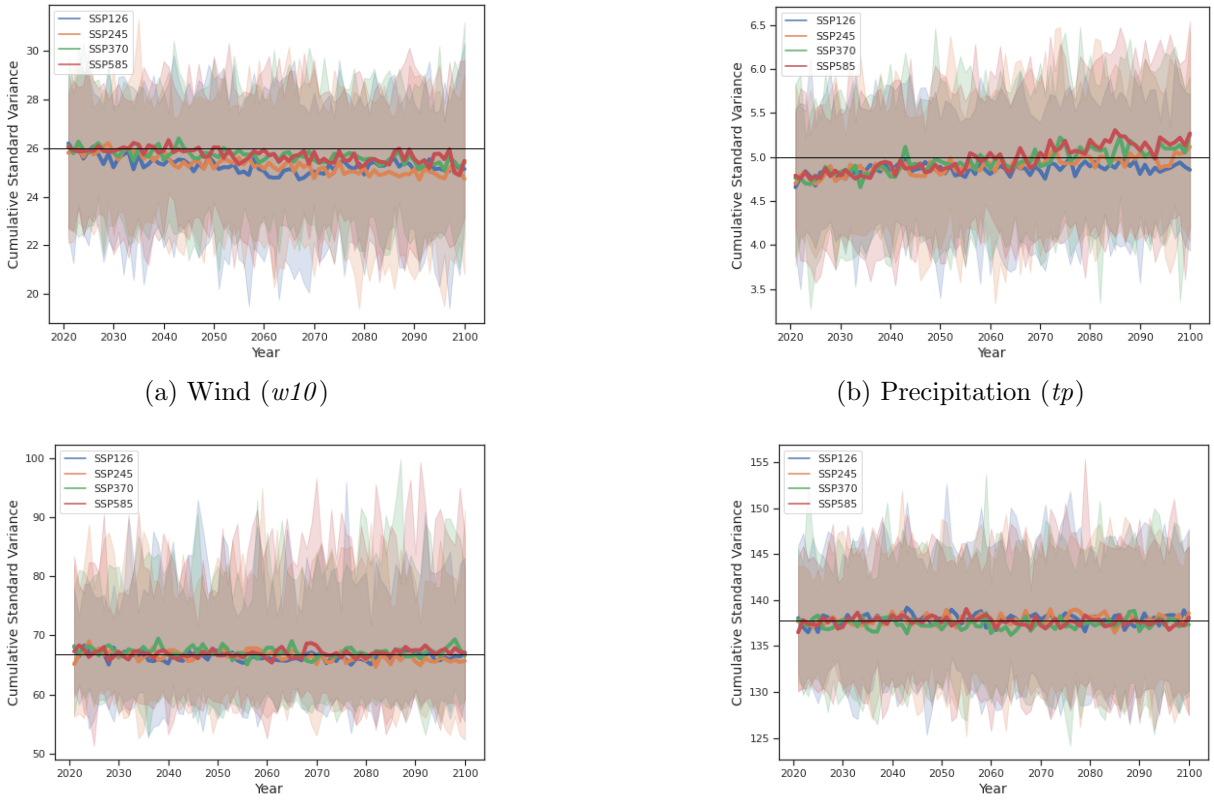


Figure 3.9: Surface downwelling shortwave radiation downscaling on 2100-May-31 under the ACCESS-CM GCM & SSP126

day-ahead LMPs under uncertain climate change.

Multicollinearity: The regression Model (3.5) encounter multicollinearity due to the correlation among climate variables in different counties, as nearby counties have similar climate patterns. To address this issue, we employ the PCA (Jolliffe, 2002), which can transform dependent elements into independent elements that are formulated as linear combinations of the original elements.

The PCA requires a hyperparameter, namely the number of principal components. We determine the the number of principal components by ensuring the retention of a sufficient number of principal components to capture 99% of the total variance in the original dataset. To select the number of principal components, we utilize the PCA function from the Scikit-learn Python package (v1.0.2; Pedregosa et al. (2011)). This package will output a PCA transformation matrix, denoted as $\mathbf{W}$, which can be used to calculate the standard variance



(a) Wind (*w10*)

(b) Precipitation (*tp*)

(c) Temperature (*t2m*)

(d) Shortwave Solar Radiation (*ssrd*)

Figure 3.10: Annual cumulative standard variances of the downscaled climate
The solid lines are the mean of cumulative standard variances across the 27 GCMs; the shaded areas are the ranges of cumulative standard variances across the 27 GCM.

of coefficients in Model (3.5). Further details are provided in Appendix B.2.

The PCA method also requires centralized data as input (Jolliffe, 2002). We centralize the (normalized) historical climate data by subtracting the mean of (normalized) historical climate data across all counties for each climate variable. Similarly, we centralize the (normalized) future climate data by subtracting the same means.

Note that the PCA is exclusively applied to the climate variables and not the dummy

variables. The purpose of PCA is designed to handle continuous variables, as indicated by Pagès (2014). On the other hand, the dummy variables are retained for predicting LMPs due to their time-varying nature.

By applying the PCA to the climate variables, we generate transformed climate variables based on the hyperparameter capturing 99% of the total variance in the original dataset. These transformed climate variables are then combined with time dummy variables and a constant to form the independent variables in regression Model (3.5).

Example of the best-fitted regression model: As an example, Figure 3.11a displays the real day-ahead LMP trajectory (depicted by the green line) and the best-fitted day-ahead LMP trajectory (depicted by the red line) for the ALTE.SHEEPSIN1 Node. The period covered in the figure spans from 2019-06-04 00:00:00 to 2019-06-08 03:00:00 (100 hours). Figure 3.11b presents the forecasted day-ahead LMP for the same node from 2100-07-28 08:00:00 to 2100-08-01 11:00:00 (100 hours) under the ACCESS-CM2 GCM & SSP126. The blue-dashed lines in both figures represent the 95% confidence intervals of the the best-fitted or forecasted day-ahead LMPs.

In the Figure 3.11a, the red lines capture the primary trend of the green line but do not achieve a perfect fit. The cosine similarity of the real day-ahead LMP trajectory (the green line) and the best-fitted day-ahead LMP trajectory (the red line) is 0.966, indicating a strong connection between the real day-ahead LMP trajectory and the best-fitted day-ahead LMP trajectory. The blue-dashed lines encompass the main part of the green line, indicating that using the 95% confidence intervals for predicting the day-ahead LMP is a more reliable approach than relying solely on the best-fitted values (represented by the red

line).

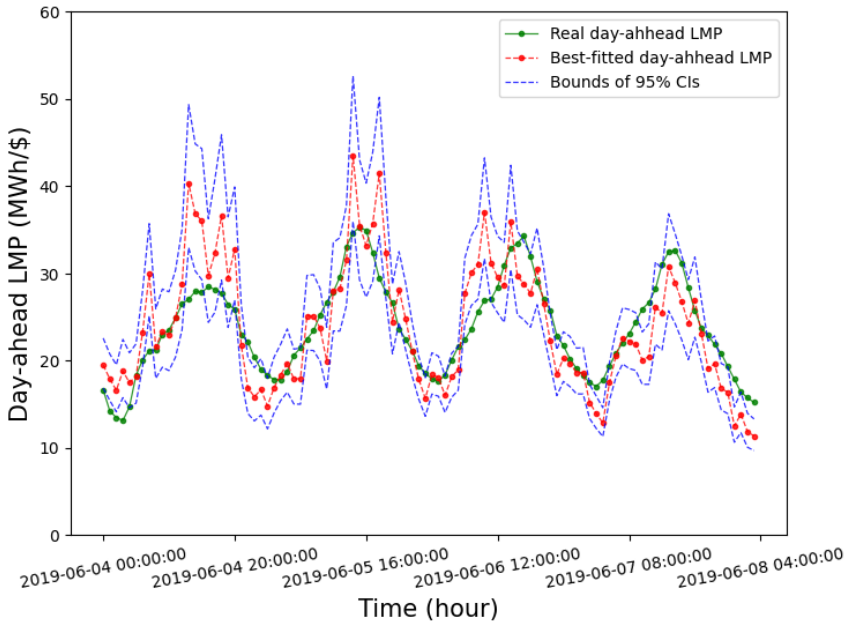
3.4.5 Day-ahead LMP Results

We analyze day-ahead LMP prediction results from two perspectives: (1) impacts of climate change on day-ahead LMPs and (2) change in day-ahead LMP variance under climate change scenarios.

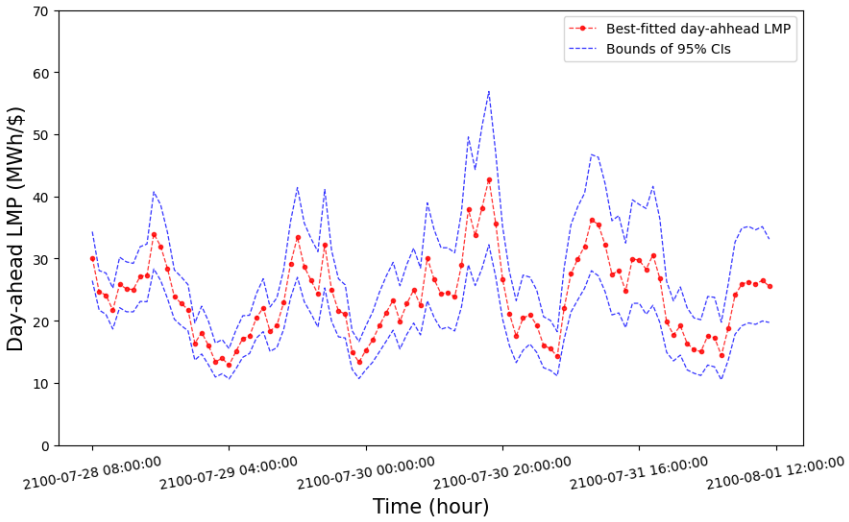
Impact of climate change on day-ahead LMPs: To investigate how climate change affects day-ahead LMPs across various regions within MISO, we conducted a statistical analysis using the best-fitted Model (3.5). This analysis encompasses 13 selected nodes in four different Zones (Zones 2, 4, 8, and 9) within MISO, as outlined in Table 3.5. Again, we choose these four Zones due to their north-to-south vertical distribution, enabling us to examine the effects of climate variations across various latitudes.

It is essential to note that the coefficients obtained for the best-fitted Model (3.5) pertain to the PCA-transformed variables rather than the real (normalized) climate variables. To derive coefficients for the real (normalized) climate variables, we need to transfer the obtained coefficients by multiplying the PCA transfer matrix. For the sake of clarity, we have averaged the transferred coefficients for each climate variable. We also calculated the standard variance for the mean of coefficients for each climate variable and presented the results in Table 3.5. The details about calculating the mean of the transferred coefficients and the standard variance can be found in the Appendix B.2.

In Table 3.5, all nodes suggest that wind speed and precipitation exhibit negative (marginal) impacts on day-ahead LMPs. This phenomenon can be attributed to higher



(a) Best-fitted day-ahead LMP



(b) Forecasted day-ahead LMP

Figure 3.11: An example of the best-fitted day-ahead LMP and forecasted day-ahead LMP (under the ACCESS-CM2 GCM & SSP126) in the ALTE.SHEEPSIN1 Node

wind speeds increasing wind power generation, thereby reducing electricity prices. Additionally, increased precipitation facilitates power plant cooling and energy storage, leading to a subsequent decrease in electricity prices. Conversely, all nodes indicate positive impacts of temperature and solar radiation on day-ahead LMPs. Higher temperatures result in increased energy consumption for cooling purposes, contributing to an increase in electricity prices. While higher solar radiation may lead to greater solar energy generation, it is also associated with elevated temperatures, which, in turn, influence electricity prices positively.

Zones 2 and 4 exhibit higher sensitivity to wind compared to Zones 8 and 9. The WPS.CRANECREK Node in Zone 2 and three out of four nodes (AMIL.TRA_ATRAE, MEC.NEALS_4, and AMIL.STWF) in Zone 4 show mean wind coefficients smaller than -1×10^{-3}. Furthermore, the AMIL.STWF Node in Zone 4 reveals an even more pronounced sensitivity, with mean wind coefficients below -4×10^{-3}. However, all nodes in Zones 8 and 9 display mean wind coefficients greater than -0.7×10^{-3} and smaller than -0.03×10^{-3}. This variation in sensitivity patterns arises because a massive portion of the MISO active wind-power capacity as of 2020 (a total of 15,720 MW) is concentrated in northern MISO regions, such as Michigan (1,715 MW, 10.9%), Iowa (3,785 MW, 24.1%) and Illinois (2,337 MW, 14.9%). In contrast, there is a much smaller capacity in southern MISO regions like Arkansas (407 MW, 2.6%) and Louisiana (less than 200 MW, 1.2%) (Sustainable FERC, 2021). Zones 2 and 4, situated in the northern MISO, are consequently more susceptible to the influence of wind power supply, resulting in heightened wind sensitivity for day-ahead LMPs in Zones 2 and 4 compared to Zones 8 and 9.

Change in day-ahead LMP variances under climate change scenarios: To examine the

Table 3.5: Impacts of climate change on different MISO nodes

Location	**Zone**	**Node Name**	**Mean of Climate Coefficients (standard variance)** $\times 10^{-3}$			
			wind	**precipitation**	**temperature**	**radiation**
Northern MISO , including WI	Zone 2	ALTE.SHEEPSIN1	-0.05 (0.03)	-1.61 (0.13)	0.78 (0.04)	0.57 (0.04)
		WPS.CRANECREK	-1.45 (0.03)	-3.18 (0.13)	0.50 (0.04)	0.96 (0.04)
		WPS.CUSTEGEN1	-0.30 (0.02)	-2.91 (0.10)	0.71 (0.03)	0.67 (0.03)
North-central MISO , including IL	Zone 4	AMIL.HEN_AMT_P	-0.63 (0.04)	-2.5 (0.15)	0.69 (0.04)	0.79 (0.05)
		AMIL.TRA_ATRAE	-2.34 (0.07)	-3.62 (0.28)	0.77 (0.08)	0.98 (0.08)
		MEC.NEALS_4	-2.25 (0.04)	-2.56 (0.18)	0.54 (0.05)	1.01 (0.05)
		AMIL.STWF	-4.61 (0.08)	-0.53 (0.31)	0.14 (0.09)	0.98 (0.09)
South-central MISO , including AR	Zone 8	EAI.ANO2	-0.48 (0.05)	-2.31 (0.20)	0.47 (0.06)	0.64 (0.06)
		EAI.INDEPEND2	-0.16 (0.02)	-2.02 (0.07)	0.38 (0.02)	0.66 (0.02)
		EAI.INDEPEND1	-0.09 (0.02)	-2.31 (0.10)	0.35 (0.03)	1.06 (0.03)
		EAI.AECCHYDRO2	-0.69 (0.06)	-2.10 (0.27)	0.11 (0.07)	0.93 (0.08)
Southern MISO , including LA	Zone 9	LEPA.HOUMA_G16	-0.03 (0.02)	-1.66 (0.14)	0.21 (0.03)	0.74 (0.03)
		LEPA.MURRAY	-0.05 (0.02)	-1.42 (0.11)	0.24 (0.02)	0.41 (0.02)

Note: The standard variances ($\times 10^{-3}$) in parentheses are for the mean of climate coefficients.

