The Economically Optimal Nitrogen Rate for Spring Wheat Production in West-Central Saskatchewan

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Agricultural and Resource Economics
University of Saskatchewan
Saskatoon

By
Aiden Sanden

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

DISCLAIMER

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Agricultural and Resource Economics
51 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A8
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

Wheat is an essential crop for global food security serving as a staple food crop in households worldwide, providing a significant portion of the daily calorie needs. While significant wheat yield advancements have been achieved over time, a gap between farmer-realized wheat yields and their genetic potential persists. Several studies have estimated that average wheat yields range from 20 to 70% of their potential, suggesting that significant contributions to global wheat production are possible. Additionally, many yield gap studies have identified fertilizer deficiencies as a major contributor to the yield gap. These studies, however, have failed to examine if it is economically desirable at the farm level to increase fertilizer use in order to reduce the yield gap. This study examines how much of the yield gap for spring wheat in west-central Saskatchewan is economically exploitable with respect to synthetic nitrogen fertilizer use. A spring wheat yield response function was estimated using field-level yield, input, and management data to determine the impact of various input levels and management characteristics on yields. Focusing specifically on nitrogen use, observed nitrogen rates were compared to estimated economically optimal nitrogen rates to determine that spring wheat fields on the most productive soils were observed having received suboptimal nitrogen rates while those on the least productive soils had received nitrogen rates beyond what was estimated to be economically optimal. Spring wheat produced on fields that followed a pulse crop exhibited the largest economically exploitable yield gap, indicating an increased yield response to applied nitrogen for subsequent crops following pulses. Producers may be able to reduce a portion of the yield gap for wheat in west-central Saskatchewan by increasing nitrogen rates when producing wheat on pulse stubble. Observed application rates of nitrogen appeared to follow recommended agronomic yield targets rather than yield potential. Additionally, spring wheat variety and the management ability of the producer significantly impacted estimated spring wheat yields.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Richard Gray, for his generous mentorship, support, and patience throughout the process of writing this thesis. I am grateful for the funding provided by Dr. Gray's Canadian Grain Policy Research Chair that supported my research. I am also grateful for funding provided by the Alberta Wheat Commission to include an economic analysis of the yield gap for wheat in western Canada.

I would like to thank my committee members Dr. Patrick Lloyd-Smith, Dr. Steve Shirtliffe, and Dr. Tristan Skolrud for their valuable contributions and feedback. I would also like to thank my external examiner, Dr. Harvey Brooks, for his valuable insights during my defence.

I am grateful to the Saskatchewan Crop Insurance Corporation (SCIC) for providing access to the detailed management and yield data that made this research possible and to Arlan Frick of (SCIC) who answered all my data related questions. I would also like to thank Lampros-Nikolas Maros, Michelle Ross, and Feryel Lassoued for their insights and support throughout the process of writing this thesis.

Finally, I would like to thank my family for their support and encouragement throughout this entire process. You all believed in me when I did not. I am especially thankful to my wife, Carley, for all your love, patience, and support throughout this process – I couldn't have done this without you.

TABLE OF CONTENTS

LIST OF TABLES

Table 4.1A: Provincial, SCIC, and SMP Representative Hectares, CWRS, 2011-2019. 32
Table 4.1B: Provincial, SCIC, and SMP Representative Hectares, CWRS, 2011-2019. 33
Table 4.2: Description of variables included in production function. 35
Table 4.3: Mean yield of Spring-wheat by year (kg ha^{-1}), RZ 16. 36
Table 4.4: Mean applied rate of Nitrogen by year (kg ha^{-1}), RZ 16. 38
Table 4.5: Mean applied rate of Phosphorous, Potassium, and Sulphur by year (kg ha^{-1}), RZ 16. 39
Table 4.6: Average Seeding Date (days before/after May 14), 2011-2019. 41

LIST OF FIGURES

Chapter 1: Introduction

1.1 Background

It is estimated that additional production of nearly 1 billion metric tonnes of cereal grains per year will be needed by 2050 to feed the world's rising population (Alexandratos N. and J. Bruinsma, 2012). Given that wheat contributes roughly one-fifth of the world's daily caloric and protein intake, wheat will play a significant role in improving global food insecurity over the next 25 years (Shiferaw et al., 2013). Given the finite supply of arable land such increases in cereal production will have to come primarily from increased yields.

Several studies have estimated that globally, actual wheat yields range from 20 to 70% of their potential (T. Fischer, Byerlee, & Edmeades, 2014; Licker et al., 2010; Lobell, Cassman, & Field, 2009; Mueller et al., 2012). While actual wheat yields in many of the world's top wheat-producing regions have increased over the last several decades, potential wheat yields have increased at a faster rate, resulting in a relatively stable and consistent yield gap for wheat globally (Hatfield & Beres, 2019). Narrowing or reducing the yield gap for wheat can help address future food demand.

Currently, the yield gap for wheat in Canada is being studied as part of the Global Yield Gap Atlas (GYGA) project to determine yield gaps for major food crops across the world. Canadian Western Red Spring (CWRS), herein after referred to as wheat or spring wheat, is the most widely grown class of wheat in Canada (Cereals Canada, 2022). Beres (2022) estimates the yield gap for CWRS wheat in Saskatchewan is in the range of the 20 to 50%. More work is needed to understand the source of this yield gap.

1.1.1 The Yield Gap

Yield potential or potential yield[1] is defined as the yield of a crop, free from the limitations of nutrients, with all pests, weeds, and diseases effectively controlled (Evans & Fisher, 1999). The yield gap is calculated as the difference between potential yield and actual observed yields. Actual yields are expected to be reduced below potential yield by growth-

---

[1] Yield potential and potential yield are used synonymously throughout this report.

reducing factors including weeds, pests, disease, and lodging (M. K. Van Ittersum & Rabbinge, 1997).

Conceptually, reaching potential yield and eliminating the yield gap requires the perfect supply of nutrients and optimal management of any growth-reducing factors to remove all yield limitations and losses. However, it is not considered technically feasible to manage all potential yield losses nor is it considered economically rational to maximize production because inputs such as fertilizer are subject to diminishing marginal productivity (Lobell et al., 2009). The term *exploitable yield* is defined as the yield that farmers can be expected to reach given realistic management of yield losses and the associated costs and diminishing returns of applied inputs (Global Yield Gap Atlas, n.d.). Intuitively, and defined by several yield gap studies, exploitable yield represents the theoretical profit maximizing yield level (Cassman, Dobermann, Walters, & Yang, 2003; Fischer et al., 2014; Lobell et al., 2009; van Dijk et al., 2017). However, given that most yield gap studies are agronomic in nature and focus on estimating agronomic yield potential, exploitable yield is commonly estimated at 75 to 85% of the estimated agronomic yield potential (Fischer et al., 2014; Global Yield Gap Atlas, n.d.; Lobell et al., 2009), rather than as a separate economic measurement.

1.2 Motivation

While crop yields are primarily an agronomic measure, yield is an outcome of producer decisions, the environment, and the physical relationship between inputs such as seed and fertilizer, and production technology such as labor and equipment. The maximum of different output combination that correspond to a specific set of inputs and production technology, is commonly referred to as a production possibility frontier (Beddow, Hurley, Pardey, & Alston, 2015). The output that makes up the frontier is referred to as technically efficient production (Beddow et al., 2015). Conceptually, technically efficient production represents a crop's yield potential for any given quantity of inputs. Achieving the yield potential, which is maximum quantity of output per unit of land, would be a point on a production possibility frontier where all inputs except land are freely available and where optimal management of any growth-reducing factors removes all yield limitations and losses.

From an economic or management perspective *allocative efficiency* is a more important concept whenever inputs are costly. Allocative efficiency occurs when inputs and outputs are allocated to produce the highest possible profit (Beddow et al., 2015). Conceptually, allocative efficiency represents the profit maximizing or economically optimal yield level for a specific set of input and output prices. The simplest micro-economic theory of the firm is based on the premise that firms maximize profits and are therefore allocatively efficient.