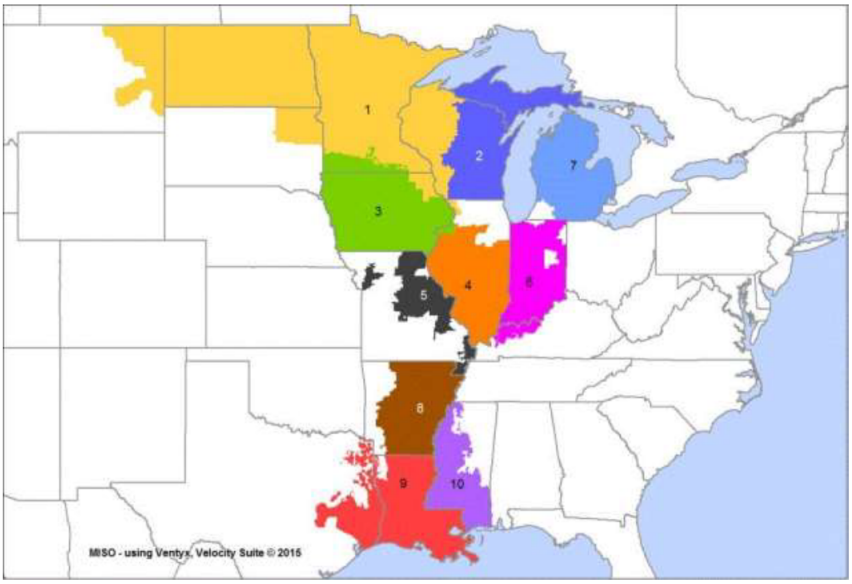


Figure 3.12: MISO Zone map (retrieved from MISO Inc. (2020))
Zone 2, including Wisconsin (WI), is shaded in blue. Zone 4, including Illinois (IL), is shaded in orange. Zone 8, including Arkansas (AR), is shaded in brown. Zone 9, including Louisiana (LA), is shaded in red.

fluctuations in day-ahead LMPs in the MISO region, we investigate the growth rates of day-ahead LMP variances under a range of climate scenarios. Calculating these growth (%) involves the following steps, which are illustrated here for the ACCESS-CM2 GCM,

in Zone 4 under SSP 585 (See Figure 3.13).

1. We calculate annual LMP variances for each node (in Zone 4) and year (from 2021 to 2100) by assessing the variance of predicted hourly LMP data for that node and year. These annual variances are then averaged across all (four) nodes within the same Zone (Zone 4) to obtain Zone-specific annual LMP variances (depicted as blue dots in Figure 3.13).

2. We establish a line trend (regression) of the 80 annual variances for Zone 4 from 2021 to 2100 (See the blue line in Figure 3.13) and project the linear-fitted variances for both 2021 and 2100.

3. We obtain the growth rates by subtracting the linear-fitted variance in 2021 from the linear-fitted variance in 2100 and then dividing the result by the linear-fitted variance in 2021.

We do these three-step calculations for each combination of 4 SSPs, 4 Zones, and 27 GCMs, resulting in a total of 432 values.

Figure 3.14 shows the growth (%) in day-ahead LMP variances across four SSPs (for the 432 values). Under the SSP126 scenario, all 27 GCMs indicate the growth rates falling in the range of $[-4.7\%, 5.2\%]$. As the emission scenario transitions from SSP126 to SSP585, the ranges of the growth rates expand. Under SSP585, the highest emission scenario, the ranges of the growth rates expand to $[-15.0\%, 3.5\%]$. All GCMs under SSP585 suggest a reduction in LMP variances (at most 15%) over the 80-year period in Zone 2, 8, and 9. However, in Zone 4, two GCMs (i.e., CMCC-CM2-SR5 and CMCC-ESM2) suggest an

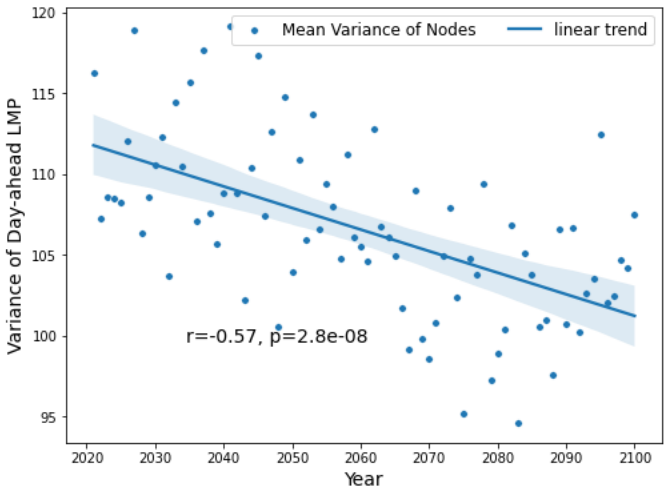
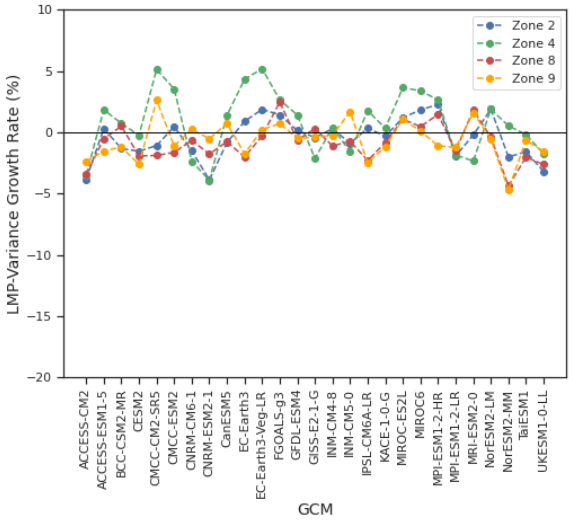


Figure 3.13: An example of a line trend analysis of annual day-ahead LMP variances for SSP585 in Zone 4 under the ACCESS-CM2 GCM

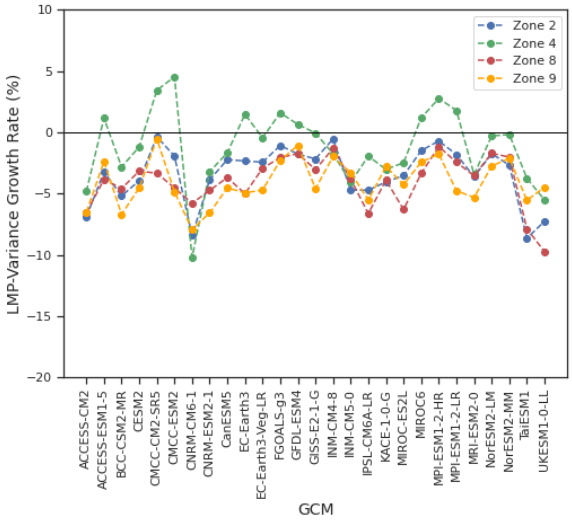
The blue line is the linear trend (regression) derived from all blue dots, with "r" denoting the Pearson correlation coefficient and "p" indicating the p-value used to test for non-correlation, both associated with the blue line.

increase of LMP variations. This expansion of growth rate ranges may be attributed to reduced wind variances and heightened precipitation variances, as the emission scenario transitions from SSP126 to SSP585.

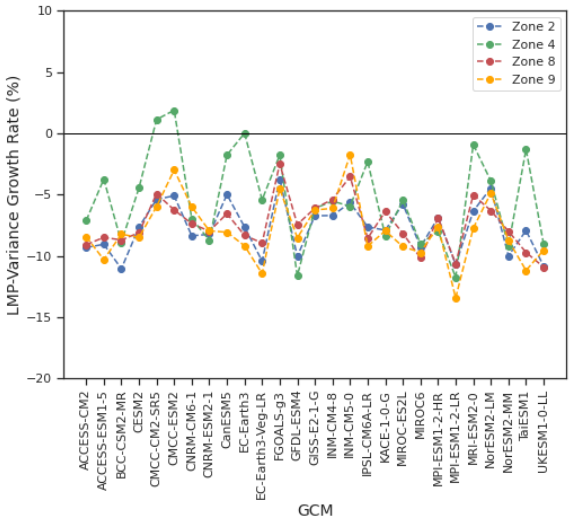
Zone 4 experiences a higher growth (or lower decrease) in LMP variances when compared to Zones 2, 8, and 9 across all SSPs. This could be attributed to the increased sensitivity of day-ahead LMP in Zone 4 to changes in wind patterns. The fluctuations in wind due to climate change may result in greater LMP variance in Zone 4 in contrast to Zones 2, 8, and 9.



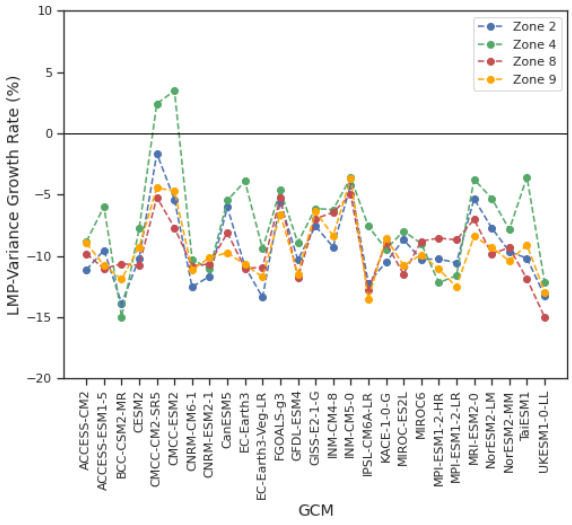
(a) SSP126



(b) SSP245



(c) SSP370



(d) SSP585

Figure 3.14: Growth (%) of day-ahead LMP variances between 2021 and 2100

3.4.6 Discussion of day-ahead LMP forecasts

Here we discuss how different climate trajectories translate into differences in day-ahead LMP variances. As greenhouse gas emissions increase (from SSP126 to SSP585), the level of LMP variances deviating from their 2021 levels also rises. For example, in Figure 3.14, under the low-emission scenario SSP126, most GCMs suggest that the LMP variances remain relatively stable as climate evolves from 2021 to 2100. However, high-emission scenario SSP585 implies that the LMP variances strongly deviate from their 2021 levels as the climate evolves. Specifically, under SSP585, most GCMs suggest a decrease in the LMP variances, mostly by around 15.0%, in Zone 2, 8, and 9, while two out of 27 GCMs suggest an increase of the LMP variances of up to 3.5% in Zone 4.

3.5 Study II: Operation Decisions in MISO under Climate Change Scenarios

In this section, we apply the robust BESS Model (3.3) to study BESS operations in electricity nodes in MISO under different climate scenarios. Specifically, we study the optimal hourly operations of a BESS in MISO from 2021 up to 2100. First, we formulate the LMP bounds as intervals for each climate scenario. After getting the bounds, we solve Model (3.3) multiple times under different scenarios to investigate operations of the BESS in the MISO case.

3.5.1 Methodology for the Empirical Study of Operation Decisions

Based on the forecasted confidence intervals for future day-ahead LMPs, we establish the day-ahead LMP bounds for the BESS operation model (see Model (3.3)). We illustrate this process using an example for a node, a GCM, and an SSP within an hour t ($t \in \mathcal{T}$) to demonstrate how to form the day-ahead LMP bound.

First, for the node, we calculate the mean of $sinh^{-1}(P_t)$ and its corresponding 95% confidence interval (CI95). We assume that future year indicators are the same as the year indicator in 2019 when predicting electricity prices, i.e., $\sum_{y=2016}^{2019} \beta_y I_y(t) = \beta_{2019}$ ($\forall t \in \mathcal{T}$). We denote the (estimated) mean of $sinh^{-1}(P_t)$ as $sinh^{-1}\bar{P}_t$ and the range of the CI95 as $2d_t$. Then, we can express the CI95 for $sinh^{-1}\bar{P}_t$ as $[sinh^{-1}\bar{P}_t - d_t, sinh^{-1}\bar{P}_t + d_t]$.

Second, we transform the confidence interval for $sinh^{-1}\bar{P}_t$ into the day-ahead LMP bounds. We define $\bar{p}_t = \frac{1}{2}[sinh(sinh^{-1}\bar{P}_t - d_t) + sinh(sinh^{-1}\bar{P}_t + d_t)]$, and $\hat{p}_t = \bar{p}_t - sinh(sinh^{-1}\bar{P}_t - d_t)$. We can model the day-ahead LMP bounds as $[\bar{p}_t - \hat{p}_t, \bar{p}_t + \hat{p}_t]$.

For each node, SSP, GCM, and each hour in the period of 2021-2100, we follow these two steps to form the price interval trajectory required for solving Model (3.3). In total, we have 1,404 ($13 \times 4 \times 27$) trajectories of hourly day-ahead LMP bounds from 2021 to 2100.

Besides the price parameters, we set the remaining parameters in Model (3.3) as follows: hourly discount rate $\delta = 1 - 2.3 \times 10^{-6}$, the BESS's energy capacity $C^S = 200$ (MWh), charging capacity in an hour $C^C = 10$ (MWh), discharging capacity in an hour $C^D = 10$

(MWh), charging efficiency $\eta^C = 0.9$, discharging efficiency $\eta^D = 0.9$, and the BESS's initial storage level $x_0 = 0$ (MWh).

We verify the 1,404 LMP bound trajectories and confirm that they all satisfy Lemma 2, ensuring that Model (3.3) can be represented as an equivalent linear model. Subsequently, we solve this equivalent linear model for each of the 1,404 LMP bounds using CPLEX, resulting in the generation of 1,404 optimal operation policies for the BESS.

3.5.2 BESS Operation Results

This section presents the results of the numerical implementation in the MISO case. We present how the operating income of the BESS evolves under climate change scenarios.

To understand how different climate scenarios affect the operating incomes of the BESS in different MSIO regions, we studied annual operating incomes from 2021 to 2100 for all 108 climate scenarios (27 GCMs $\times$ 4 SSPs). The annual operating incomes are obtained from the Model (3.3) for each node and climate scenario.

Figure 3.15 presents the annual operating incomes for all four Zones. Within the figure, each solid line represents the mean of annual operating incomes across 27 GCMs and nodes in the Zone for a given SSP. Dashed lines, sharing the same color as the solid line, represent the ranges (bounds) of these annual operating incomes for the given SSP. Among the four Zones, Zone 4 exhibits the highest mean operating incomes, likely due to its highest changes in day-ahead LMP variance among the four zones. The lower bound of the annual operating incomes in Zone 4 is even close to the upper bound of annual operating incomes in Zone 9. Zone 2 shows the second-largest mean operating incomes with the largest variance around

the mean. The mean annual operating incomes in Zones 8 and 9 are lower than Zones 2 and 4.