A number of empirical studies have found that farmers often fail to produce at the economically optimal (allocatively efficient) level. The reasons given less than efficient input-usage include risk aversion, moral hazard, information or knowledge constraints, perceptions of the upcoming year, various resource constraints, or alternative objectives (Gómez-Limón, Riesgo, & Arriaza, 2004; Horowitz & Lichtenberg, 1993; Kelly, Adesina, & Gordon, 2003; Meijer, Catacutan, Ajayi, Sileshi, & Nieuwenhuis, 2015; Reader, Revoredo-Giha, Lawrence, Hodge, & Lang, 2018; Tong, Swallow, Zhang, & Zhang, 2019).

Previous studies looking at fertilizer profitability, particularly nitrogen, have found that producers often apply rates beyond what is economically optimal (Rajsic & Weersink, 2008; Reader et al., 2018). In the case of sub-Saharan Africa (SSA) where it is generally perceived that producers under apply fertilizers causing a significant yield gap, economic analysis shows that producers in many regions of Kenya are approaching economically optimal rates of nitrogen and in some cases exceeding what is economical (Sheahan, Black, & Jayne, 2012).

Further research is needed to better understand the economic components of the yield gap in Saskatchewan. Saskatchewan has been a leader in the development and adoption of the zero tillage which has radically transformed the cropping systems to include a diversity of crops and intensified crop production. Despite these changes and the existence of field level data, the allocative efficiency of nitrogen fertilizer use in wheat production has not been explored within the context of modern cropping systems. An empirical study is needed to understand how the on-farm use of nitrogen fertilizer contributes to the yield gaps observed in Saskatchewan spring-wheat production.

1.3 Objectives

The goal of this thesis is to quantify the economic optimal nitrogen rate for spring wheat production in Saskatchewan. This goal will be achieved by focusing on the following objectives:

1. To estimate the spring wheat yield response to fertilizer and chemical inputs, field management, weather, and soil characteristics.

2. To determine the economic optimal nitrogen rate for spring wheat production in Saskatchewan from the estimated spring-wheat yield response and assumed input and output prices.
3. To compare the estimated economic optimal nitrogen rate to actual observed nitrogen rates for spring wheat production in Saskatchewan to determine if an exploitable yield gap exists.

Identifying how much of the yield gap for spring wheat in west-central Saskatchewan is economically exploitable with respect to nitrogen fertilizer will help to inform researchers and policymakers of what potential constraints to production are contributing to the yield gap in Saskatchewan.

1.4 Overview of Methods

To achieve the objectives listed above, we use field-level input, management, soil, weather, and yield data from the Saskatchewan Crop Insurance Corporation (SCIC) to estimate a spring wheat yield response function for production in Saskatchewan. A linear regression model using fixed effects for year was employed for estimation using 7,537 observations from over 5,000 fields in Saskatchewan from 2011-2019. Focusing specifically on applied synthetic nitrogen fertilizer, the economically optimal nitrogen rate was determined where the value of the estimated marginal product of nitrogen was equal to the unit cost for nitrogen. While residual nitrogen already present in the soil was not accounted for in the model, the yield response to applied nitrogen was conditioned by the field's previous crop and soil productivity rating to capture potential differences. The model also captured the yield response of several other important input and management characteristics, including different spring-wheat varieties, producer management ability, chemical use, and rainfall, contributing to future research on factors constraining spring wheat yields below potential. Additionally, the average change in expected profitability between observed and optimal nitrogen rates was estimated for each previous crop and soil grouping to highlight the impact of optimizing nitrogen usage for spring-wheat producers in west-central Saskatchewan.

Currently in Saskatchewan, Agriculture and Agri-Food Canada (AAFC) is estimating the yield gap for spring wheat using the GYGA approach, explained in more detail in section 2.2.3. The GYGA approach begins with the determination of climatic zones (CZs) which are characterized predominantly by homogenous climate conditions. Once CZs have been established, reference weather stations (RWS) within the zone are designated to ensure the availability of historical and accurate weather data. A buffer zone of roughly 100-kilometers in radius is drawn around each RWS to capture a representative sample of crop data within each CZ. Together, weather data from the RWS, combined with soil, cropping patterns, common seeding date, and dominant variety are then used as inputs into crop simulation models. Potential yield is determined for each buffer zone. Actual yields, averaged across the designated buffer zone are then applied to form the yield gap as the difference between actual yield and potential yield. Early results suggest that spring wheat yields in Saskatchewan are 50 to 80% of estimated potential (Beres, 2022). Focusing specifically on AAFC Buffer Zone 21 in the west-central region of Saskatchewan, spring wheat yields are estimated to be 71% of potential yield, estimated to be 3,813 kilograms per hectare (Beres, 2022). AAFC Buffer Zone 21 was chosen for comparison because it matches up closely geographically with SCIC grain risk zone 16 which includes additional yield and management data. A map of the designated AAFC Buffer Zones is included in Appendix B, Figure B.1.

1.5 Organization of the Study

This thesis is organized as follows: Chapter 2 discusses background information on yield gap analysis and a review of existing yield gaps and determined causes worldwide. Chapter 3 describes the theoretical framework of the model, including a discussion of necessary data. Chapter 4 discusses the data selected for this analysis and describes the variables included in the model estimation. Chapter 5 describes the model and chosen functional form in more detail, including the results of several specification tests. Chapter 6 presents the empirical results, including the estimated optimal nitrogen rates and predicted corresponding yield gaps. Lastly, chapter 7 concludes the thesis.

Chapter 2: Background and literature review

2.1 Introduction

This chapter is divided into two main sections. The first section includes background information on the theory of yield gap analysis, and the defining and measuring of potential yield. This section concludes with a brief discussion of the Global Yield Gap Atlas (GYGA) approach to estimating and benchmarking yield gaps globally. The second section reviews existing yield gap studies to provide insight into the estimated yield gap for wheat globally, as well as the determination of potential constraints contributing to the yield gap. Lastly, because this thesis focuses on the optimization of applied nitrogen fertilizer as it relates to the yield gap, fertilizer-use profitability literature is reviewed with a particular interest on relevant methods and the interpretation of results.

2.2 Background of yield gap analysis

This section discusses important concepts and measurements pertaining to yield gap analysis. The yield gap was first defined by Gomez (1977), as the difference between actual farm yields and nearby experimental station yields. Since then, yield gaps have been analyzed for multiple crops throughout the world (Beza, Silva, Kooistra, & Reidsma, 2017). Today, there is a renewed interest in yield gaps by researchers as a potential tool to help improve global food security. Yield gaps are a particularly useful metric for assessing the potential capacity of agricultural production systems and provide methods for identifying existing constraints to production at a local, national, and global level. This section begins with a background discussion of yield potential, exploitable yield, and actual yields which makes up the basic construction of the yield gap. Given an apparent lack of standardization in current yield gap analysis methods, the Global Yield Gap Atlas approach is highlighted to provide insight into the direction of future yield gap analysis that addresses existing shortcomings.

2.2.1 Concepts related to the Yield Gap

The yield gap is made up of several concepts with definitions that are inconsistent across studies. Given that this thesis focuses primarily on exploitable yield, this section works to defines the key concepts of the yield gap that are consistent with the approach of the Global Yield Gap Atlas.

The *yield gap* is defined as the difference between potential yield and the actual observed yield, commonly averaged across a region. *Potential yield* is defined as the yield of crop adapted to the environment it is grown in, free from limitations in nutrients and water, and free of yield losses caused by pests, diseases, or lodging (Evans & Fisher, 1999).

Given its definition, potential yield is determined solely by *growth-defining* factors which include specific genetic characteristics, solar radiation and atmospheric CO_2 levels, and temperature (M. K. Van Ittersum & Rabbinge, 1997). It is a location and time specific measure but is not dependent on soil characteristics, assuming that the addition of nutrients and water at some level can alleviate most soil constraints. Potential yield is also sensitive to planting date given that specific genetic plant characteristics will dictate the crops maturity date (Lobell et al., 2009).