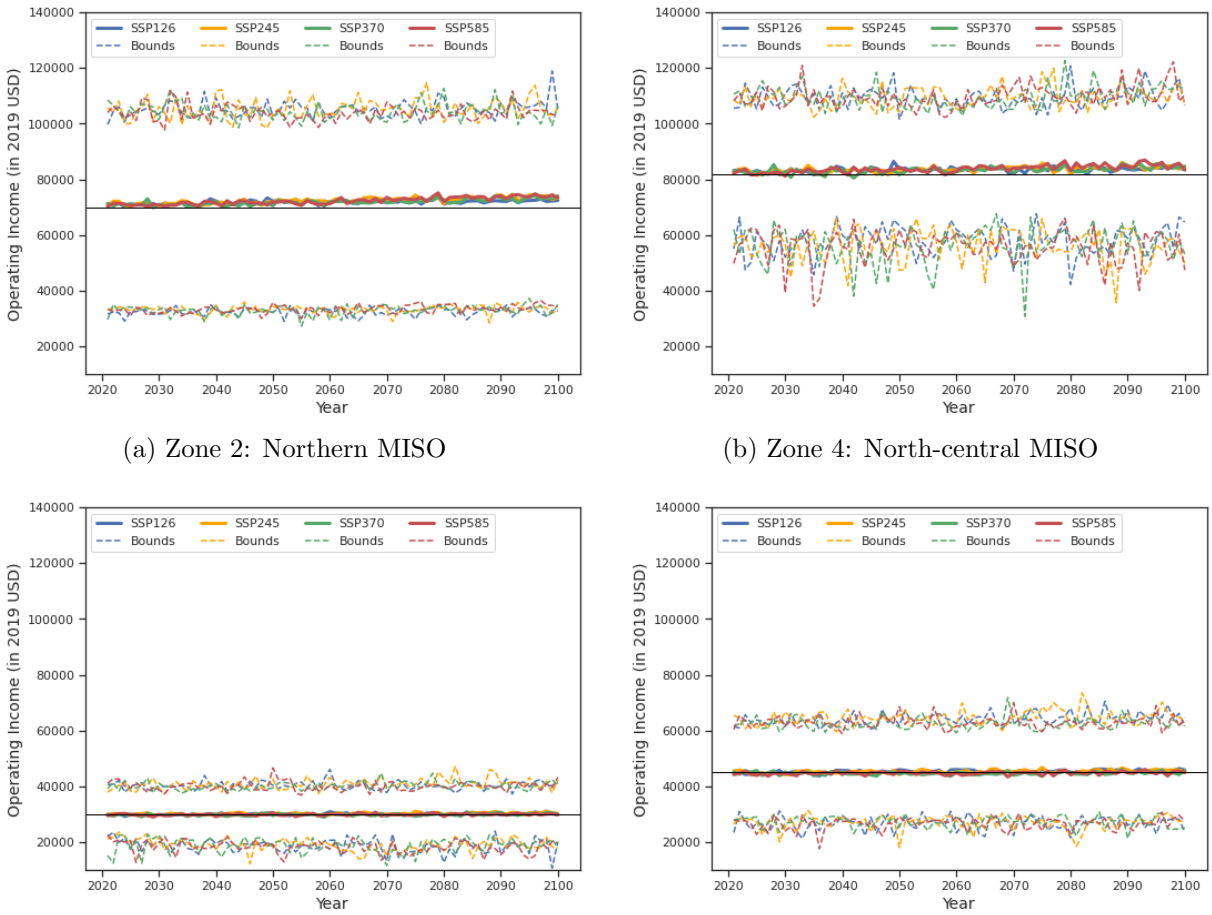
Figure 3.15 also illustrates that the mean operating incomes are relatively stable across different SSPs and the 80 years. However, there are substantial variations in mean operating incomes, indicating significant income disparities across the 27 GCMs. For instance, Zone 2 demonstrates the widest range of operating incomes, reaching up to 79,000 USD, closing even its mean operating incomes, which are approximately 71,000 USD.

It is also important to study how different climate scenarios affect change in operating incomes. To understand this change, we study the operating income growths (%) for all 27 GCMs under 4 SSPs separately. The way to calculate the operating income growth is similar to calculate the growth of day-ahead LMP variances (as discussed in Section 3.4.5) but with the focus shifted to operating incomes rather than day-ahead LMP variances. Here is a summary of the process illustrated for a SSP, in a Zone under a GCM:

1. Obtain annual Zone-specific operating incomes by averaging the annual operating incomes of individual nodes within the Zone. Calculate these annual Zone-specific operating incomes for an 80-year period spanning from 2021 to 2100. For comparison purposes, we measure all the operating incomes in 2019 USD.

2. Fit a linear trend (regression) to the (Zone-specific) annual operating incomes to project operating incomes for both 2021 and 2100.

3. Calculate the operating income growth as the quotient of the difference between the projected 2100 and 2021 operating incomes, divided by the projected 2021 operating income.



(a) Zone 2: Northern MISO

(b) Zone 4: North-central MISO

(c) Zone 8: South-central MISO

(d) Zone 9: Southern MISO

Figure 3.15: Projected annual operating incomes in MISO Zones from 2021 to 2100
Each solid line is the mean of operating incomes across the 27 GCMs and nodes in the Zone, and dashed lines with the same color as the solid line are the ranges (bounds) of operating incomes across the 27 GCMs and nodes in the Zone, for a SSP.

We perform the three-step calculation for each combination of 27 GCMs, 4 Zones, and 4 SSPs, resulting in 432 values.

Figure 3.16 presents operating incomes growth (%) of the BESS across SSP126, SSP245, SSP370, and SSP585 (for the 432 values). Under SSP126 (see Figure 3.16a), all GCMs suggest an operating income growth within the range of [-4%,9%] in Zones 2, 4, 8, and 9.

Similar trends are observed for SSP245 and SSP370 (see Figures 3.16b and 3.16c). However, in Zone 2, one GCM suggests operating income growth exceeding 10% for SSP245, and two GCMs suggest the same for SSP370. Under SSP370, the MPI-ESM1-2-HR GCM suggests operating income growths less than -6% in Zone 9.

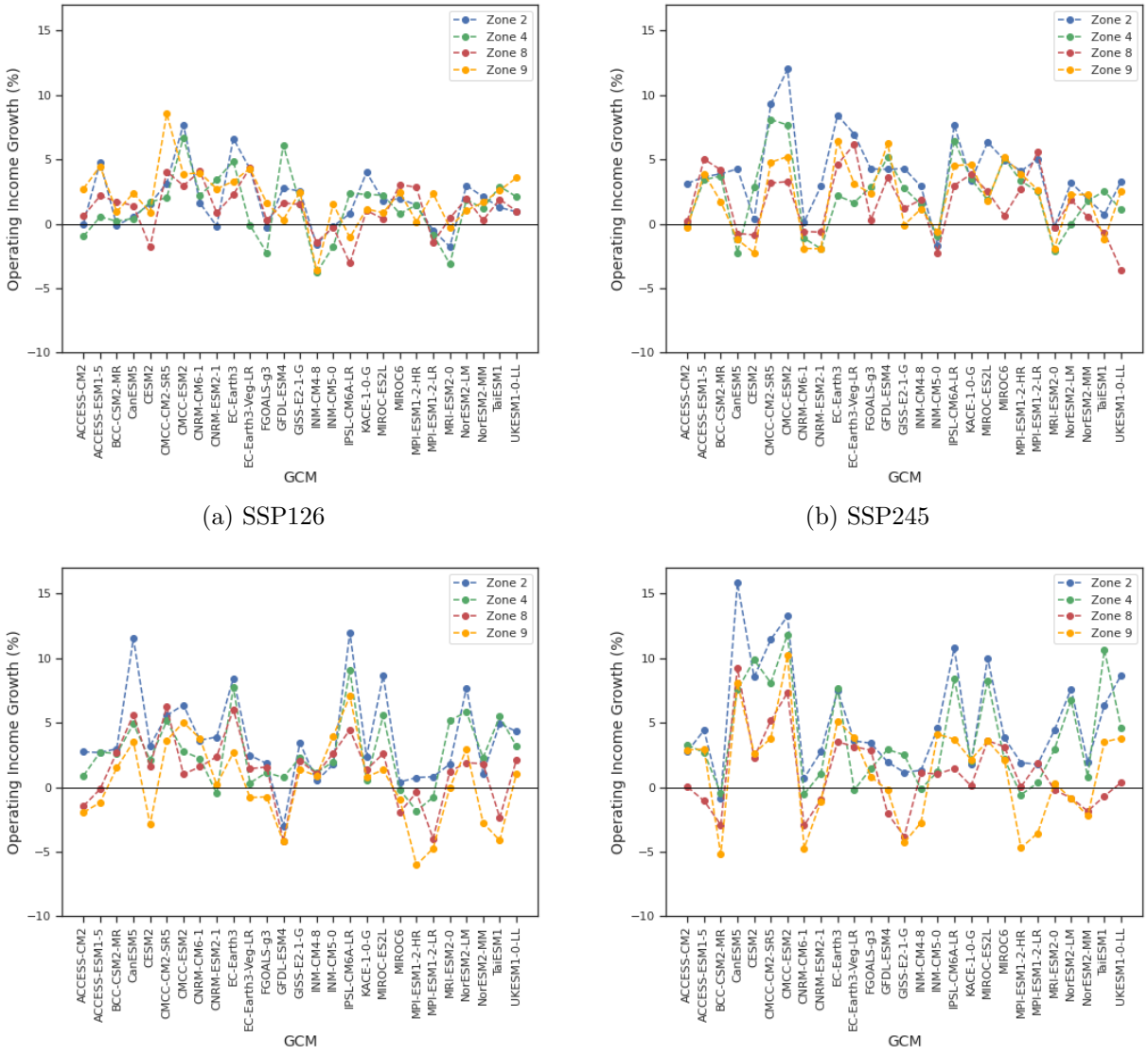
For SSP585 (see Figure 3.16d), nine GCMs suggest an increase of more than 5% in operating incomes in Zone 4, of which two GCMs anticipate an increase in operating income by more than 10%. The CanESM5 GCM suggests operating income growth exceeding 15% in Zone 2. In Zone 9, the BCC-CSM2-MR GCM suggest reductions in operating income surpassing 5%.

All these patterns in the operating income growth underscore that the BESS's operating income growth varies across GCMs and SSPs. In general, the ranges of operating income growth increase when greenhouse gas emissions escalate (from SSP126 to SSP585).

3.5.3 Discussion of Forecasted BESS Operations

The empirical study elucidates how hourly climate trajectories, generated by various climate models, impact the operating incomes of the BESS through arbitrage in MISO. These results provide insights into how challenging it is to quantify disagreements between forecasts and translate them into different operating incomes.

The challenge of addressing aggregated forecasts becomes apparent when comparing the operating incomes derived from different GCMs, SSPs, and Zones. Most GCMs suggest that the operating income from 2021 to 2100, under SSP126, remains relatively stable over time, although several GCMs suggests a minor increase of approximately 9% or a decrease



(a) SSP126

(b) SSP245

(c) SSP370

(d) SSP585

Figure 3.16: Operating income growth (%) between 2021 and 2100 under each SSP

of about 4%. In contrast, the operating income growth rates associated with the other three SSPs show more significant divergence. Under SSP585, the divergence in operating

income growth rates among different GCMs and Zones could surpass 20%. For example, in Figure 3.16d, the operating income growth rate in Zone 2, suggested by the CanESM5 GCM, exceeds that in Zone 9, suggested by the BCC-CSM2-MR GCM, by more than 20%.

Zone 4 achieves the highest operating incomes through arbitrage among the four Zones. The highest income is attributed to high sensitivity to wind variance and high LMP variance, creating more opportunities for profitable arbitrage. Investing in BESSs for arbitrage in Zone 4 may be a promising choice, considering the anticipated decrease in BESS prices in the future.

However, due to the high variance of operating incomes among different climate scenarios, BESS operators or investors should consider multiple climate models in their decision-making process to adapt to uncertain climate change.

3.6 Conclusion

This chapter contributes on two main fronts: Firstly, it introduces a scalable methodology tailored for assessing operational decisions within a timeframe and scope suitable for climate adaptation research within the electricity market. Secondly, it applies these methodologies to investigate the dependency of the operating incomes of a BESS on valuations derived from diverse climate forecasts.

The empirical study uses a regression model and a robust optimization model to study BESS operations over 80 years within the MISO from 2021 to 2100. The study suggests that differences across climate scenarios have observable impacts on day-ahead LMP variance

and the operating incomes of BESSs. As greenhouse gas emission scenarios increase, such as the transition from SSP126 to SSP585, climate change leads to increased fluctuations in electricity prices and the operating incomes of the BESS in the MISO market.

Chapter 4

Conclusions and Future Research

This chapter summarizes the contributions of this dissertation and outlines directions for future research.

4.1 Conclusions

Efficient climate change adaptation involves decision-making over a long-time horizon with uncertainty. This dissertation presents a framework that combines climate data, regression models, and robust optimization to study climate adaptation with applications in two areas: land investment and electricity storage systems. This framework extends the application of operations research in these two areas by connecting the broader climate perspective to specific operational decisions. The results obtained from these two applications provide valuable insights for a wide range of stakeholders, including investors, government bodies, and the public.

The main takeaway from these applications underscores the significance of incorporating a wide range of specific climate scenarios, along with detailed temporal and geographic data, when making climate adaptation decisions. The following two subsections provide detailed conclusions for the two applications of the framework.

4.1.1 Climate Change Adaptation in Land Investment

We conducted a study on climate adaptation in land investment over a 68-year horizon in the Mississippi River Basin (MRB). Our approach involved applying regression models to predict future land price trajectories under different climate change scenarios. We introduced a robust model parameterized by the land price trajectories to analyze a land investor's decisions regarding land acquisition, including when and where to buy or sell land and how much to buy or sell. We demonstrated that these robust models have linear equivalence and can be easily solved. Furthermore, our models extend the models found in the literature. Our numerical study in the MRB case revealed that land investment varies across climate scenarios and spatial areas. A high greenhouse gas emission scenario (e.g., RCP8.5) diminishes land investment opportunities across the MRB, as this scenario significantly decreases land values over the next 80 years.

4.1.2 Climate Change Adaptation in Electricity Storage Systems

We conducted a study on the arbitrage incomes of a battery energy storage system (BESS) over an 80-year horizon in the Midcontinent Independent System Operator (MISO) market

under climate change. We proposed a climate downscaling method to temporally downscale climate data to the desired resolution for our regression model that links electricity prices and climate variables. We applied the regression models to predict future electricity price trajectories under different climate change scenarios. Further, we proposed a robust model parameterized by the electricity price trajectories to analyze a BESS's operational decisions, including when to charge and discharge the BESS and how much to charge and discharge. Our numerical study in the MISO case suggests that climate change impacts the variation in electricity prices and BESS's operating incomes, both of which exhibit differences across various climate scenarios and geographical regions. As greenhouse gas emission scenarios increase (e.g., from SSP126 to SSP585), climate change introduces greater variability in electricity prices and the operating incomes of the BESS in the MISO market.