While potential yield is an appropriate measure for irrigated crops, rainfed crops are constrained by *growth-limiting* factors including water and nutrient-supply levels (M. K. Van Ittersum & Rabbinge, 1997). *Water-limited potential yield* is used as the benchmark for dryland production systems, following the same definition as potential yield with the addition of limitations in water supply and the added dependency of soil characteristics which may impact the water holding capacity of the soil (Martin K. Van Ittersum et al., 2013). *Actual yields* are reduced below potential yield or water-limited potential yield by *growth-reducing* factors including weeds, pests, disease, and lodging (M. K. Van Ittersum & Rabbinge, 1997). Actual yields are considered time and location specific, calculated as the average yield under the typical management characteristics for seeding date, seeding rate, fertilizer management, and crop protection practices (Martin K. Van Ittersum et al., 2013). Therefore, the yield gap is calculated as the difference between potential yield (irrigated crops) or water-limited potential yield (rainfed), and average actual farm yields, depicted in Figure 2.1 below.

Figure 2.1: Factors that reduce yields below that of potential yield down to actual yields (a). The yield gap between actual yields and potential yield, and the exploitable yield gap between actual yields and exploitable yield (80% of potential) (b).

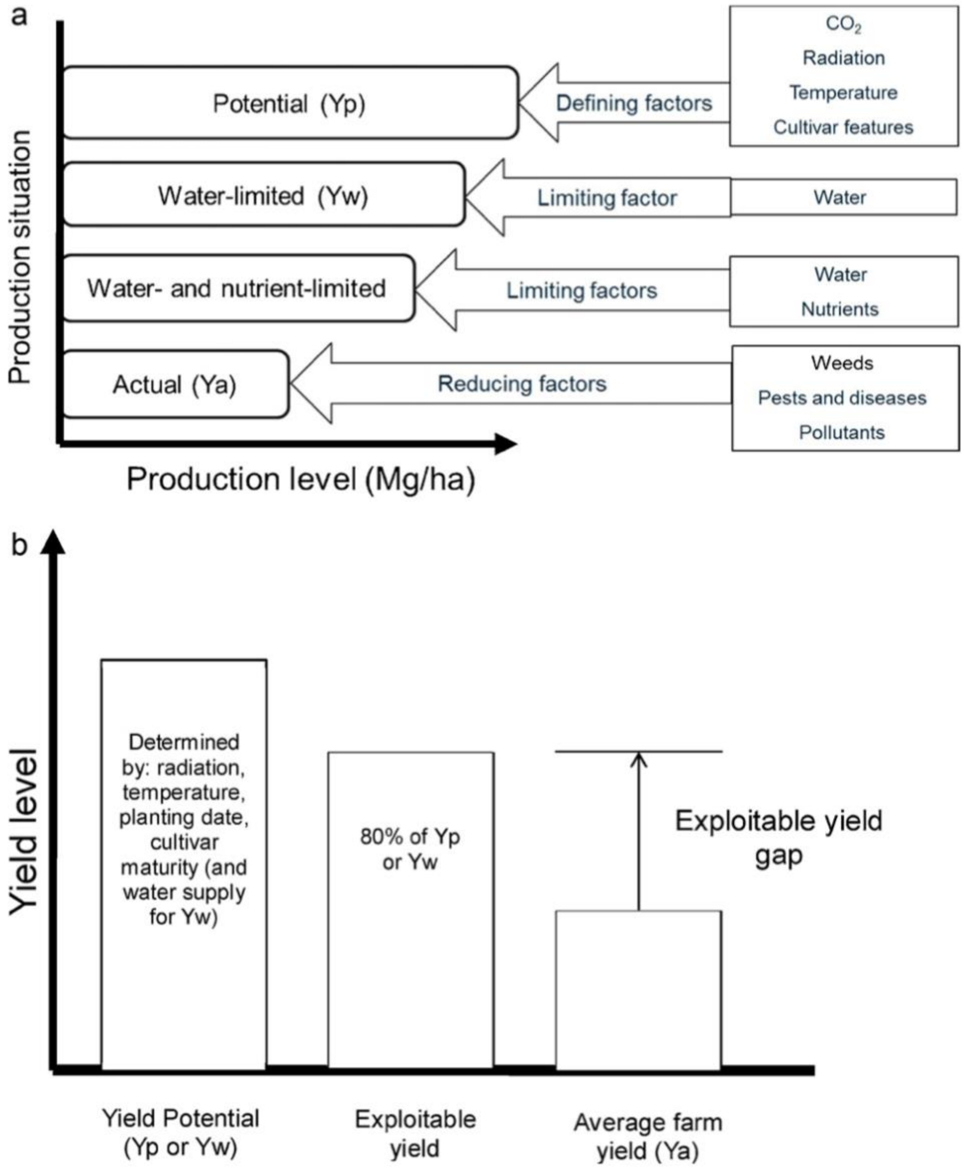


Image source: Lobell et al., 2009; van Ittersum et al., 2013.

Potential yield (including water-limited potential) is defined by the environment that the crop is grown in, while actual yields are both limited and reduced below potential. Therefore, reaching potential yield requires the perfect supply of nutrients and management of any growth-reducing factors to remove all yield limitations and losses. Additionally, it is not considered economical to invest continuously in inputs as the yield response of all inputs eventually diminishes (Lobell et al., 2009). Given that actual yields have been observed to plateau between 75 to 85% of potential, *exploitable yield* is commonly estimated at 80% of the determined potential yield, however, rarely with any economic considerations (Cassman et al., 2003; Global Yield Gap Atlas, n.d.).

Therefore, exploitable yield represents the point where it is no longer economical for yields to be increased, or put differently, for the yield gap to be closed. The gap between potential yield and what is considered to be exploitable is expected to be made larger by risk and uncertainty, especially under rainfed cropping systems (Fischer et al., 2014). It is worth noting that while exploitable yield, by definition, represents the realistic management ability of producers and the economically optimal yield level, it does not account for any regional or farm-level economic factors related to production in its calculation. Given that actual yields are the result of farmer decisions, understanding what drives these decisions can help to explain why yield gaps may exist from a socioeconomic perspective, and how much is truly exploitable to be closed (Beddow et al., 2015).

2.2.2 Measures of Potential Yield

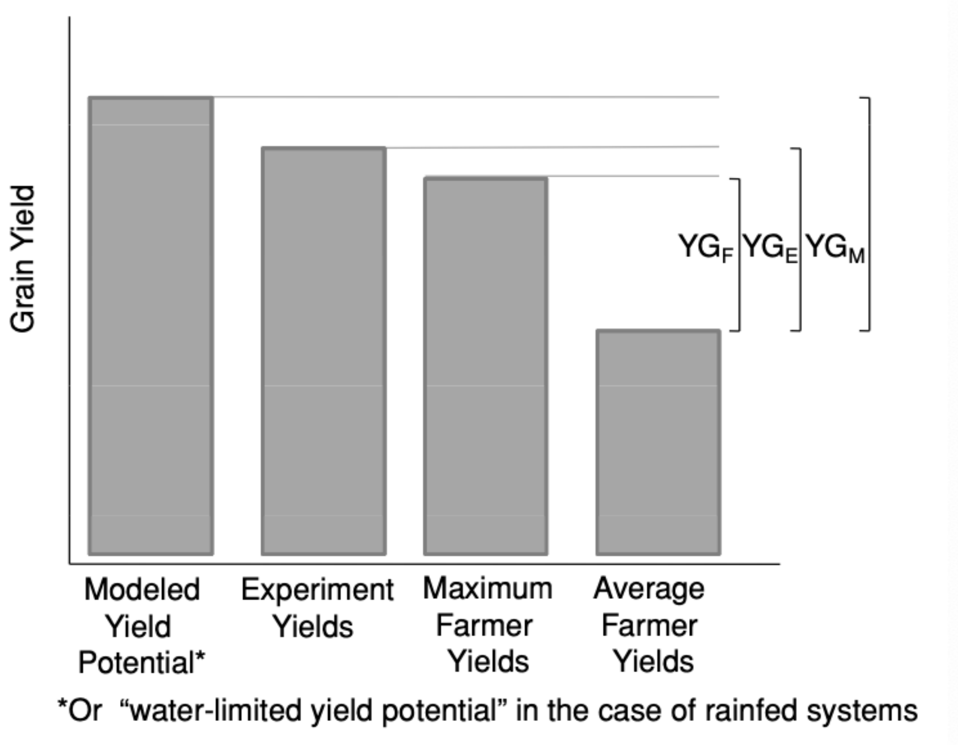
As discussed in section 2.2.1, the upper boundary of the yield gap is determined by a crop's potential. Potential yield is by definition, the yield of a specific crop free from the limitations of water, nutrients, pests, and disease, dictated only by solar radiation, and temperature. It requires perfect management in both the supply of nutrients and management against yield reductions from weeds, pests, disease, and lodging, throughout the entire growing season. This level of management is considered largely impossible at the field level and is difficult to consistently replicate at the smaller experimental plot level, yet the measure of potential yield as it relates to the yield gap remains an important metric for determining the potential capacity of agriculture production. Therefore, potential yield estimated using crop simulation models allows for various growth conditions and management abilities to be simulated rather than necessarily observed (Lobell et al., 2009). Additional approaches to measuring potential include the use of experimental plot yields, highest yield contests, and maximum observed farm yields from local surveys or reports, all of which are discussed in further detail below.