4.2 Future Research

While our two main applications of the proposed framework are attractive, this dissertation also has certain limitations. We summarize these limitations and suggest directions for further exploration in the future.

Firstly, we acknowledge that the assumption of a future investment environment or electricity market structure identical to the current conditions can be overly simplistic. Over time, these environments and structures evolve and change. To address this, additional prediction models can be integrated with the price regression models to more accurately forecast future parameters for the robust optimization models.

Secondly, our robust optimization models could benefit from the consideration of more real-world settings or constraints to enhance the models' applicability. For instance, in the robust model for land investment, incorporating factors such as labor costs and farm machinery transit costs could provide a more comprehensive measure of the profitability of land investment. In the case of the electricity storage system, it would be beneficial to include the real-time electricity market in the electricity price model and the robust model. This is crucial because the real-time electricity market can also be profitable for electricity storage systems.

Thirdly, it is worth noting that our proposed framework does not account for the rationality of stakeholders. Our models assume rational stakeholders, while real-world stakeholders often exhibit irrational behavior. Furthermore, stakeholders in practice may not always possess complete information about the current and future environment. Future research could explore these scenarios involving irrational stakeholders and incomplete information to enhance the realism of our models.

Fourth, it is interesting to study the relationship between confidence levels and operating results, especially determining at which confidence level the operating patterns undergo significant changes. Answers to these questions may help operators choose the proper confidence levels for their applications.

References

Adopted, IPCC. Climate change 2014 synthesis report. *IPCC: Geneva, Szwitzerland*, 2014.

Alevin NRDCINC1. Midcontinent Independent System Operator (MISO) borders, 2021. https://www.arcgis.com/home/item.html?id=8fcdeeeb2b5449559985552a99e2e4b1. Accessed on September/19/2022.

Afshin Amiraslany. *The impact of climate change on Canadian agriculture: a Ricardian approach*. PhD thesis, University of Saskatchewan, 2010.

American Meteorological Society AMS. downwelling, 2012. https://glossary.ametsoc.org/wiki/Downwelling. (Accessed on April/12/2023).

Maryam Arbabzadeh, Ramteen Sioshansi, Jeremiah X Johnson, and Gregory A Keoleian. The role of energy storage in deep decarbonization of electricity production. *Nature communications*, 10(1):3413, 2019.

The European Wind Energy Association. How efficient are wind turbines?, 2023. DOI:https://www.ewea.org/wind-energy-basics/faq/. (Accessed on June/8/2023).

Atalay Atasu, Charles J Corbett, Ximin Huang, and L Beril Toktay. Sustainable operations management through the perspective of manufacturing & service operations management. *Manufacturing & service operations management*, 22(1):146–157, 2020.

Martin B Bagaram and Sándor F Tóth. Multistage sample average approximation for harvest scheduling under climate uncertainty. *Forests*, 11(11):1230, 2020.

François Bareille and Raja Chakir. The impact of climate change on agriculture: A repeat-ricardian analysis. *Journal of Environmental Economics and Management*, 119:102822, 2023.

Alan Barreca, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro. Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, 124(1):105–159, 2016.

Jens Beckert. *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Harvard University Press, 2016.

Stephen E Belcher, Jacob N Hacker, and Dianne S Powell. Constructing design weather data for future climates. *Building services engineering research and technology*, 26(1): 49–61, 2005.

N Bellouin, WJ Collins, ID Culverwell, PR Halloran, SC Hardiman, TJ Hinton, CD Jones, RE McDonald, AJ McLaren, FM O'Connor, et al. The HadGEM2 family of met office unified model climate configurations. *Geoscientific Model Development*, 4(3):723–757, 2011.

Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical programming*, 88:411–424, 2000.

Martin Beniston, David B Stephenson, Ole B Christensen, Christopher AT Ferro, Christoph Frei, Stéphane Goyette, Kirsten Halsnaes, Tom Holt, Kirsti Jylhä, Brigitte Koffi, et al. Future extreme events in european climate: an exploration of regional climate model projections. *Climatic change*, 81:71–95, 2007.

Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization and network flows. *Mathematical programming*, 98(1):49–71, 2003.

Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1): 35–53, 2004.

Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

Ajay Gajanan Bhave, Declan Conway, Suraje Dessai, and David A Stainforth. Barriers and opportunities for robust decision making approaches to support climate change adaptation in the developing world. *Climate Risk Management*, 14:1–10, 2016.

John R Birge and Francois Louveaux. *Introduction to Stochastic Programming*. Springer Science & Business Media, 2011.

John R Birge, Ali Hortaçsu, Ignacia Mercadal, and J Michael Pavlin. Limits to arbitrage in electricity markets: A case study of miso. *Energy Economics*, 75:518–533, 2018.

Board of Governors of the Federal Reserve System. 2020 statement on longer-run goals and monetary policy strategy, Washington, DC, 2021. https://www.federalreserve.gov/. Accessed on May/06/2022.

Brad Boehmke and Brandon M Greenwell. *Hands-on machine learning with R*. CRC press, 2019.

Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Martina Bozzola, Emanuele Massetti, Robert Mendelsohn, and Fabian Capitanio. A ricardian analysis of the impact of climate change on italian agriculture. *European Review of Agricultural Economics*, 45(1):57–79, 2018.

C Bracken. Downscaled CMIP3 and CMIP5 climate projections-addendum release of downscaled CMIP5 climate projections (LOCA) and comparison with preceding information. *Technical Service Center, Bureau of Reclamation, US Department of the Interior: Denver, CO, USA*, 2016.

Teresa Armada Brás, Júlia Seixas, Nuno Carvalhais, and Jonas Jägermeyr. Severity of drought and heatwave crop losses tripled over the last five decades in europe. *Environmental Research Letters*, 16(6):065012, 2021.

M. Tyson Brown. Nerc assessment highlights potential electricity reliability concern for central u.s., 2022. https://www.eia.gov/todayinenergy/detail.php?id=52618. (Accessed on June/24/2023).

Edward A Byers, Gemma Coxon, Jim Freer, and Jim W Hall. Drought and climate change impacts on cooling water shortages and electricity prices in great britain. *Nature Communications*, 11(1):2239, 2020.

Jose A Carta, Penelope Ramirez, and Sergio Velazquez. A review of wind speed probability distributions used in wind energy analysis: Case studies in the canary islands. *Renewable and sustainable energy reviews*, 13(5):933–955, 2009.

Shankar N Chandramowli and Frank A Felder. Impact of climate change on electricity systems and markets–a review of models and forecasts. *Sustainable Energy Technologies and Assessments*, 5:62–74, 2014.

Yongfu Chen, Zhigang Wu, Katsuo Okamoto, Xinru Han, Guoying Ma, Hsiaoping Chien, and Jing Zhao. The impacts of climate change on crops in china: A ricardian analysis. *Global and Planetary Change*, 104:61–74, 2013.

Yupeng Chen, Garud Iyengar, and Chun Wang. Robust inventory management: A cycle-based approach. *Manufacturing & Service Operations Management*, 25(2):581–594, 2023.

Bolong Cheng and Warren B Powell. Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming. *IEEE Transactions on Smart Grid*, 9(3):1997–2005, 2016.

Yuan Shih Chow and Henry Teicher. *Probability theory: independence, interchangeability, martingales*. Springer Science & Business Media, 2003.

CMIP6 Data Request. CMIP6 data request, 2023. https://clipc-services.ceda.ac.uk/dreq/mipVars.html. (Accessed on April/11/2023).

CoinNews Media Group. U.S. inflation calculator, 2022. https://www.usinflationcalculator.com. Accessed on May/08/2022.

Matthew Collins, Reto Knutti, Julie Arblaster, Jean-Louis Dufresne, Thierry Fichefet, Pierre Friedlingstein, Xuejie Gao, William J Gutowski, Tim Johns, Gerhard Krinner, et al. Long-term climate change: projections, commitments and irreversibility. In *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1029–1136. Cambridge University Press, 2013.

Sara M Constantino and Elke U Weber. Decision-making under the deep uncertainty of climate change: The psychological and political agency of narratives. *Current opinion in psychology*, 42:151–159, 2021.

Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. Arima models to predict next-day electricity prices. *IEEE transactions on power systems*, 18 (3):1014–1020, 2003.

Climate Data Store Copernicus Climate Change Service. ERA5 hourly data on single levels from 1940 to present., 2023. DOI:10.24381/cds.adbb2d47. (Accessed on March/26/2023).

Jennifer Cronin, Gabrial Anandarajah, and Olivier Dessens. Climate change impacts on the energy system: a review of trends and gaps. *Climatic change*, 151:79–93, 2018.

Klaus Deininger. Challenges posed by the new wave of farmland investment. *The Journal of Peasant Studies*, 38(2):217–247, 2011.

Klaus Deininger and Derek Byerlee. *Rising Global Interest in Farmland: Can It Yield Sustainable and Equitable Benefits?* World Bank Publications, 2011.

E Delage. Quantitative risk management using robust optimization. *Lecture Notes, HEC Montréal*, 2021.

Temesgen Tadesse Deressa. *Measuring the economic impact of climate change on Ethiopian agriculture*, volume 4342. World Bank Publications, 2007.

Olivier Deschênes and Michael Greenstone. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1):354–385, 2007.

Jay L Devore. *Probability and Statistics for Engineering and the Sciences*. Cengage learning, 2011.

Luca Di Corato and Dimitrios Zormpas. Investment in farming under uncertainty and decoupled support: a real options approach. *European Review of Agricultural Economics*, 49(4):876–909, 2022.

Peter K Dunn, Gordon K Smyth, et al. *Generalized linear models with examples in R*, volume 53. Springer, 2018.

M Ek, T Murdock, S Sobie, B Cavka, B Coughlin, and R Wells. Future weather files to support climate resilient building design in vancouver. In *Proceedings of the 1st International Conference on New Horizons in Green Civil Engineering (NHICE-01), Victoria, BC, Canada*, pages 25–27, 2018.

Energy Information Administration. Electricity generation, capacity, and sales in the united states, 2023a. https://www.eia.gov/energyexplained/electricity/electricity-in-the-us-generation-capacity-and-sales.php. Accessed on August/31/2023.

Energy Information Administration. What is u.s. electricity generation by energy source?, 2023b. https://www.eia.gov/tools/faqs/faq.php?id=427&t=3. Accessed on August/31/2023.

Prince M Etwire, David Fielding, and Viktoria Kahui. Climate change, crop selection and agricultural revenue in ghana: A structural ricardian analysis. *Journal of Agricultural Economics*, 70(2):488–506, 2019.

Veronika Eyring, Sandrine Bony, Gerald A Meehl, Catherine A Senior, Bjorn Stevens, Ronald J Stouffer, and Karl E Taylor. Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016.

Madeleine Fairbairn. 'like gold with yield': Evolving intersections between farmland and finance. *The Journal of Peasant Studies*, 41(5):777–795, 2014.

Madeleine Fairbairn. Fields of gold. In *Fields of Gold*. Cornell University Press, 2020.

Madeleine Fairbairn, Jim LaChance, Kathryn Teigen De Master, and Loka Ashwood. In vino veritas, in aqua lucrum: Farmland investment, environmental uncertainty, and groundwater access in california's cuyama valley. *Agriculture and Human Values*, pages 1–15, 2020.

Madeleine Fairbairn, Jim LaChance, Kathryn Teigen De Master, and Loka Ashwood. In vino veritas, in aqua lucrum: Farmland investment, environmental uncertainty, and groundwater access in california's cuyama valley. *Agriculture and Human Values*, 38(1): 285–299, 2021.

Jing-Li Fan, Jia-Wei Hu, and Xian Zhang. Impacts of climate change on electricity demand in china: An empirical estimation based on panel data. *Energy*, 170:880–888, 2019.

Shu Fan, Chengziong Mao, and Luonan Chen. Next-day electricity-price forecasting using a hybrid network. *IET generation, transmission & distribution*, 1(1):176–182, 2007.

Carlo Fezzi and Ian Bateman. The impact of climate change on agriculture: Nonlinear effects and aggregation bias in ricardian models of farmland values. *Journal of the Association of Environmental and Resource Economists*, 2(1):57–92, 2015.

Nuala Fitton, Peter Alexander, Nigel Arnell, Bojana Bajzelj, Katherine Calvin, Jonathan Doelman, James S Gerber, Petr Havlik, Tomoko Hasegawa, Mario Herrero, et al. The vulnerabilities of agricultural land and food production to future water scarcity. *Global Environmental Change*, 58:101944, 2019.

Gregory Flato, Jochem Marotzke, Babatunde Abiodun, Pascale Braconnot, Sin Chan Chou, William Collins, Peter Cox, Fatima Driouech, Seita Emori, Veronika Eyring, et al. Evaluation of climate models. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 741–866. Cambridge University Press, 2014.

Guobin Fu, Neil R Viney, and Stephen P Charles. Evaluation of various root transformations of daily precipitation amounts fitted with a normal distribution for australia. *Theoretical and applied climatology*, 99:229–238, 2010.

Liping Gao, Bo Tao, Yunxuan Miao, Lihua Zhang, Xia Song, Wei Ren, Liyuan He, and Xiaofeng Xu. A global data set for economic losses of extreme hydrological events during 1960-2014. *Water Resources Research*, 55(6):5165–5175, 2019.

Jordi Garcia-Gonzalo, Cristóbal Pais, Joanna Bachmatiuk, and Andrés Weintraub. Accounting for climate change in a forest planning stochastic optimization model. *Canadian Journal of Forest Research*, 46(9):1111–1121, 2016.

David EHJ Gernaat, Harmen Sytze de Boer, Vassilis Daioglou, Seleshi G Yalew, Christoph Müller, and Detlef P van Vuuren. Climate change impacts on renewable energy supply. *Nature Climate Change*, 11(2):119–125, 2021.

José Portela González, Antonio Muñoz San Muñoz San Roque, and Estrella Alonso Perez. Forecasting functional time series with a new hilbertian armax model: Application to electricity price forecasting. *IEEE Transactions on Power Systems*, 33(1):545–556, 2017.

Stelios Grafakos, Giulia Viero, Diana Reckien, Katie Trigg, Vincent Viguie, Andrew Sudmant, Catherine Graves, Aoife Foley, Oliver Heidrich, JM Mirailles, et al. Integration of mitigation and adaptation in urban climate change action plans in europe: A systematic assessment. *Renewable and Sustainable Energy Reviews*, 121:109623, 2020.

Jonathan M Gregory, Stephen M Griffies, Chris W Hughes, Jason A Lowe, John A Church, Ichiro Fukimori, Natalya Gomez, Robert E Kopp, Felix Landerer, Gonéri Le Cozannet,

et al. Concepts and terminology for sea level: Mean, variability and change, both local and global. *Surveys in Geophysics*, 40:1251–1289, 2019.

Ryan Hafen. *housingData: U.S. Housing Data from 2008 to 2016*, 2016. URL https://CRAN.R-project.org/package=housingData. R package version 0.3.0.

Stéphane Hallegatte. Strategies to adapt to an uncertain climate change. *Global environmental change*, 19(2):240–247, 2009.

Stéphane Hallegatte, Ankur Shah, Casey Brown, Robert Lempert, and Stuart Gill. Investment decision making under deep uncertainty–application to climate change. *World Bank Policy Research Working Paper*, (6193), 2012.

Lee Hannah, Guy F Midgley, and D Millar. Climate change-integrated conservation strategies. *Global Ecology and Biogeography*, 11(6):485–495, 2002.

MA Hannan, SB Wali, PJ Ker, MS Abd Rahman, M Mansor, VK Ramachandaramurthy, KM Muttaqi, TMI Mahlia, and ZY Dong. Battery energy-storage system: A review of technologies, optimization objectives, constraints, approaches, and outstanding issues. *Journal of Energy Storage*, 42:103023, 2021.

Charles R. Harris, K. Jarrod Millman, Stéfan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi,

Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585:357–362, 2020. doi: 10.1038/s41586-020-2649-2.

Ed Hawkins and Rowan Sutton. The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1108, 2009.

Robert J. Hijmans. *raster: Geographic Data Analysis and Modeling*, 2023. URL https://CRAN.R-project.org/package=raster. R package version 3.6-14.

Joy Hill, Jordan Kern, David E Rupp, Nathalie Voisin, and Gregory Characklis. The effects of climate change on interregional electricity market dynamics on the us west coast. *Earth's Future*, 9(12):e2021EF002400, 2021.

Ove Hoegh-Guldberg, Daniela Jacob, M Bindi, S Brown, I Camilloni, A Diedhiou, R Djalante, K Ebi, F Engelbrecht, J Guiot, et al. Impacts of 1.5 °C global warming on natural and human systems. *Global warming of 1.5 °C. An IPCC Special Report*, 2018.

Mohammad Shakhawat Hossain, Muhammad Arshad, Lu Qian, Harald Kächele, Imran Khan, Md Din Il Islam, and M Golam Mahboob. Climate change impacts on farmland value in bangladesh. *Ecological Indicators*, 112:106181, 2020.

IPCC. IPCC Fourth Assessment Report: Climate Change 2007, 2007. https://archive.ipcc.ch/publications_and_data/ar4/wg2/en/frontmattersg.html. Accessed on Oct/08/2023.

IPCC. The Intergovernmental Panel on Climate Change, 2023. https://www.ipcc.ch/. Accessed on July/27/2023.

ISO New England. Faqs: Locational marginal pricing, 2023. https://www.iso-ne.com/participate/support/faq/lmp. Accessed on October/05/2023.

Ian T Jolliffe. *Principal component analysis*. Springer, 2 edition, 2002.

A.D. Jones and C.P. Underwood. A thermal model for photovoltaic systems. *Solar Energy*, 70(4):349–359, 2001. ISSN 0038-092X. doi: https://doi.org/10.1016/S0038-092X(00)00149-3. URL https://www.sciencedirect.com/science/article/pii/S0038092X00001493.

Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311, 1984.

David L Kelly, Charles D Kolstad, and Glenn T Mitchell. Adjustment costs from environmental change. *Journal of Environmental Economics and Management*, 50(3):468–495, 2005.

Seama Koohi-Fayegh and Marc A Rosen. A review of energy storage types, applications and recent developments. *Journal of Energy Storage*, 27:101047, 2020.

Magnus Korpaas, Arne T Holen, and Ragne Hildrum. Operation and sizing of energy storage for wind power plants in a market system. *International Journal of Electrical Power & Energy Systems*, 25(8):599–606, 2003.

John Kotcher, Edward Maibach, Jeni Miller, Eryn Campbell, Lujain Alqodmani, Marina Maiero, and Arthur Wyns. Views of health professionals on climate change and health: a multinational survey study. *The Lancet Planetary Health*, 5(5):e316–e323, 2021.

Elmar Kriegler, Brian C O'Neill, Stephane Hallegatte, Tom Kram, Robert J Lempert, Richard H Moss, and Thomas Wilbanks. The need for and use of socio-economic scenarios for climate change analysis: a new approach based on shared socio-economic pathways. *Global Environmental Change*, 22(4):807–822, 2012.

GV Brahmendra Kumar and K Palanisamy. A review of energy storage participation for ancillary services in a microgrid environment. *Inventions*, 5(4):63, 2020.

Brian Kuns, Oane Visser, and Anders Wästfelt. The stock market and the steppe: The challenges faced by stock-market financed, nordic farming ventures in russia and ukraine. *Journal of Rural Studies*, 45:199–217, 2016.

Taesam Lee and Vijay P Singh. *Statistical downscaling for hydrological and environmental applications*. CRC press, 2018.

Flavio Lehner, Clara Deser, Nicola Maher, Jochem Marotzke, Erich M Fischer, Lukas Brunner, Reto Knutti, and Ed Hawkins. Partitioning climate projection uncertainty with multiple large ensembles and cmip5/6. *Earth System Dynamics*, 11(2):491–508, 2020.

Tania Murray Li. Transnational farmland investment: A risky business. *Journal of Agrarian Change*, 15(4):560–568, 2015.

Xin-Zhong Liang, You Wu, Robert G Chambers, Daniel L Schmoldt, Wei Gao, Chaoshun Liu, Yan-An Liu, Chao Sun, and Jennifer A Kennedy. Determining climate effects on us total agricultural productivity. *Proceedings of the National Academy of Sciences*, 114 (12):E2285–E2292, 2017.

Haiyu Liao, Bixiong Huang, Yan Cui, Huan Qin, Xintian Liu, and Huayuan Xu. Research on a fast detection method of self-discharge of lithium battery. *Journal of Energy Storage*, 55:105431, 2022.

Jian Liu, Xin-yue Sun, Rui Bo, Siyuan Wang, and Meng Ou. Economic dispatch for electricity merchant with energy storage and wind plant: State of charge based decision making considering market impact and uncertainties. *Journal of Energy Storage*, 53: 104816, 2022.

Heman D Lohano and Robert P King. A stochastic dynamic programming analysis of farmland investment and financial management. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 57(4):575–600, 2009.

Jason MacDonald, Peter Cappers, Duncan Callaway, and Sila Kiliccote. Demand response providing ancillary services: A comparison of opportunities and challenges in the US wholesale markets, 2012. https://gridwiseac.org/pdfs/forum_papers12/macdonald_paper_gi12.pdf. Accessed on October/05/2023.

André Magnan and Sean Sunley. Farmland investment and financialization in saskatchewan, 2003–2014: An empirical analysis of farmland transactions. *Journal of Rural Studies*, 49:92–103, 2017.

Emanuele Massetti, Robert Mendelsohn, and Shun Chonabayashi. How well do degree days over the growing season capture the effect of climate on farmland values? *Energy Economics*, 60:144–150, 2016.

Dylan McConnell, Tim Forcey, and Mike Sandiford. Estimating the value of electricity storage in an energy-only wholesale market. *Applied Energy*, 159:422–432, 2015.

Linda Mearns, Seth McGinnis, Daniel Korytina, Raymond Arritt, Sebastien Biner, Melissa Bukovsky, Hsin-I Chang, Ole Christensen, Daryl Herzmann, Yanjun Jiao, Slava Kharin, Michael Lazare, Grigory Nikulin, Minwei Qian, John Scinocca, Katja Winger, Chris Castro, Anne Frigon, and William Gutowski. The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, Boulder CO, 2017. https://doi.org/10.5065/D6SJ1JCH. Accessed on April/28/2022.

Robert Mendelsohn and Michelle Reinsborough. A ricardian analysis of us and canadian farmland. *Climatic Change*, 81(1):9–17, 2007.

Robert Mendelsohn, William D Nordhaus, and Daigee Shaw. The impact of global warming on agriculture: a ricardian analysis. *The American Economic Review*, pages 753–771, 1994.

Robert O Mendelsohn and Ariel Dinar. *Climate Change and Agriculture: An Economic Analysis of Global Impacts, Adaptation and Distributional Effects*. Edward Elgar Publishing, 2009.

Julia L Michalak, Joshua J Lawler, David R Roberts, and Carlos Carroll. Distribution and protection of climatic refugia in north america. *Conservation Biology*, 32(6):1414–1425, 2018.

Adrien Michel, Varun Sharma, Michael Lehning, and Hendrik Huwald. Climate change

scenarios at hourly time-step over switzerland from an enhanced temporal downscaling approach. *International Journal of Climatology*, 41(6):3503–3522, 2021.

ISO Midcontinent. Energy and operating reserve markets business practice manual, 2016.

Torben K Mideksa and Steffen Kallbekken. The impact of climate change on the electricity market: A review. *Energy policy*, 38(7):3579–3585, 2010.

Marija Miletić, Hrvoje Pandžić, and Dechang Yang. Operating and investment models for energy storage systems. *Energies*, 13(18):4600, 2020.

MISO. Buying and selling energy, 2022. https://www.misoenergy.org/stakeholder-engagement/training2/learning-center/market-basics/. Accessed on September/19/2022.

MISO. About MISO, 2023. https://www.misoenergy.org/about/. Accessed on August/31/2023.

MISO Energy. MISO data, 2023. https://www.misoenergy.org/markets-and-operations. (Accessed on April/6/2023).

MISO Inc. Map of MISO Zone boundaries, 2020. https://docs.misoenergy.org/legalcontent/Attachment_VV_-_MAP_of_Local_Resource_Zone_Boundaries.pdf. Accessed on October/06/2023.

Frank Moraes. 15 US states where agriculture drives the economy, 2022. Data retrieved from https://commodity.com/blog/state-economies-agriculture/. Accessed on July/17/2022.

Richard H Moss, Jae A Edmonds, Kathy A Hibbard, Martin R Manning, Steven K Rose, Detlef P Van Vuuren, Timothy R Carter, Seita Emori, Mikiko Kainuma, Tom Kram, et al. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756, 2010.

Mavuto M Mukaka. A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal*, 24(3):69–71, 2012.

Machiel Mulder and Bert Scholtens. The impact of renewable energy on electricity prices in the netherlands. *Renewable energy*, 57:94–100, 2013.

United States Department of Agriculture National Agricultural Statistics Service. 2017 census of agriculture, 2017. Data retrieved from https://quickstats.nass.usda.gov/#ACF3D153-DC4C-37FF-87B2-3419A0088827. Accessed on July/17/2022.

National Agricultural Statistics Service, United States Department of Agriculture. Ag land, incl buildings-asset value, measured in $/acre, 2017. Data retrieved from https://quickstats.nass.usda.gov/results/A0BD07A9-C755-3AE0-A070-CDF0E40E3DDF. Accessed on April/12/2021.

National Oceanic and Atmospheric Administration. Climate Models, 2023. https://www.climate.gov/maps-data/climate-data-primer/predicting-climate/climate-models. Accessed on Oct/08/2023.

Natural Resources Defense Council. MISO's clean energy dreams hinge on a state-of-the-art grid, 2023. https://www.nrdc.org/bio/natalie-mcintire/misos-clean-energy-dreams-hinge-state-art-grid. Accessed on August/31/2023.

Richard B Neale, Chih-Chieh Chen, Andrew Gettelman, Peter H Lauritzen, Sungsu Park, David L Williamson, Andrew J Conley, Rolando Garcia, Doug Kinnison, Jean-Francois Lamarque, et al. Description of the ncar community atmosphere model (cam 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, 1(1):1–12, 2010.

The Economist New frontiers. Climate change will alter where many crops are grown, 2021. https://www.economist.com/. Accessed on July/17/2022.

Chau Trinh Nguyen and Frank Scrimgeour. Measuring the impact of climate change on agriculture in vietnam: A panel ricardian analysis. *Agricultural Economics*, 53(1):37–51, 2022.

Tu A Nguyen, Raymond H Byrne, Ricky J Concepcion, and Imre Gyuk. Maximizing revenue from electrical energy storage in miso energy & frequency regulation markets. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2017.

Chao Ning and Fengqi You. Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, 125:434–448, 2019.

Nuveen. Nuveen knows responsible investing, 2018. https://www.tiaa.org/public/pdf/addressing_the_climate_change.pdf.

Office of Energy Efficiency and Renewable Energy . Nrel study identifies the opportunities and challenges of achieving the U.S. transformational goal of 100% clean electricity by 2035, 2022. https://www.energy.gov/eere/articles/

nrel-study-identifies-opportunities-and-challenges-achieving-us-transformational-goal. Accessed on August/31/2023.

Brian C O'Neill, Claudia Tebaldi, Detlef P Van Vuuren, Veronika Eyring, Pierre Friedlingstein, George Hurtt, Reto Knutti, Elmar Kriegler, Jean-Francois Lamarque, Jason Lowe, et al. The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9):3461–3482, 2016.

Ariel Ortiz-Bobea. The role of nonfarm influences in ricardian estimates of climate change impacts on us agriculture. *American Journal of Agricultural Economics*, 102(3):934–959, 2020.

Ariel Ortiz-Bobea, Toby R Ault, Carlos M Carrillo, Robert G Chambers, and David B Lobell. Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 11(4):306–312, 2021.

Rajendra K Pachauri, Myles R Allen, Vicente R Barros, John Broome, Wolfgang Cramer, Renate Christ, John A Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta, et al. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc, 2014.

Jérôme Pagès. *Multiple factor analysis by example using R*. CRC Press, 2014.

Marvin J Painter. Is farmland as good as gold? *Economics Research International*, 2011, 2011.

Camille Parmesan, Mike D Morecroft, and Yongyut Trisurat. *Climate change 2022: Impacts, adaptation and vulnerability*. PhD thesis, GIEC, 2022.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Andrés F Peñaranda, David Romero-Quete, and Camilo A Cortés. Grid-scale battery energy storage for arbitrage purposes: A colombian case. *Batteries*, 7(3):59, 2021.

David W Pierce, Daniel R Cayan, and Bridget L Thrasher. Statistical downscaling using localized constructed analogs (LOCA). *Journal of Hydrometeorology*, 15(6):2558–2585, 2014.

Andreas Poullikkas. A comparative overview of large-scale battery systems for electricity storage. *Renewable and Sustainable energy reviews*, 27:778–788, 2013.

Frederick Quaye, Denis Nadolnyak, and Valentina Hartarska. Climate change impacts on farmland values in the southeast united states. *Sustainability*, 10(10):3426, 2018.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

Md Mustafizur Rahman, Abayomi Olufemi Oni, Eskinder Gemechu, and Amit Kumar. Assessment of energy storage technologies: A review. *Energy Conversion and Management*, 223:113295, 2020.