The least common method of measuring potential yield is the approach of surveying local farm yield data to determine the maximum yields achieved among actual farmers within a particular region. One of the key issues of this approach is that not all fields are intensively managed in regard to fertilizer use, and pest and disease control, to ensure that yield losses are minimized, and therefore true yield potential is rarely observed (Lobell et al., 2009). While this method provides the smallest measure of the yield gap between potential and actual yields, it most likely underestimates potential yield given that even the highest actual farm yields are still expected to be constrained by economic factors if not also by other biotic or abiotic factors not adequately managed at the field level.

Yield gap studies have also leveraged local field or plot-level experimental yields. While this method is similar to that of maximum farmer yields in the sense that actual observed yields are used to estimate potential yield, field or plot-level experimental production typically takes place on a small acreage scale where both yield-limiting and yield-reducing factors can be effectively controlled (Martin K. Van Ittersum et al., 2013). Additionally, experimental research stations are commonly exempt from the same economic constraints faced by farmers. For field experiments to produce robust measures of potential yield, experiments need to be conducted over multiple seasons to control for climate variations from year to year that may influence what is considered optimal management in each season (Lobell et al., 2009; Martin K. Van Ittersum et al., 2013). This presents a significant challenge for researchers to limit yield losses consistently year after year. Lastly, yield contests focus on maximizing yields rather than profits, which is assumed to be the goal of producers under normal circumstances.

Crop simulation models address the shortcomings of using observed maximum or experimental yields to estimate potential. Simulation models represent the estimated response of a crop to various factors and conditions modeled throughout the growing season. They typically require localized climate data including daily minimum and maximum air temperature, solar radiation, precipitation, humidity, and field level management data regarding the depth and date of seeding, as well as the maturity rating of the variety that was planted. Additionally, for water-limited potential yield to be simulated, models typically require data including soil texture, typical root depth of the crop, and how much moisture is available at the beginning of the season for best estimates (Lobell et al., 2009; Martin K. Van Ittersum et al., 2013). Given these requirements, data limitations are one of the obvious challenges of using crop simulation models to estimate potential yield. The more data provided, the more accurate the results are considered. Estimates are most robust when using daily values for climate factors as well as detailed soil information. Additionally, because simulation models attempt to remove yield-limiting and yield-reducing factors to estimate potential yield, estimates may be significantly larger than what is observed at the farm level, especially in areas of low-input farming systems. However, potential yield as determined by a crop simulation model is considered most accurate to the true agronomic potential of the crop (Lobell et al., 2009).

Figure 2.2: Framework comparing the three common measures of potential yield relative to actual yields. YGM represents the model-based yield gap, YGE represents the experimental yield gap, and YGF represents the maximum farmer yield gap.



*Or "water-limited yield potential" in the case of rainfed systems

Image source: Lobell et al., 2009.

As illustrated in Figure 2.2, the yield gap is expected to vary in size depending on which method is used to estimate potential yield. While each method has its limitations, together they validate the estimated potential yield at each level. Simulation models are expected to produce the highest estimates of potential yield but can be validated by experimental and yield-contest results that are impacted by both seasonal and management conditions. Most studies fail to estimate more than one potential yield measure and are therefore unable to be compared or validated in this way (Lobell et al., 2009).

While the availability of data limits the application of the methods mentioned above to a local scale, the use of aggregated datasets for cropping area, yield, and climate characteristics has allowed for yield gap analysis to take place at a more global scale. Provincial/state or nationally aggregated data is more readily available. The dataset compiled by Monfreda, Ramankutty, & Foley (2008) from FAOSTAT reported yields, is commonly applied in global yield gap studies (Martin K. Van Ittersum et al., 2013). Yield and crop areas are measured at the national and subnational level for a variety of crops. The main limitations of global yield gap studies are that while extensive yield data exists, they rarely distinguish between irrigated and rainfed production systems, and other management practices which can make the resulting yield gap appear much larger than it is (Snyder, Miththapala, Sommer, & Braslow, 2017). It is worth noting that while several regions may share the same climatic zone, they can have significantly different management practices and socioeconomic conditions that constrain yields below potential (Martin K. Van Ittersum et al., 2013). For example, borders define different socioeconomic conditions, but climate characteristics are not confined by political borders. While these limitations are important to consider, global studies can be of help to compare regions and determine areas where large yield gaps exist, requiring more research, commonly referred to as a top-down approach.

2.2.3 The Global Yield Gap Atlas approach

To address the scale limitations of crop simulation models and the inaccuracy concerns of global yield gap approaches that use coarse, aggregated datasets, the Global Yield Gap Atlas (GYGA) approach was established. The GYGA approach uses local weather, soil, and cropping data to estimate potential yield using crop simulation models, validated by local experts, and through observed yield levels. This approach allows for accurate estimates to then be scaled upwards to comparable climatic zone levels, commonly referred to as a bottom-up approach. This section describes the steps employed in the GYGA approach in further detail.

The GYGA approach begins with the determination of climatic zones (CZs) which are characterized predominantly by homogenous climate conditions including growing degree days

(GDD), temperature seasonality, and the ratio of in-season precipitation to in-season potential evapotranspiration. Once CZs have been established, reference weather stations (RWS) within the zone are designated to ensure the availability of historical and accurate weather data. A buffer zone of roughly 100-kilometers in radius is drawn around each RWS to capture a representative sample of crop data within each CZ. Distinctions are made between zones that are irrigated and rainfed. Dominant soil types and cropping patterns are then identified within each buffer zone based on localized data. Together, weather data from the RWS, combined with soil, cropping patterns, common seeding date, and dominant variety are then used as inputs into crop simulation models. Potential yield is determined for each soil type by cropping pattern combination so that a buffer zone with two dominant soil types and two common cropping patterns will result in four simulated potential yields. Actual yields, averaged across the designated buffer zone are then applied to form the yield gap as the difference between actual yield and potential yield. If actual yields are not available at the same scale as the buffer zone, more aggregated yields are applied as a second-best option.

Currently in Saskatchewan, Agriculture and Agri-Food Canada (AAFC) is estimating the yield gap for spring-wheat using the GYGA approach. Early results suggest that spring-wheat yields in Saskatchewan are 50 to 80% of estimated potential (Beres, 2022). In west-central Saskatchewan in AAFC Buffer Zone 21, spring wheat yields are 71% of potential yield, estimated to be 3,813 kilograms per hectare (Beres, 2022). As discussed in section 1.4 above, AAFC Buffer Zone 21 was chosen for comparison because it matches up closely geographically with SCIC grain risk zone 16 which includes additional yield and management data. A map of the designated AAFC Buffer Zones is included in Appendix B, Figure B.1.