Sanchayita Rajkhowa and Jyotirmoy Sarma. Climate change and flood risk, global climate change. In *Global climate change*, pages 321–339. Elsevier, 2021.

Md Masud Rana, Moslem Uddin, Md Rasel Sarkar, Sheikh Tanzim Meraj, GM Shafiullah, SM Muyeen, Md Ariful Islam, and Taskin Jamal. Applications of energy storage systems in power grids with and without renewable energy integration—a comprehensive review. *Journal of Energy Storage*, 68:107811, 2023.

Michelle J Reinsborough. A ricardian model of climate change in Canada. *Canadian Journal of Economics/Revue canadienne d'économique*, 36(1):21–40, 2003.

Keywan Riahi, Detlef P Van Vuuren, Elmar Kriegler, Jae Edmonds, Brian C O'neill, Shinichiro Fujimori, Nico Bauer, Katherine Calvin, Rob Dellink, Oliver Fricko, et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, 42:153–168, 2017.

CE Richards, RC Lupton, and Julian M Allwood. Re-framing the threat of global warming: an empirical causal loop diagram of climate change, food insecurity and societal collapse. *Climatic Change*, 164(3-4):49, 2021.

Gerard H Roe and Marcia B Baker. Why is climate sensitivity so unpredictable? *Science*, 318(5850):629–632, 2007.

I Santiago, A Moreno-Munoz, P Quintero-Jiménez, F Garcia-Torres, and MJ Gonzalez-Redondo. Electricity demand during pandemic times: The case of the covid-19 in spain. *Energy policy*, 148:111964, 2021.

João A Santos, Helder Fraga, Aureliano C Malheiro, José Moutinho-Pereira, Lia-Tânia Dinis, Carlos Correia, Marco Moriondo, Luisa Leolini, Camilla Dibari, Sergi Costafreda-Aumedes, et al. A review of the potential climate change impacts and adaptation options for european viticulture. *Applied Sciences*, 10(9):3092, 2020.

Pedro Santos, Igor Rezende, Tiago Soares, and Vladimiro Miranda. Evaluation of different bidding strategies for a battery energy storage system performing energy arbitrage–a neural network approach. In *2023 19th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2023.

Wolfram Schlenker, W Michael Hanemann, and Anthony C Fisher. The impact of global warming on us agriculture: an econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1):113–125, 2006.

Gary D Schnitkey, C Robert Taylor, and Peter J Barry. Evaluating farmland investments considering dynamic stochastic returns and farmland prices. *Western Journal of Agricultural Economics*, pages 143–156, 1989.

Uwe Schulzweida. Cdo user guide, October 2022. URL https://doi.org/10.5281/zenodo.7112925.

Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

S Niggol Seo, Robert Mendelsohn, Ariel Dinar, Rashid Hassan, and Pradeep Kurukulasuriya. A ricardian analysis of the distribution of climate change impacts on agriculture

across agro-ecological zones in africa. *Environmental and Resource Economics*, 43(3): 313–332, 2009.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

Robert H Shumway, David S Stoffer, and David S Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.

Ramteen Sioshansi, Paul Denholm, Thomas Jenkin, and Jurgen Weiss. Estimating the value of electricity storage in pjm: Arbitrage and some welfare effects. *Energy economics*, 31(2):269–277, 2009.

Shawn R Smith, Mark A Bourassa, and Ryan J Sharp. Establishing more truth in true winds. *Journal of Atmospheric and Oceanic Technology*, 16(7):939–952, 1999.

Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

Kepa Solaun and Emilio Cerdá. Climate change impacts on renewable energy generation. a review of quantitative projections. *Renewable and sustainable energy Reviews*, 116: 109415, 2019.

S&P Global. US renewables tracker: Miso takes over lead in q4 renewables output, wind soars, 2023. https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/electric-power/030323-us-renewables-tracker-miso-takes-over-lead-in-q4-renewables-output-wind-soars. Accessed on August/31/2023.

Alisa Spiegel, Wolfgang Britz, Utkur Djanibekov, and Robert Finger. Stochastic-dynamic modelling of farm-level investments under uncertainty. *Environmental Modelling & Software*, 127:104656, 2020.

Statista. Revenue of the electricity industry in the united states from 1970 to 2021, 2023. https://www.statista.com/statistics/190548/revenue-of-the-us-electric-power-industry-since-1970/. Accessed on August/31/2023.

Graeme L Stephens, Martin Wild, Paul W Stackhouse, Tristan L'Ecuyer, Seiji Kato, and David S Henderson. The global character of the flux of downward longwave radiation. *Journal of Climate*, 25(7):2329–2340, 2012.

Yang Su, Benoit Gabrielle, and David Makowski. The impact of climate change on the productivity of conservation agriculture. *Nature Climate Change*, 11(7):628–633, 2021.

Sustainable FERC. Updated interactive map shows clean energy projects withdrawn from miso queue, 2021. https://sustainableferc.org/wp-content/uploads/2021/01/SFP-MISO-Queue-Map-Update-2-pager-11-9-20.pdf. Accessed on September/26/2023.

Karl E Taylor, Ronald J Stouffer, and Gerald A Meehl. An overview of cmip5 and the experiment design. *Bulletin of the American meteorological Society*, 93(4):485–498, 2012.

Claudia Tebaldi, Kevin Debeire, Veronika Eyring, Erich Fischer, John Fyfe, Pierre Friedlingstein, Reto Knutti, Jason Lowe, Brian O'Neill, Benjamin Sanderson, et al. Cli-

mate model projections from the scenario model intercomparison project (scenariomip) of cmip6. *Earth System Dynamics Discussions*, 2020:1–50, 2020.

The Office of Energy Efficiency and Renewable Energy. How does solar work?, 2023. https://www.energy.gov/eere/solar/how-does-solar-work. Accessed on June/15/2023.

The United States Census Bureau. Land in farms (NAICS) 2002 (acres) (adjusted), 2012. Data retrieved from https://www2.census.gov/prod2/statcomp/usac/excel/AGN01.xls. Accessed on August/17/2022.

The World Bank. Agriculture and food, 2022. https://www.worldbank.org/en/topic/agriculture/overview. Accessed on July/17/2022.

Bridget Thrasher, Weile Wang, Andrew Michaelis, Forrest Melton, Tsengdar Lee, and Ramakrishna Nemani. Nasa global daily downscaled projections, CMIP6. *Scientific Data*, 9(1):262, 2022.

Léonard Tschora, Erwan Pierre, Marc Plantevit, and Céline Robardet. Electricity price forecasting on the day-ahead market using machine learning. *Applied Energy*, 313:118752, 2022.

Monica G Turner, W John Calder, Graeme S Cumming, Terry P Hughes, Anke Jentsch, Shannon L LaDeau, Timothy M Lenton, Bryan N Shuman, Merritt R Turetsky, Zak Ratajczak, et al. Climate change, ecosystems and abrupt change: science priorities. *Philosophical Transactions of the Royal Society B*, 375(1794):20190105, 2020.

UN Environmental Programme. Adaptation gap report 2022, 2022. https://www.unep.org/resources/adaptation-gap-report-2022. Accessed on October/22/2023.

U.S. Department of Agriculture. What is agriculture's share of the overall U.S. economy?, 2022. https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=58270. Accessed on July/17/2022.

Steven Van Passel, Emanuele Massetti, and Robert Mendelsohn. A ricardian analysis of the impact of climate change on european agriculture. *Environmental and Resource Economics*, 67(4):725–760, 2017.

Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009. ISBN 1441412697.

Detlef P Van Vuuren, Jae Edmonds, Mikiko Kainuma, Keywan Riahi, Allison Thomson, Kathy Hibbard, George C Hurtt, Tom Kram, Volker Krey, Jean-Francois Lamarque, et al. The representative concentration pathways: an overview. *Climatic Change*, 109: 5–31, 2011.

Mauricio Varas, Sergio Maturana, Rodrigo Pascual, Ignacio Vargas, and Jorge Vera. Scheduling production for a sawmill: A robust optimization approach. *International Journal of Production Economics*, 150:37–51, 2014.

Fernando Badilla Veliz, Jean-Paul Watson, Andres Weintraub, Roger J-B Wets, and David L Woodruff. Stochastic optimization models in forest planning: a progressive hedging solution approach. *Annals of Operations Research*, 232(1):259–274, 2015.

Oane Visser. Persistent farmland imaginaries: celebration of fertile soil and the recurrent ignorance of climate. *Agriculture and Human Values*, pages 1–14, 2020.

EM Volodin, NA Dianskii, and AV Gusev. Simulating present-day climate with the inmcm4. 0 coupled model of the atmospheric and oceanic general circulations. *Izvestiya, Atmospheric and Oceanic Physics*, 46:414–431, 2010.

Alexandra Von Meier. *Electric power systems: a conceptual introduction*. John Wiley & Sons, 2006.

Russell S Vose, Scott Applequist, Mike Squires, Imke Durre, Matthew J Menne, Claude N Williams Jr, Chris Fenimore, Karin Gleason, and Derek Arndt. NOAA's gridded climate divisional dataset (CLIMDIV), 2014a. Data retrieved from NOAA National Climatic Data Center https://doi:10.7289/V5M32STR. Accessed on April/12/2021.

Russell S Vose, Scott Applequist, Mike Squires, Imke Durre, Matthew J Menne, Claude N Williams Jr, Chris Fenimore, Karin Gleason, and Derek Arndt. Improved historical temperature and precipitation time series for us climate divisions. *Journal of Applied Meteorology and Climatology*, 53(5):1232–1251, 2014b.

Ronald E Walpole, Raymond H Myers, Sharon L Myers, and Keying Ye. *Probability and statistics for engineers and scientists*, volume 5. Macmillan New York, 1993.

Robert Walton. MISO opens energy, operating reserve markets to storage with over 13 GW in interconnection queue, 2022. https://www.utilitydive.com/news/miso-opens-reserve-markets-storage/631289/. (Accessed on June/24/2023).

WDC Climate. World Data Center for Climate, 2023. https://www.wdc-climate.de/ui/q?query. (Accessed on April/26/2023).

Albert Weiss and Cynthia J Hays. Calculating daily mean air temperatures by different methods: implications from a non-linear algorithm. *Agricultural and Forest Meteorology*, 128(1-2):57–65, 2005.

Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1081, 2014.

World Meteorological Organization. Updated 30-year reference period reflects changing climate, 2021. https://public.wmo.int/en/media/news/updated-30-year-reference-period-reflects-changing-climate. (Accessed on March/27/2023).

Wendong Yang, Shaolong Sun, Yan Hao, and Shouyang Wang. A novel machine learning-based electricity price forecasting model based on optimal model selection strategy. *Energy*, 238:121989, 2022.

Zhang Yang, Li Ce, and Li Lian. Electricity price forecasting by a hybrid model, combining wavelet transform, arma and kernel-based extreme learning machine methods. *Applied Energy*, 190:291–305, 2017.

Zhenguo Yang, Jun Liu, Suresh Baskaran, Carl H Imhoff, and Jamie D Holladay. Enabling renewable energy—and the future grid—with advanced electricity storage. *Jom*, 62(9): 14–23, 2010.

Dimitrios Zafirakis, Konstantinos J Chalvatzis, Giovanni Baiocchi, and Georgios Daskalakis. The value of arbitrage for energy storage: Evidence from european electricity markets. *Applied energy*, 184:971–986, 2016.

Ying Zhang and Jayashankar M Swaminathan. Improved crop productivity through optimized planting schedules. *Manufacturing & Service Operations Management*, 22(6): 1165–1180, 2020.

Yangfang Zhou, Alan Scheller-Wolf, Nicola Secomandi, and Stephen Smith. Managing wind-based electricity generation in the presence of storage and transmission capacity. *Production and Operations Management*, 28(4):970–989, 2019.

Florian Ziel and Rick Steinert. Probabilistic mid-and long-term electricity price forecasting. *Renewable and Sustainable Energy Reviews*, 94:251–266, 2018.

APPENDICES

Appendix A

Appendix: Chapter 2

A.1 Histogram of Farmland Prices

Figure A.1a presents a histogram of the discounted farmland price data for the years 1997, 2002, 2007, 2012, and 2017, as mentioned in Section 2.4.2. These data exhibit right skewness with a long right tail. Some of these samples exceed $10,000, with two exceeding $25,000.

Figure A.1b shows a histogram plot of the log-transformed (discounted) farmland price data. The histogram indicates that the log-transformed farmland prices are less skewed and exhibit symmetry around the value 8. These patterns indicate a logarithmic regression model is preferable in our case.

A.2 Proof of Theorem 1

We rely on duality to provide linear transformations for both the non-linear objective function (2.4a) and the non-linear constraints (2.4b) into their corresponding linear forms. These transformations effectively convert non-convex Model (2.4) into its linear equivalences. We depend on strong duality to demonstrate that these linear equivalents generate the same optimal objective value as the optimal objective value of Model (2.4).