2.3 Literature Review

This section reviews relevant literature on how big the yield gap for wheat it is globally, why it may exist, and what opportunities exist for closing it. It begins with a review of several wheat yield gap studies from key growing areas around the world. While the objective of this thesis focuses on actual yields as they relate to what is determined to be exploitable, the review of yield gaps globally and what may cause them serves to illustrate the many possible exploitable constraints producers face. Little work has been done to estimate the yield gap for wheat within Saskatchewan; however, some global studies have estimated yield trends for Canada using aggregated data. Additionally, potential explanations as to why these gaps may exist provided from the literature are discussed. Lastly, the prospects of closing such gaps are discussed. This section reviews the empirical application of the methods and theory described in section 2.2 above.

2.3.1 How big is the yield gap for wheat?

The yield gap for wheat varies greatly throughout the literature, in size, calculation, and how it is defined. This section discusses measures of the yield gap for wheat from global studies mostly. While this review is not exhaustive of all yield gap studies for wheat, it helps to show the overall trend of actual yields compared to potential yield overtime, and how large the yield gap for wheat is. It is important to again note that this thesis focuses specifically on Canadian Western Red Spring wheat, whereas global studies may not necessarily make the same distinction.

Hatfield and Beres (2019) employed a global yield gap study using FAOSTAT and U.S. National Agriculture Statistics Service (NASS) data for the top ten wheat producing countries, from 1960 to 2017. Using a quantile regression at the 95th percentile, annual yields near the upper frontier for each country represent what they define as attainable yield in which annual weather was not a limitation to yield. The yield gap is calculated as the difference between attainable and actual yield, divided by attainable yield. Average yield gaps over the total period, range from 0 in Germany to 0.24 in Australia and Canada. The slope of attainable yield over time is assumed to represent the technology increase for a country, while the slope of actual yield over time represents the increase in yields from the combination of both weather effects and the adoption of now available technology (Hatfield & Beres, 2019). While actual yields in all countries have shown continual increases, attainable yields have increased at a larger rate from technological advances, resulting in the yield gap remaining consistent over time. Beddow et al. (2015) point out that if the yield gap shifts upwards from improvements in both actual and potential (attainable as it is referred to here) yield, rather than grows larger, food security will not necessarily be negatively affected.

To account for more localized trends, Hatfield and Beres (2019) employ state and county-level production data from the NASS for the top three wheat producing states in the US (Kansas, North Dakota, and Washington). They also employ wheat production data from Saskatchewan, all showing the same result as at the national level with a consistent yield gap. Yields at the state and province level highlight that in-season weather effects reduce the full impact of technologies that increase attainable yield, affecting the size of the yield gap (Hatfield and Beres, 2019).

Alternatively, Licker et al. (2010) compared aggregated crop yields by climate zones globally to effectively remove the influence of climate on yield. They employ the global crop dataset from Monfreda et al. (2008) for the year 2000, dividing the world into 100 individual climate zones based on growing degree days (GDD) and soil moisture index, defined as the ratio of evapotranspiration to potential evapotranspiration. Climatic potential yield is defined as the 90th percentile yield within a climate zone, with the yield gap calculated as the difference between climatic potential and actual yields. Therefore, yields from two opposite sides of the world, belonging to one climate zone are compared to estimate the magnitude of factors not related to climate that impact crop yields. The authors suggest that factors responsible for the yield gap, unrelated to climate, appear to follow political borders rather than bio-physical ones implying that they may be out of the farmer's control. Licker et al. (2010) concludes that approximately 60% more wheat could be produced globally if climatic potential were reached. Within Canada, the gap between actual yields and climatic potential ranges from approximately 25% in Western Canada up to approximately 75% in Eastern Canada.

In their analysis of yield gaps for the three major cereal crops, Fischer, Byerlee, & Edmeades (2009) review a series of wheat yield case studies for important wheat mega-environments (WME's). These regions include Mexico, India, Australia, and North Dakota, that make up a large part of global wheat production. However, Canada is not included, as a significant producer and exporter of wheat. They define the yield gap as the difference between potential yield and farm yields, represented as a percentage of farm yield. Throughout the WME's, the yield gap ranges from 35 to 50%. The lowest yield gap was found in India where crops are commonly irrigated at 35%, while the highest gap was found in North Dakota under rainfed conditions at 50%. Fischer et al. (2009) estimate potential yield using variety trial data from breeders and national databases. They find that progress in both actual yields and potential yield have resulted in a consistent yield gap over time, with only significant improvements in the rate of change for farm yields in Australia. These results are similar to those reported by Hatfield and Beres (2019).

Alternatively, in their analysis of the yield gap for several crops, Fischer et al. (2014) report that the yield gap for wheat in several growing regions is slowly closing. Interestingly, they also report that the yield gap for winter wheat in the UK and France is getting larger as actual farm yields have slowed dramatically while potential yield continues to grow. This is attributed to a small gap existing between actual and potential yield, discussed as a lack of an exploitable yield gap in the previous section. The yield gap is again represented as a percentage of farm yield, ranging from 30% in parts of Europe to as high as 69% in rainfed Saskatchewan. Potential yield is estimated using variety trial data from breeders when available and comparing to both official and survey reports for actual yields. Fischer et al. (2014) estimate that actual yields are increasing at least twice as fast as potential yield in both Mexico and Australia, yet a significant yield gap is still estimated in these two regions (Lobell et al., 2009; Monjardino, Hochman, & Horan, 2019). Similar to Fischer et al. (2009), and Hatfield and Beres (2019), the yield gap has remained consistent over time for most regions due to similar growth in actual and potential yield in many regions.

In a more localized approach, Monjardino et al. (2019) employ a data rich method similar to the GYGA approach, to estimate water-limited potential yield for wheat in Australia. The authors exploit more detailed soil, climate, and crop data from statistical local areas (SLA's), the smallest administrative units in Australia. Buffer zone radiuses from reference weather stations (RWS) are shrunk to approximately 20 km, opposed to 100 km under the GYGA protocol. The authors use the APSIM crop simulation model to estimate water-limited potential yield. 245 SLA's are estimated to compare actual yields to potential yield. They find that actual yields have increased linearly with potential yield over time, and that farmers are achieving about 50% of potential. The authors compare this gap to both estimates from the GYGA protocol, and from EarthStat global yield and potential yield maps, confirming that when estimates are upscaled to the national level they are similar (Monjardino et al., 2019). The estimated trendline of actual and potential yield is similar to those found by Hatfield and Beres (2019).

Fischer et al. (2014) found a smaller yield gap for Australia in their case study, noting that actual yield growth has accelerated at a rate faster than that of potential yield. They estimated water-limited potential yield to be 2.6 t/ha, while Monjardino et al. (2019) estimated water-limited potential yield to be closer to 3.5 t/ha. This estimate was similar to what was found by the GYGA when mapping Australia. The difference in estimates may be explained by Fischer et al. (2014) using national variety trial yield data to measure potential yield, while Monjardino et al. (2019) and the GYGA protocol use the application of crop simulation models. This concept is discussed in the previous section where simulation models tend to estimate larger potential yield measures compared to actual observed field-level trials. As previously mentioned, simulation models have the ability to mute all yield reducing factors to truly estimate maximum potential.

The yield gap for wheat varies globally, however it is difficult to compare as both the definition and calculation of potential yield varies between studies, even for global analyses. What is similar among the findings is that the trend of actual yields and potential yield has increased over time and the yield gap has remained statistically consistent. This indicates that gains in potential yield are being realized at the farm level, just not to the full extent. Climate variability may reduce the full extent of yield gains from new technology (Hatfield and Beres, 2019).

2.3.2 Why does the yield gap exist?

The causes of yield gaps vary considerably across the globe and are often multiple and interconnected within a region (Fischer et al., 2014). Given this, it is difficult to easily determine what the underlying causes of the gap are and how to go about alleviating these constraints. Studies at different scales cause factors to differ in importance and there is no standard approach for evaluating these causes (Beza et al., 2017). Therefore, this section of the review discusses various methods employed in the literature to evaluate the causes of yields gaps in different regions and highlights the current identified causes. These methods include surveys and on-farm experiments, regression analysis of yield heterogeneity and management, satellite imaging, crop simulation models, surveys of local experts, and global studies that leverage remote sensing or aggregated datasets.

In their global analysis of the climatic yield gap, Licker et al. (2010) effectively removed the influence of climate on yield to highlight the portion of the yield gap resulting from factors other than climate such as agriculture management. They use aggregated production data, land-use data on irrigated acres, and climate data. Interestingly, they find that the differences in high and low yielding zones follow political borders rather than biophysical ones. In the case of Western Europe and Eastern Europe, the authors find that both regions share a similar climate yet yields in Eastern Europe are significantly lower than in Western Europe. They suggest that this is the result of the collapse of the Soviet Union and a dramatic increase in the cost of fertilizer prices following the removal of subsidies. Mueller et al. (2012) employ the same dataset from Monfreda et al. (2008), building off the work of Licker et al. (2010). They find that yield variability can be explained by differences in nutrient management and water availability. The yield gap could be narrowed by increasing the amount of irrigated hectares by 25% and increasing nitrogen use by about 30% in underachieving areas (Mueller et al., 2012). While these studies are global, they highlight regions where a greater focus is needed showing the benefit of such global studies.

Neumann, Verburg, Stehfest, & Müller (2010) employ the same dataset from Monfreda et al. (2008), using a stochastic frontier analysis (SFA) to show the efficient yield level at different average weather variables. Yield observations off the upper boundary (frontier) are deemed inefficient. Inefficient points are then related to several management and soil variables to determine the cause. Fertilizer use was not included due to a lack of available data. They find the average efficiency for wheat to be 0.64. van Dijk et al. (2017) employ the same approach, using the Living Standards in Agriculture (LSIA) survey data from Tanzania to estimate a frontier and inefficiency gap. They disaggregate the gap into four smaller gaps. They measure the gap between the actual observed yield and the frontier (technically efficient), the gap between the technically efficient yield and the allocatively efficient yield, the gap between the allocatively efficient yield and the maximum technically efficient yield, and the gap between the maximum technically efficient yield and potential yield above the frontier, only reachable through new technologies and the latest crop varieties. van Dijk et al. (2017) find that the biggest contributors to the inefficiency gap is a lack of accessible technology and difficulty reaching the economic optimum (allocative efficiency). Given a lack of data, a partial approach is used to determine the economic optimum or allocatively efficient point focusing only on nitrogen fertilizer.

Chapagain & Good (2015) employ an analysis of variance (ANOVA) method for grain yield by variety, year, location, and their interaction. Climate and soil data are taken from Environment Canada, while yield data is derived from Alberta Regional Variety Trial data. They focus on estimating attainable yield defined as the yield under proper nutrient management and crop protection, and maximum attainable yield under optimal variety use and proper nutrient and crop management. The authors conclude that yields can be improved by proper nutrient management including the use of soil testing, and by using varieties that are specifically located for the growing region. Nitrogen rates are generally chosen based on the growers past experiences, yield targets, and what the neighbour is doing (Chapagain & Good, 2015).

In their national analysis of which factors constrain wheat yields in Australia, Monjardino et al. (2019), simulate the impact on yield that suboptimal management practices have compared to optimal practices that are used to determine the water-limited potential yield. They look at the effects of suboptimal nitrogen rates, delayed seeding date, reduced weed control, and low seeding rate. Using the simulated water-limited potential yield as the benchmark of optimal practices, the authors find that the average nitrogen rate in Australia of 45 kg/ha reduced yield potential by 40%, delayed seeding date by two weeks reduced potential by 7%, reduced weed control reduced potential by 26%, and low seeding rates reduced potential by 12%. These results show that nitrogen use impacts almost half of the yield gap in Australian wheat production (Monjardino et al., 2019), as estimated by a crop simulation model.

Employing the use of a survey of experts, (Waddington, Li, Dixon, Hyman, & de Vicente, 2010) review the response of over 650 surveyed experts to help determine the causes of yield losses for major food crops and growing regions in Africa and parts of Asia. Constraints include those that are abiotic, biotic, management, and socioeconomic related. They find from the survey results that wheat yield constraints included nitrogen deficiencies and poor management, as well as in season drought conditions and poor irrigation management, similar to that found by Mueller et al. (2012). Ultimately, constraints differ by region with management constraints more impactful in parts of Asia and socioeconomic constraints more impactful in Africa.

In their review of 50 peer reviewed agronomic articles, Beza et al. (2017) assess data availability, collection methods, and the most considered and explaining factors of the yield gap. They review studies at various scales, highlighting that scale does impact the importance of factors. The authors find that management and soil factors are included more than farm and socioeconomic factors, yet farm and socioeconomic factors did help to explain the causes of the yield gap when included. Data availability is the major limitation of many yield gap analyses, as suggested in other studies, with no standardized approach to collection or upscaling. Beza et al. (2017) suggest that a bottom-up approach leveraging local data and experts can help to alleviate data limitations and highlight the issues driving the yield gap at a local level, given the importance of farm characteristics and socioeconomic conditions in determining the yield gap.

Fertilization was the most included management factor, included in nearly half of the reviewed studies. While it explained a portion of the yield gap in many cases, the timing, costs, and efficiency had as much or more explaining power (Beza et al., 2017)). The explanatory power is related to location. Fertilizer deficiencies helped to explain the gap in Africa but less so in Asia. Given that the majority of these studies are agronomic in nature, farm and socioeconomic factors are less often included (Beza et al., 2017).

The underlying causes of yield gaps vary across regions even when the role of climate is controlled for. These causes are often multiple and interconnected meaning there is no "silver-bullet" solution to closing the yield gap. The yield gap is estimated to be larger for regions in developing countries and those under rainfed production systems. The most likely causes of the yield gap in wheat include poor nutrient management, lack of irrigation, and a lack of applied technologies such as the newest varieties (Beza et al., 2017; Mueller et al., 2012; van Dijk et al., 2017; Waddington et al., 2010). It is unclear as to what drives these constraints. While developing countries lack access to technology and crop inputs, developed countries also face risk and uncertainty when applying inputs or adopting new technologies throughout the growing season, especially under dryland conditions.

2.3.4 Fertilizer-Use Profitability Literature

The previous section discussed several of the underlying causes driving yield gaps. While it is difficult to compare causes across regions, one common factor among studies that looked at multiple regions was that of poor nitrogen management and deficiencies. This is not necessarily surprising due to the importance of nitrogen in plant growth. Given that several yield gap analyses reviewed in the previous sections are agronomic in nature, they do not include further discussion on the underlying causes of the determined constraints driving yield gaps, such as nitrogen management at the farm level. This section reviews literature on the profitability of fertilizer-use to better understand what factors influence fertilizer input-use decisions.

Rajsic and Weersink (2008) look at what factors cause Ontario maize farmers to apply nitrogen at levels beyond what is recommended by government extension services. They estimate and compare the *ex-post* economically optimal rate of nitrogen to the *ex-ante* recommended nitrogen rate for multiple sites, years, and yield-response functions to try and explain over-application. Producers may feel that the provincially recommended nitrogen rate, which is based on a quadratic-plateau yield-response function is not representative of their farm and therefore, should apply more (Rajsic & Weersink, 2008). Yield responses are estimated for a linear-plateau, quadratic, quadratic-plateau, and Mitscherlich functional forms. The authors do not find any statistical or economic basis for choosing one functional form over another, meaning the perception of the yield response function does not explain over-application. They do, however find a flat pay-off function on several of the sites with the difference between the economic optimum and the recommended rate being less than $10 per hectare. This suggests that the cost of over-application is low and is perceived as less costly than under-applying, especially in a favorable year. These results also suggest that given a relatively flat pay-off function, employing a service such as soil testing for more accurate fertilizer information may not be worth it.

The findings of Rajsic and Weersink (2008) are consistent with the findings of Reader et al. (2018), who find that producers in the UK apply inputs beyond what is economically optimal. Farmers appear to attempt to maximize yields rather than profits. This behaviour is explained by potential optimism by the producer about the upcoming year, avoidance of the risk of under-application as it is viewed as more costly than over-applying, and the view that recommended rates are not representative of the specific farm or field. These studies look at explanations for the over-application of fertilizer and attempt to provide rational for what drives input-use decisions.