First, we convert the objective function (2.4a) of Model (2.4) to a linear equivalence. The objective function can be simplified as follows:

$$\max_{\{b_{tc},s_{tc}|\forall t\in\mathcal{T}\setminus\{T\},c\in\mathcal{C}\}} \Big\{ \sum_{c\in\mathcal{C}\setminus\{m\}} \bar{p}_{Tc} \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tc}-s_{tc}) + p_{Tm} \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tm}-s_{tm}) + \min_{\{z_{Tc}|c\in\mathcal{C}\setminus\{m\}\}} - \sum_{c\in\mathcal{C}\setminus\{m\}} z_{Tc}\hat{p}_{Tc} \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tc}-s_{tc}) \Big\} \tag{A.1}$$

Let $\boldsymbol{b}_t = (b_{t1}, b_{t2}, ..., b_{t(m-1)})$ and $\boldsymbol{s}_t = (s_{t1}, s_{t2}, ..., s_{t(m-1)})$ $(\forall t \in \mathcal{T} \setminus \{T\})$. Given any $\boldsymbol{b} = (\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_{T-1})$ and $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, ..., \boldsymbol{s}_{T-1})$, we define:

$$\gamma_T(\boldsymbol{b},\boldsymbol{s}) = \Big\{ \min_{\{z_{Tc}|c\in\mathcal{C}\setminus\{m\}\}} - \sum_{c\in\mathcal{C}\setminus\{m\}} \hat{p}_{Tc} z_{Tc} \sum_{t\in\mathcal{T}\setminus\{T\}} (b_{tc}-s_{tc}) \Big| \sum_{c\in\mathcal{C}\setminus\{m\}} z_{Tc} \le \Gamma_T , 0 \le z_{Tc} \le 1, \forall c \in \mathcal{C}\setminus\{m\} \Big\} \tag{A.2}$$

, where the dual of this problem is:

$$\max_{\{h_T, q_{Tc} | c \in \mathcal{C} \setminus \{m\}\}} \quad \Gamma_T h_T + \sum_{c \in \mathcal{C} \setminus \{m\}} q_{Tc} \tag{A.3a}$$

$$s.t. \quad h_T + q_{Tc} \leq -\hat{p}_{Tc} \sum_{t \in \mathcal{T} \setminus T} (b_{tc} - s_{tc}) \quad \forall c \in \mathcal{C} \setminus \{m\} \tag{A.3b}$$

$$h_T \leq 0 \tag{A.3c}$$

$$q_{Tc} \leq 0 \quad \forall c \in \mathcal{C} \setminus \{m\} \tag{A.3d}$$

As we want the variables h_T and q_{Tc} to be nonnegative, we can further transform Problem (A.3) as follows:

$$\max_{\{h_T, q_{Tc} | c \in \mathcal{C} \setminus \{m\}\}} \quad -\Gamma_T h_T - \sum_{c \in \mathcal{C} \setminus \{m\}} q_{Tc} \tag{A.4a}$$

$$s.t. \quad h_T + q_{Tc} \geq \hat{p}_{Tc} \sum_{t \in \mathcal{T} \setminus T} (b_{tc} - s_{tc}) \quad \forall c \in \mathcal{C} \setminus \{m\} \tag{A.4b}$$

$$h_T \geq 0 \tag{A.4c}$$

$$q_{Tc} \geq 0 \quad \forall c \in \mathcal{C} \setminus \{m\} \tag{A.4d}$$

Similarly, we can convert the constraints (2.4b) to linear constraints. In fact, constraints

(2.4b) can be simplified as follows:

$$\sum_{c\in\mathcal{C}\setminus\{m\}}[(1+r_b^l)\bar{p}_{tc}b_{tc}-(1-r_s^l)\bar{p}_{tc}s_{tc}]+(1+r_b^\beta)p_{tm}b_{tm}-(1-r_s^\beta)p_{tm}s_{tm}$$
$$+\max_{\{z_{tc}|c\in\mathcal{C}\setminus\{m\}\}}\Big\{\sum_{c\in\mathcal{C}\setminus\{m\}}[(1+r_b^l)\hat{p}_{tc}b_{tc}+(1-r_s^l)\hat{p}_{tc}s_{tc}]z_{tc}\Big\}\leq B_t\quad\forall t\in\mathcal{T}\setminus\{T\}\tag{A.5}$$

For each $t\in\mathcal{T}\setminus\{T\}$, we define:

$$\gamma_t(\boldsymbol{b}_t,\boldsymbol{s}_t)=\max_{\{z_{tc}|c\in\mathcal{C}\setminus\{m\}\}}\quad\sum_{c\in\mathcal{C}\setminus m}[(1+r_b^l)\hat{p}_{tc}b_{tc}+(1-r_s^l)\hat{p}_{tc}s_{tc}]z_{tc}\tag{A.6a}$$
$$s.t.\quad\sum_{c\in\mathcal{C}\setminus m}z_{tc}\leq\Gamma_t\tag{A.6b}$$
$$0\leq z_{tc}\leq 1\quad\forall c\in\mathcal{C}\setminus\{m\}\tag{A.6c}$$

, where the dual of this problem is:

$$\min_{\{h_t,q_{tc}|c\in\mathcal{C}\setminus\{m\}\}}\quad\Gamma_t h_t+\sum_{c\in\mathcal{C}\setminus\{m\}}q_{tc}\tag{A.7a}$$
$$s.t.\quad h_t+q_{tc}\geq(1+r_b^l)\hat{p}_{tc}b_{tc}+(1-r_s^l)\hat{p}_{tc}s_{tc}\quad\forall c\in\mathcal{C}\setminus\{m\}\tag{A.7b}$$
$$h_t\geq 0\tag{A.7c}$$
$$q_{tc}\geq 0\quad\forall c\in\mathcal{C}\setminus\{m\}\tag{A.7d}$$

Since Model (A.6) is feasible and bounded for all $\Gamma_t\in[0,m-1]$, by the strong duality, both Model (A.6) and (A.7) have the same optimal objective value. We have that $\gamma_t(\boldsymbol{b}_t,\boldsymbol{s}_t)$ equals to the objective function value of Model (A.7).

By replacing Model (A.1) and (A.4) in (2.4a), and replacing Model (A.5) and (A.7) in (2.4b), Model (2.4) is equivalent to Model (2.5). □

A.3 Proof of Lemma 1

The central concept of the proof relies on the utilization of a counterexample to demonstrate that $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ is not an optimal solution in cases where land buying and selling occur within the same county and during the same period.

For notational simplicity, we let $\boldsymbol{b}^* = (b^*_{11}, b^*_{12}, ..., b^*_{1m}, b^*_{21}, ..., b^*_{(T-1)m})$, $\boldsymbol{s}^* = (s^*_{11}, s^*_{12}, ..., s^*_{1m}, s^*_{21}, ..., s^*_{(T-1)m})$, $\boldsymbol{h}^* = (h^*_1, h^*_2, ..., h^*_T)$, and $\boldsymbol{q}^* = (q^*_{11}, q^*_{12}, ..., q^*_{1(m-1)}, q^*_{21}, ..., q^*_{T(m-1)})$. Assume $\exists\, u \in \mathcal{T}\backslash\{T\}$ and $\exists\, v \in \mathcal{C}$ in the optimal solution $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ such that $b_{uv} \cdot s_{uv} > 0$. Then, $b_{uv} > 0$ and $s_{uv} > 0$.

Case 1: $v \in \mathcal{C}/\{m\}$. Let e_{ij} ($i \in \mathcal{T} \setminus \{T\}$ and $j \in \mathcal{C}$) be a $(T-1)m$-dimensional row vector with $(i \cdot j)$-th element being 1 and others being 0. Let $\underline{\boldsymbol{b}} = \boldsymbol{b}^* - e_{uv} min\{b^*_{uv}, s^*_{uv}\} + e_{um} \frac{(r^l_b + r^l_s)\bar{p}_{uv} min\{b^*_{uv}, s^*_{uv}\}}{(1+r^{\beta}_b)p_{um}}$ and $\underline{\boldsymbol{s}} = \boldsymbol{s}^* - e_{uv} min\{b^*_{uv}, s^*_{uv}\}$. We can easily see that $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ satisfy (2.5c), (2.5d), (2.5e), (2.5f), (2.5g), and (2.5h). Regarding the constraints (2.5b), if

$t \in \mathcal{T}/\{u, T\}$, the constraints hold for $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$; if $t = u$, we have:

$$\sum_{c \in \mathcal{C} \setminus \{v,m\}} [(1 + r_b^l)\bar{p}_{uc} b_{uc}^* - (1 - r_s^l)\bar{p}_{uc} s_{uc}^*]$$
$$+ (1 + r_b^l)\bar{p}_{uv}(b_{uv}^* - min\{b_{uv}^*, s_{uv}^*\}) - (1 - r_s^l)\bar{p}_{uv}(s_{uv}^* - min\{b_{uv}^*, s_{uv}^*\})$$
$$+ (1 + r_b^\beta)p_{um}(b_{um}^* + \frac{(r_b^l + r_s^l)\bar{p}_{uv} min\{b_{uv}^*, s_{uv}^*\}}{(1 + r_b^\beta)p_{um}}) - (1 - r_s^\beta)p_{um} s_{um}^* + \Gamma_u h_u^* + \sum_{c \in \mathcal{C} \setminus \{m\}} q_{uc}^* \tag{A.8}$$

$$= \sum_{c \in \mathcal{C} \setminus \{m\}} [(1 + r_b^l)\bar{p}_{uc} b_{uc}^* - (1 - r_s^l)\bar{p}_{uc} s_{uc}^*] + (1 + r_b^\beta)p_{um} b_{um}^* - (1 - r_s^\beta)p_{um} s_{um}^*$$
$$+ \Gamma_u h_u^* + \sum_{c \in \mathcal{C} \setminus \{m\}} q_{uc}^* \leq B_u \tag{A.9}$$

Therefore, $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ satisfies the constraints (2.5b) and is a feasible solution for Model (2.5).

Then, we want to prove that $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ gives a higher objective value than $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$, i.e., the latter is not an optimal solution. By replacing $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ with $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ in the objective function (2.5a), the value of the second item of the objective function (i.e., $p_{Tm} \sum_{t \in \mathcal{T} \setminus \{T\}} (b_{tm} - s_{tm})$) increases by $p_{Tm} \frac{(r_b^l + r_s^l)\bar{p}_{uv} min\{b_{uv}^*, s_{uv}^*\}}{(1 + r_b^\beta)p_{tm}} > 0$ (as $\bar{p}_{uv}, b_{uv}^*, s_{uv}^*, p_{Tm}, r_b^l$, and $r_s^l > 0$) and other items of the objective function remain the same. Therefore, $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ is not the optimal solution for Model (2.5), conflicting with the initial condition that it is the optimal solution.

Case 2: $v = m$. Let $\underline{\boldsymbol{b}} = \boldsymbol{b}^* + e_{um}(-min\{b_{um}^*, s_{um}^*\} + \frac{(r_b^\beta + r_s^\beta) min\{b_{um}^*, s_{um}^*\}}{1 + r_b^\beta})$ and $\underline{\boldsymbol{s}} = \boldsymbol{s}^* - e_{um} min\{b_{tm}^*, s_{tm}^*\}$. We can easily see that $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ satisfies (2.5c), (2.5d), (2.5e), (2.5f), (2.5g), and (2.5h). Regarding the constraints (2.5b) and given $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$, if

$t \in \mathcal{T}/\{u, T\}$, the constraints hold for $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$; if $t = u$, we have:

$$\sum_{c \in \mathcal{C}\setminus\{m\}} [(1 + r_b^l)\bar{p}_{uc} b_{uc}^* - (1 - r_s^l)\bar{p}_{uc} s_{uc}^*] + (1 + r_b^\beta) p_{um} (b_{um}^* - min\{b_{um}^*, s_{um}^*\}$$
$$+ \frac{(r_b^\beta + r_s^\beta) min\{b_{um}^*, s_{um}^*\}}{1 + r_b^\beta}) - (1 - r_s^\beta) p_{um} (s_{um}^* - min\{b_{um}^*, s_{um}^*\}) + \Gamma_u h_u^* + \sum_{c \in \mathcal{C}\setminus\{m\}} q_{uc}^* \tag{A.10}$$

$$= \sum_{c \in \mathcal{C}\setminus\{m\}} [(1 + r_b^l)\bar{p}_{uc} b_{uc}^* - (1 - r_s^l)\bar{p}_{uc} s_{uc}^*] + (1 + r_b^\beta) p_{um} b_{um}^* - (1 - r_s^\beta) p_{um} s_{um}^* + \Gamma_u h_u^*$$
$$+ \sum_{c \in \mathcal{C}\setminus\{m\}} q_{uc}^* \leq B_u \tag{A.11}$$

Therefore, $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ satisfies the constraints (2.5b) and is a feasible solution for Model (2.5).

Then, we want to prove that $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ gives higher objective value than $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$, i.e., the latter is not an optimal solution. By replacing $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ with $(\underline{\boldsymbol{b}}, \underline{\boldsymbol{s}}, \boldsymbol{h}^*, \boldsymbol{q}^*)$ in the objective function (2.5a) and given $r_b^\beta + r_s^\beta > 0$, the value of the second item of the objective function (i.e., $p_{Tm} \sum_{t \in \mathcal{T}\setminus\{T\}} (b_{tm} - s_{tm})$) increases by $p_{Tm} \frac{(r_b^\beta + r_s^\beta) \cdot min\{b_{tm}^*, s_{tm}^*\}}{1 + r_b^\beta} > 0$ (as $b_{um}^*, s_{um}^*, p_{Tm} > 0$) and other items of the objective function stay the same. Therefore, $(\boldsymbol{b}^*, \boldsymbol{s}^*, \boldsymbol{h}^*, \boldsymbol{q}^*)$ is not the optimal solution for Model (2.5), conflicting with the initial condition that it is the optimal solution. □

A.4 Generating Seasonal Temperature and Precipitation

We obtain the seasonal temperature and precipitation values from the monthly historical weather observations and climate projections, as detailed in Section 2.4.2, in the following two steps. First, using the relevant exponential moving averages, we smooth monthly historical observations and future projections for temperature and precipitation. Second, using the seasonal accumulative averages, we further smooth the data from the first step to get seasonal average observations and projections for temperature and precipitation.

A.4.1 Monthly Smoothing

We exponentially smooth the monthly average temperature and monthly total precipitation in each month in each county in the MRB. We index the 12 months from January to December with $o \in \{1, 2, .., 12\}$. We let $T_{c,o,t}$ (resp. $R_{c,o,t}$) be monthly temperature (resp. precipitation) at county c ($c \in \mathcal{C} \setminus \{m\}$) in month o in year t ($t \in \{1997, 1998, ..., 2099\}$), where "resp." denotes respectively. Let $\hat{T}_{c,o,t}$ (resp. $\hat{R}_{c,o,t}$) be the smoothed monthly temperature (resp. precipitation). Let b_0 ($0 < b_0 < 1$) be a discount rate, η be a positive scalar, and K be a smoothing window size. We express the monthly exponential smoothing formulas for temperature and precipitation as follows:

$$\hat{T}_{c,o,t} = \sum_{k=0}^{K-1} \eta(1-b_0)^k T_{c,o,t+k}, \forall c \in \mathcal{C} \setminus \{m\}, o \in \{1, 2, .., 12\}, t \in \{1997, 1998, ..., 2099\} \quad \text{(A.12)}$$

$$\hat{R}_{c,o,t} = \sum_{k=0}^{K-1} \eta(1-b_0)^k R_{c,o,t+k}, \forall c \in \mathcal{C} \setminus \{m\}, o \in \{1,2,..,12\}, t \in \{1997,1998,...,2099\} \quad \text{(A.13)}$$

The temperature (resp. precipitation) is recalculated using the exponential moving average, in which the weights decrease exponentially as temperature (resp. precipitation) comes from further in the future — the furthest temperature (resp. precipitation) has the smallest weight. We use (A.12) and (A.13) to smooth temperature and precipitation observations over 1997 and 2020 to get monthly smoothed historical observations. In practice, we let $\hat{K} = \min\{2020 - t + 1, K\}$ as a new window size to smooth each year in the period. Similarly, we smooth temperature and precipitation projections over 2023 and 2099 to get monthly smoothed projections. In application, we choose $\eta = \frac{1}{\sum_{k=0}^{K-1}(1-b_0)^k}$ being a normalization factor (letting the total weights to be one), $b_0 = 0.02$ (same as the inflation rate), and $K = 10$.

A.4.2 Seasonal Smoothing

We further seasonally average the monthly smoothed temperature and precipitation by using the cumulative average. Index the four seasons from spring to winter with $i \in I = \{1,...,4\}$. Let $\tilde{T}_{c,i,t}$ and $\tilde{R}_{c,i,t}$ be seasonal temperature and precipitation at county c in season i in year t. We express the seasonal-average temperature and precipitation as follows:

$$\tilde{T}_{c,i,t} = \frac{1}{3}\sum_{k=0}^{2} \hat{T}_{c,3i+k,t} \quad \forall c \in \mathcal{C} \setminus \{m\}, i \in \mathcal{I}, t \in \{1997,1998,...,2099\} \quad \text{(A.14)}$$

$$\tilde{R}_{c,i,t} = \frac{1}{3}\sum_{k=0}^{2} \hat{R}_{c,3i+k,t} \quad \forall c \in \mathcal{C} \setminus \{m\}, i \in \mathcal{I}, t \in \{1997, 1998, ..., 2099\} \tag{A.15}$$

, where $\hat{T}_{c,13,t} = \hat{T}_{c,1,t}$, $\hat{T}_{c,14,t} = \hat{T}_{c,2,t}$, $\hat{R}_{c,13,t} = \hat{R}_{c,1,t}$, and $\hat{R}_{c,14,t} = \hat{R}_{c,2,t}$ for notational convenience.

We use (A.14) and (A.15) for two things: 1) to seasonally smooth historical data for model fitting and 2) to seasonally smooth climate projections for model prediction.

A.5 Economic Indicator Parameters

For solving purposes, we collected three reliable economic indicators for the parameters of Model (2.5). First, we set the buying rate and selling rate at 0.01 and 0.06 respectively based on Lohano and King (2009). Second, we set the constant inflation rate β as 0.02 during 2023-2090 based on Board of Governors of the Federal Reserve System (2021). Third, we assume the buying cost rate on risk-free assets, i.e., r_b^β, is 0 and the selling cost rate on risk-free assets, i.e., r_s^β, is 0.01. We set the external income in period t as $B_t = 60,000$ USD ($\forall t \in \mathcal{T} \setminus \{T\}$). We summarize the values of parameters used in Model (2.5) in Table A.1.

A.6 Cardinality

We summarize the cardinality of the data used in the MRB case in Table A.2.

Table A.1: Parameter values

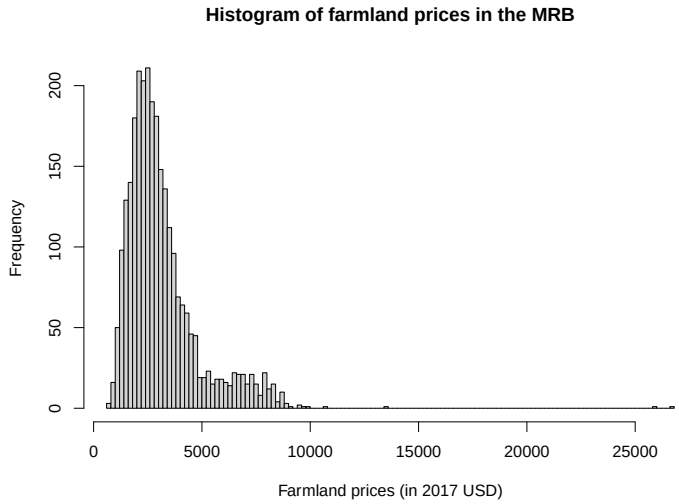
Parameters	Value
B_t	60,000
β	0.02
r_b^l	0.01
r_s^l	0.06
r_b^β	0
r_s^β	0.01

Table A.2: The cardinality of used data

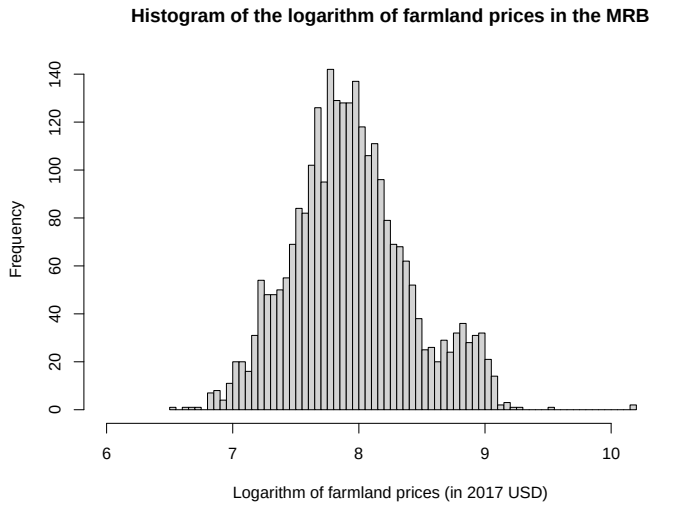
Set	Cardinality
$\mathcal{C}$	545 + 1 (1 represents the artificial county m)
$\mathcal{T}$	68 (from year 2023 to year 2090)

A.7 Marginal Impacts in 2090

We present the marginal impacts of mean climate on farmland values in the MRB in 2090 under RCP4.5 and RCP8.5 in Figure A.2 and A.3, respectively.

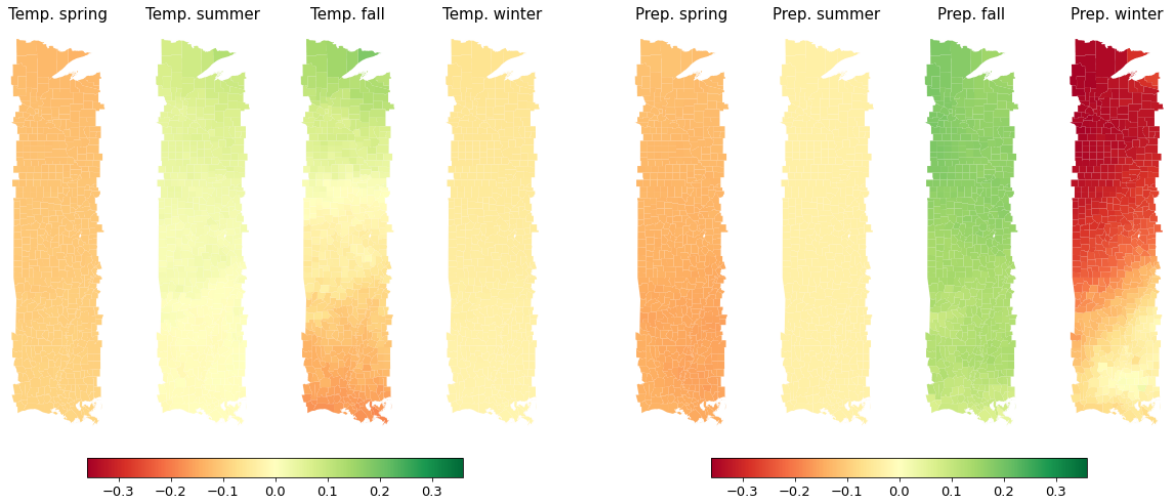


(a) Original



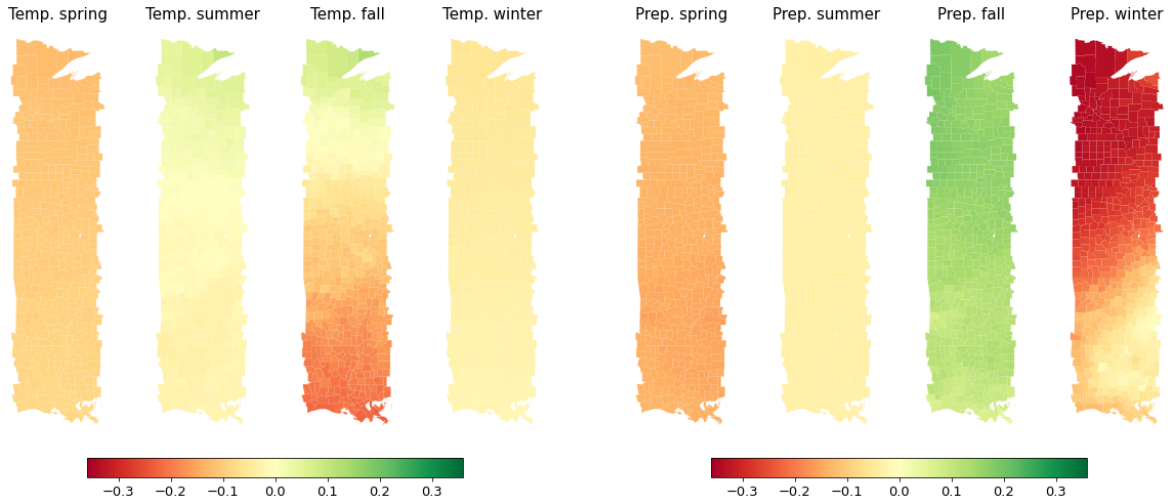
(b) Log-transformed

Figure A.1: Histogram of original and log-transformed farmland prices in the MRB



(a) Temperature (Temp.) (b) Precipitation (Prep.)

Figure A.2: Marginal impact in percentage of farmland values of mean climate in 2090 under RCP4.5



(a) Temperature (Temp.) (b) Precipitation (Prep.)

Figure A.3: Marginal impact in percentage of farmland values of mean climate in 2090 under RCP8.5

Appendix B

Appendix: Chapter 3

B.1 Proof of Lemma 2

Proof: We prove the Lemma through a contradiction. Assume that there is a $\mu \in \mathcal{T}$ such that both $q_\mu^C > 0$ and $q_\mu^D > 0$.

Case 1: $\eta^C q_t^{C\star} > q_t^{D\star}$. Let $\underline{q_t}^C = q_t^{C\star} - q_t^{D\star}/\eta^C$ if $t = \mu$ and $\underline{q_t}^C = q_t^{C\star}$ if $t \in \mathcal{T}\backslash\mu$. Similarly, let $\underline{q_t}^D = 0$ if $t = \mu$ and $\underline{q_t}^D = q_t^{D\star}$ if $t \in \mathcal{T}\backslash\mu$. We can easily see that the solution $\{\underline{q_t}^C, \underline{q_t}^D, \forall t \in \mathcal{T}\}$ satisfy Constraints (3.3b), (3.3d), and (3.3e). By replacing the solution $\{q_t^{C\star}, q_t^{D\star}, \forall t \in \mathcal{T}\}$ with $\{\underline{q_t}^C, \underline{q_t}^D, \forall t \in \mathcal{T}\}$ in the objective function (3.3a), the objective function increases by $q_t^{C\star}[(\bar{p}_t + \hat{p}_t)/\eta^C - (\bar{p}_t - \hat{p}_t)\eta^D] > 0$. Therefore, $\{q_t^{C\star}, q_t^{D\star}, \forall t \in \mathcal{T}\}$ is not the optimal solution for Model (3.3), conflicting with the initial condition that it is the optimal solution.

Case 2: $\eta^C q_t^{C\star} \leq q_t^{D\star}$. Let $\underline{q_t}^C = 0$ if $t = \mu$ and $\underline{q_t}^C = q_t^{C\star}$ if $t \in \mathcal{T}\backslash\mu$. Similarly, let

$\underline{q_t}^D = q_t^{D\star} - \eta^C q_t^{C\star}$ if $t = \mu$ and $\underline{q_t}^D = q_t^{D\star}$ if $t \in \mathcal{T} \backslash \mu$. We can easily see that the solution $\{\underline{q_t}^C, \underline{q_t}^D, \forall t \in \mathcal{T}\}$ satisfy Constraints (3.3b), (3.3d), and (3.3e). By replacing the solution $\{q_t^{C\star}, q_t^{D\star}, \forall t \in \mathcal{T}\}$ with $\{\underline{q_t}^C, \underline{q_t}^D, \forall t \in \mathcal{T}\}$ in the objective function (3.3a), the objective function increases by $q_t^{C\star}[(\bar{p}_t + \hat{p}_t) - (\bar{p}_t - \hat{p}_t)\eta^C\eta^D] > 0$. Therefore, $\{q_t^{C\star}, q_t^{D\star}, \forall t \in \mathcal{T}\}$ is not the optimal solution for Model (3.3), conflicting with the initial condition that it is the optimal solution. □

B.2 Regression based on Principle Component Analysis

Let X be an $n \times m$-dimensional matrix representing the dataset, where n is the number of samples, and each sample has m features. Let L be a positive integer. Let W be an $m \times L$-dimensional matrix, which we refer to as the transformation matrix (see Section 3.4.4). Let $\boldsymbol{\beta}$ be an L-dimensional vector. Let $\boldsymbol{y}$ be an n-dimensional vector. Let $\boldsymbol{a}$ be an n-dimensional vector with all entries being constant a. Let I be an $n \times k$-dimensional (dummy) matrix with all entries being 0 or 1, where k is the number of dummy variables. Let $\boldsymbol{\beta}_I$ be a k-dimensional vector. Let $\boldsymbol{\epsilon}$ be an n-dimensional vector representing random errors. The regression model based on the Principle Component Analysis used in Section 3.4.4 can be expressed as follows:

$$\boldsymbol{y} = \boldsymbol{a} + XW\boldsymbol{\beta} + I\boldsymbol{\beta}_I + \boldsymbol{\epsilon} \tag{B.1}$$

Let $\hat{\boldsymbol{\beta}}$ be the estimator of $\boldsymbol{\beta}$ using the ordinary least squares (OLS) method provided by the Python package *statsmodels* (Seabold and Perktold, 2010).

We denote B as the covariance matrix for $\boldsymbol{\beta}$. To estimate B, we use the Python package *statsmodels* and denote the estimator as $\hat{B}$. The coefficients for X can be expressed as $W\hat{\boldsymbol{\beta}}$ and are referred to as the (estimated) original coefficients.

We want to calculate the standard variance for each of the four climate variables (i.e., wind, precipitation, temperature, and (shortwave) radiation). With loss of generality, let

$$\begin{bmatrix} W_w \\ W_p \\ W_t \\ W_r \end{bmatrix} \tag{B.2}$$

be a row division of matrix W, where W_w is the transformation matrix for wind, W_p is the transformation matrix for precipitation, W_t is the transformation matrix for temperature, and W_r is the transformation matrix for (shortwave) radiation. Since we use each climate variable for the same MISO counties, the divided four transformation matrices have the same dimensions, each being $\frac{m}{4} \times L$, where $\frac{m}{4}$ is a positive integer.

Let $\mathbf{1}$ be an $\frac{m}{4}$-dimensional vector with all entries set to 1. We can estimate the mean of the original coefficients for each climate variable as $\frac{4}{m}\mathbf{1}^\top W_i\hat{\boldsymbol{\beta}}$, where $i \in \{w, t, p, r\}$.

Next, we can calculate the standard variance of the mean of original coefficients for

each climate variable as follows:

$$[Var(\frac{4}{m}\mathbf{1}^\top W_i\boldsymbol{\beta})]^{\frac{1}{2}} = \frac{4}{m}[\mathbf{1}^\top W_i B W_i^\top \mathbf{1}]^{\frac{1}{2}}, \quad i \in \{w, t, p, r\} \tag{B.3}$$

, where Var represents the variance. We can estimate the numerical value of the standard variance as $\frac{4}{m}[\mathbf{1}^\top W_i \hat{B} W_i^\top \mathbf{1}]^{\frac{1}{2}}$, where $i \in \{w, t, p, r\}$.