Contrary to over-application, it is generally perceived that Sub-Saharan African (SSA) farmers under apply fertilizers. However, it has not been widely studied from an economic perspective. Sheahan et al. (2012) analyze production data in Kenya to determine the profitability of nitrogen rates on maize production. The authors determine the optimal condition for nitrogen for risk neutral producers where the marginal value-to-cost ratio of nitrogen and maize is equal to 1, and equal to 2 for risk-averse producers. This can be explained as risk-neutral producers demanding an equal payoff for applied nitrogen, and risk-averse producers requiring a 2 to 1 payoff for applied nitrogen. They find that most producers are approaching economically optimal nitrogen rates, and in the highest yielding regions, producers are over-applying nitrogen from an economic perspective. These results suggest that even though nitrogen deficiencies are generally regarded as a significant yield gap constraint for SSA, if rates are approaching economic optimums, then there are additional constraints limiting how much of the yield gap is economically exploitable with respect to nitrogen use in SSA. This analysis provides an important partial measure of the exploitable yield gap for SSA, suggesting that it is significantly below the 80% benchmark.

2.4 Conclusion

While potential yield is predominantly a theoretical measure, unreachable by most producers, it is crucial for determining the capacity of agriculture production and the growth rate of yields that can be expected at the farm level. The yield gap for wheat globally stands at roughly 50% and closing slowly at a rate of roughly 0.4% per year given increases in actual yields relative to increases in potential yield over time (Fischer et al., 2014). At this rate of closure, exploitable yield will be reached in roughly 60 years. The yield gap is attributed mostly to poor nutrient management, lack of irrigation, and a lack of applied technologies such as newer varieties (Beza et al., 2017; Mueller et al., 2012; Sheahan et al., 2012; van Dijk et al., 2017; Waddington et al., 2010). While low rates of inputs may reduce yields within a region, the underlying cause may be that it is most profitable at the farm level, and exploitable yield is reached. Additionally, in western Canada, only 30% of producers reported soil testing regularly, raising the question that producers may not be fully aware of what their soil requires for nutrients. It is important to determine if the closure of such gaps is economically or socially feasible. This review suggests that there is room for improvement in wheat yields theoretically, yet a deeper understanding of the economic constraints at the farm level that drive such decisions as input usage is needed to capture what portion of the yield gap can feasibly be closed.

Chapter 3: Theoretical Framework

3.1 Introduction

This chapter discusses the theoretical framework that is applied to estimate the economic optimal nitrogen rate for spring wheat production in west-central Saskatchewan. The economic optimal nitrogen rate is estimated from field-level yield, management, and soil data, in addition to climatic conditions across multiple years. The estimated economic optimal nitrogen rate is then compared to actual observed nitrogen rates for spring wheat production to determine if an exploitable yield gap exists. The estimated spring wheat yield response function is also used to estimate the nitrogen rate required to reach maximum yield with respect to nitrogen, holding all other inputs constant. This chapter begins with a discussion of several assumptions that are made, followed by a breakdown of the spring wheat production function and conditions for optimizing inputs. A detailed discussion of estimation takes place in chapter 5.

3.2 Underlying Theory

3.2.1 Assumptions

In an effort to estimate the economic optimal nitrogen rate, several simplifying assumptions are made regarding spring wheat production in Saskatchewan. It is assumed that the primary objective of Saskatchewan spring wheat producers is to maximize profits. These producers are assumed to be risk neutral in the use of variable inputs including fertilizer and crop protection products. It is assumed that a cropping year is the short run. Estimation of the economic optimal nitrogen rate is focused on a given field in a given cropping year and therefore it is assumed that, in the short run, fixed inputs are held constant. This includes such decisions as the crop chosen (spring wheat) and the fields crop rotation. Additional assumptions specific to data sources and usage take place in section 4.2 Data Sources of the following chapter.

3.2.2 Spring Wheat Production Function

Spring wheat production is dependent on several factors including variable inputs, management characteristics, and physical conditions. In equation 3.1 below, the yield y of a spring wheat field i at time t is a function of multiple vectors characterized as,

$$y_{it} = f(v_{it}, m_{it}, s_{it}) \qquad 3.1$$

where v_{it} is a vector comprised of variable inputs including applied fertilizers and crop protection products. Vector m_{it} is comprised of management factors including chosen crop rotation, seeding date, and the variety of spring wheat chosen to grow. Lastly, vector $s_{it}$is comprised of the physical conditions of spring wheat field i at time t including beginning topsoil moisture conditions, accumulated rainfall, and soil productivity.

3.2.3 Expected Profit Function

Given that the producer does not yet know the final output price, the expected profit function equation 3.2 below, is based on the producer's expectation of spring wheat prices characterized as,

$$\mathrm{E}[\pi_{it}(p_t, w_t, r_t, v_{it}, m_{it}, s_{it})] = \mathrm{E}[p_t] \cdot f(v_{it}, m_{it}, s_{it}) - \sum(w_t \cdot v_{it}) - \sum(r_t \cdot F_{it}) \qquad 3.2$$

where p is a scalar of the spring wheat output price, based on the producer's expectation of spring wheat prices, which is assumed to be exogenously determined. Vector F includes fixed inputs while vector w includes prices for variable inputs, and vector r includes prices for fixed inputs.

3.2.4 Profit Maximization

As discussed in section 3.2.1 Assumptions, it is assumed that the goal of the producer is to maximize expected profit for each field each year. In the short-run, variable input levels will be chosen to maximize expected profit described as,

$$\mathrm{E}\big[\pi_{it}\big(p_t, w_t, v_{itj}\big)\big] = \underset{v_{itj}}{\max}\ \mathrm{E}[p_t] \cdot f\big(v_{itj} \big| v_{it}, m_{it}, s_{it}\big) - \sum(w_t \cdot v_{it}) - \sum(r_t \cdot F_{it}) \qquad 3.3$$

where subscript j denotes a specific variable input and spring wheat production is dependent on the chosen level of variable inputs for field i in year t. In the short run, a producer will maximize expected profits by choosing the optimal level of only variable inputs (v_{it}) for field i in year t. Management factors (m_{it}), physical conditions of the field (s_{it}), and fixed inputs (F_{it}) are fixed for each field i in year t.

3.2.5 First Order Condition

The first order condition of the short run profit maximization equation is,

$$\frac{\partial \mathrm{E}[\pi_{it}(p_t, w_t, v_{it})]}{\partial v_{itj}} = \max_{v_{itj}} \mathrm{E}[p_t] \cdot f'\left(v_{itj} \middle| v_{it}, m_{it}, s_{it}\right) - w_{tj} = 0, \text{for all } j \in 1, \dots, n \qquad 3.4$$

derived by taking the partial derivative of the expected profit function with respect to each variable input (v_{itj}).

3.2.6 Optimal Condition for Variable Inputs

Using equation 3.4 above, the optimal condition for a variable input can be determined. Given the focus of this thesis pertains to nitrogen fertilizer use, the optimal condition for variable input Nitrogen ($j = N$) can be derived by,

$$\mathrm{E}[p_t] \cdot f'_{v_{itN}}\left(v_{itN} | v_{itj \neq N}, m_{it}, s_{it}\right) = w_{tN} \qquad 3.5$$

where the optimal condition for nitrogen is where the expected value of marginal product is equal to the input cost for nitrogen (w_{tN}). The value of marginal product for nitrogen can be shown as,

$$\mathrm{E}[p_t] \cdot MP_{v_{itN}} = w_{tN} \qquad 3.6$$

derived as the expected price of output multiplied by the expected marginal product for nitrogen. This can be further simplified as shown in equation 3.7 below.

$$\mathrm{E}\left[VMP_{v_{itN}}\right] = w_{tN} \qquad 3.7$$

Lastly, equation 3.7 can be rewritten so that the optimal condition for variable input nitrogen is where the expected marginal product $MP_{v_{itN}}$ is equal to the input cost output price ratio shown as,

$$MP_{v_{itN}} = \frac{w_{tN}}{\mathrm{E}[p_t]} \qquad 3.8$$

which will be used for the estimation of the economic optimal nitrogen rate for spring wheat production in west-central Saskatchewan.

3.2.7 Maximum Condition for Variable Inputs

While it is assumed that maximizing expected profit is the objective of the producer, the estimation of the yield gap requires the consideration objective of maximizing yield. Using the first order condition, equation 3.4 above, the maximum condition for a variable input can also be determined where output is maximized for a given variable input. The maximum condition for variable input nitrogen ($j = N$) can be derived by,

$$MP_{v_{itN}} = 0 \tag{3.9}$$

where the maximum condition for nitrogen is where the expected marginal product is equal to zero, where an additional unit of the variable input nitrogen does not produce an additional unit of output.

Chapter 4: Data and Sample Selection

4.1 Introduction

This chapter discusses the main sources of data for this analysis including the criteria used for the specific sample of data selected. Additionally, this chapter includes a detailed description of the specific variables formulated from the sample and used in estimation of the production function, specified in Chapter 5.

4.2 Data

This section discusses the sources of the field level yield, management, soil, rainfall, and topsoil moisture data as well as the relevant price data used in estimation and analysis. The field level yield, management, and soil data was generously provided by the Saskatchewan Crop Insurance Corporation (SCIC). SCIC collected the data from producers through the administering of the provincial Agri-insurance program for the period of 2010-2019. Price data for both the bulk cost of nitrogen fertilizer and the expected farm gate price for spring wheat was included in analysis. The bulk cost of nitrogen was collected from the monthly farm input price survey administered by Alberta Agriculture and Forestry. The expected farm gate spring wheat price was collected from the Saskatchewan Crop Planning Guide which is published annually by the Saskatchewan Ministry of Agriculture. In-season rainfall data was generously provided by the Saskatchewan Ministry of Agriculture, who collect and publish in-season weekly rainfall reports by rural municipality (RM) from April until November each year. Lastly, topsoil moisture ratings were also generously provided by the Saskatchewan Ministry of Agriculture, who collect and publish in-season weekly topsoil moisture conditions by RM from April until November of each year.

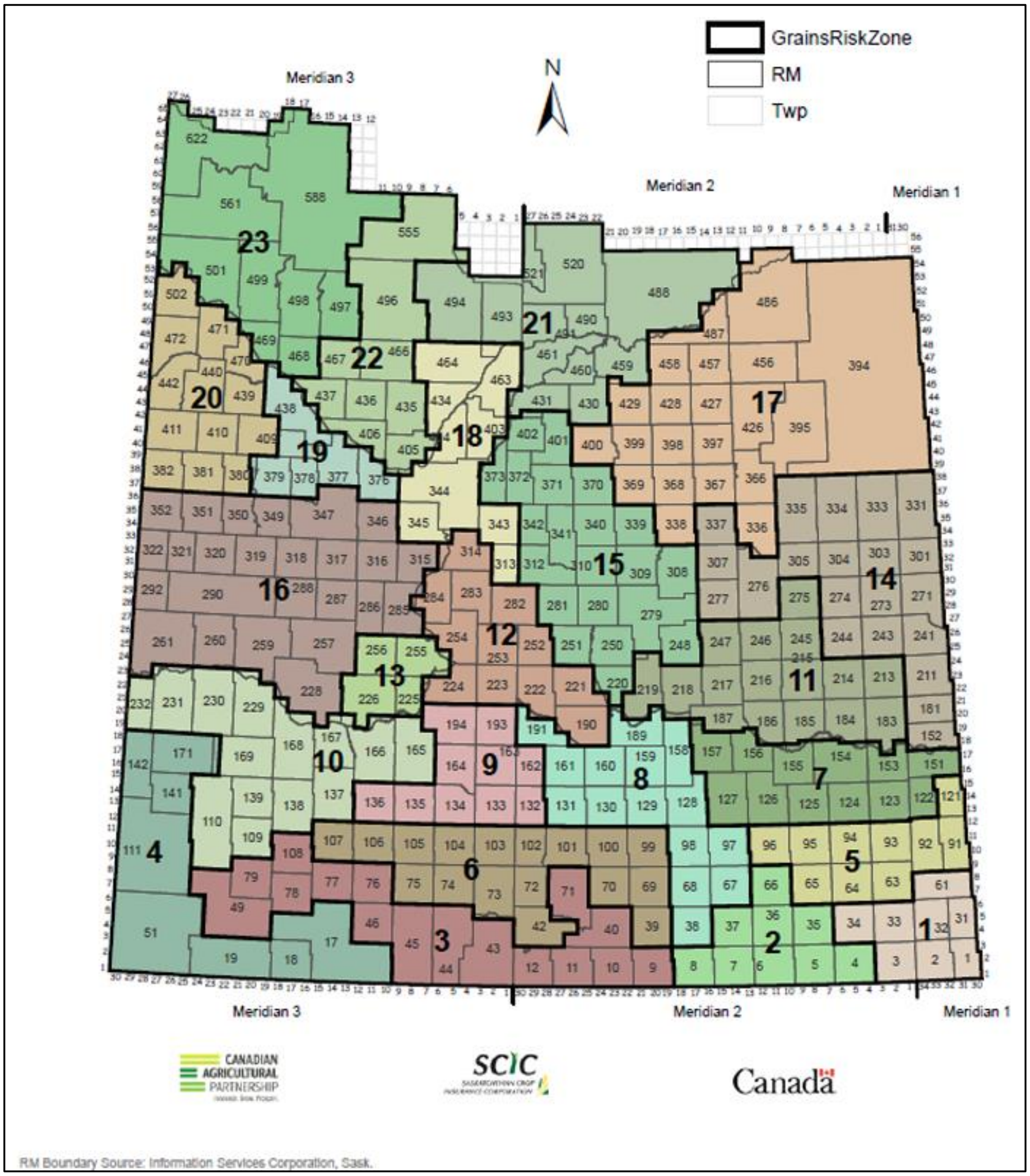
4.2.1 Field Level Data

Field level yield, management, and soil data provided by SCIC is generated annually by producers who enrol in provincial crop insurance. SCIC data is submitted by producers through first a Seeded Acreage Report, and then secondly a Production Declaration following harvest. The Seeded Acreage Report requires producers to include at a minimum, seeded acres by legal land description (LLD) or quarter-section (160 acres) and crop type. The Production Declaration requires producers to report at a minimum, their entire production weight (in kilograms) for each of their insured crops. Producers also have the option to submit additional information by LLD including seed variety, applied fertilizer and chemicals, and yields, all specific to each field or LLD. Producers who provide this additional information are enrolled in the Saskatchewan Management Plus (SMP) program which provides them with additional year-to-year field level summary data and comparable crop performance to their risk area, all at no additional cost. Although management and production information is submitted entirely by the producer, SCIC performs random audits and verifications to ensure reported data is accurate and truthful. Additionally, producers are incentivized to report actual yields, as inflating declared yields would increase their costs of insuring in the following year and deflating reported yields would lower their production guarantee reducing the chance of a claim. Therefore, the data employed in this analysis is assumed to be accurate and correctly specified field level management and production data in Saskatchewan.

For the administration of crop insurance across Saskatchewan, SCIC has divided the province into 23 grain risk zones (Figure 4.1 below). Crops grown within these zones are considered to have the same production risk based on climate, topography, and historical yield factors (Dumanski, J., Cann, M. and Wolynetz, 1992). Several characteristics of insurance vary by risk zone including production guarantees, what crops can be insured, seeding date deadlines, and the cost of insurance premiums.

The SCIC data used for this analysis was selected for spring wheat fields in grain risk zone 16 (see Figure 4.1 below) for the period of 2011-2019, from producers enrolled in the SMP program. As previously discussed in section 1.4, grain risk zone 16 most closely corresponds to AAFC Buffer Zone 21, the region where AAFC is estimating wheat yield potential. Variables from the SCIC dataset include yield, applied fertilizers rates and chemicals, variety, previous crop, producer management index, seeding date, soil rating, RM, and year. An information sharing agreement between SCIC, the University of Saskatchewan, and the thesis author was signed to ensure confidentiality and secure data management.

Figure 4.1: Map of SCIC Grain Risk Zones of Saskatchewan



RM Boundary Source: Information Services Corporation, Sask.

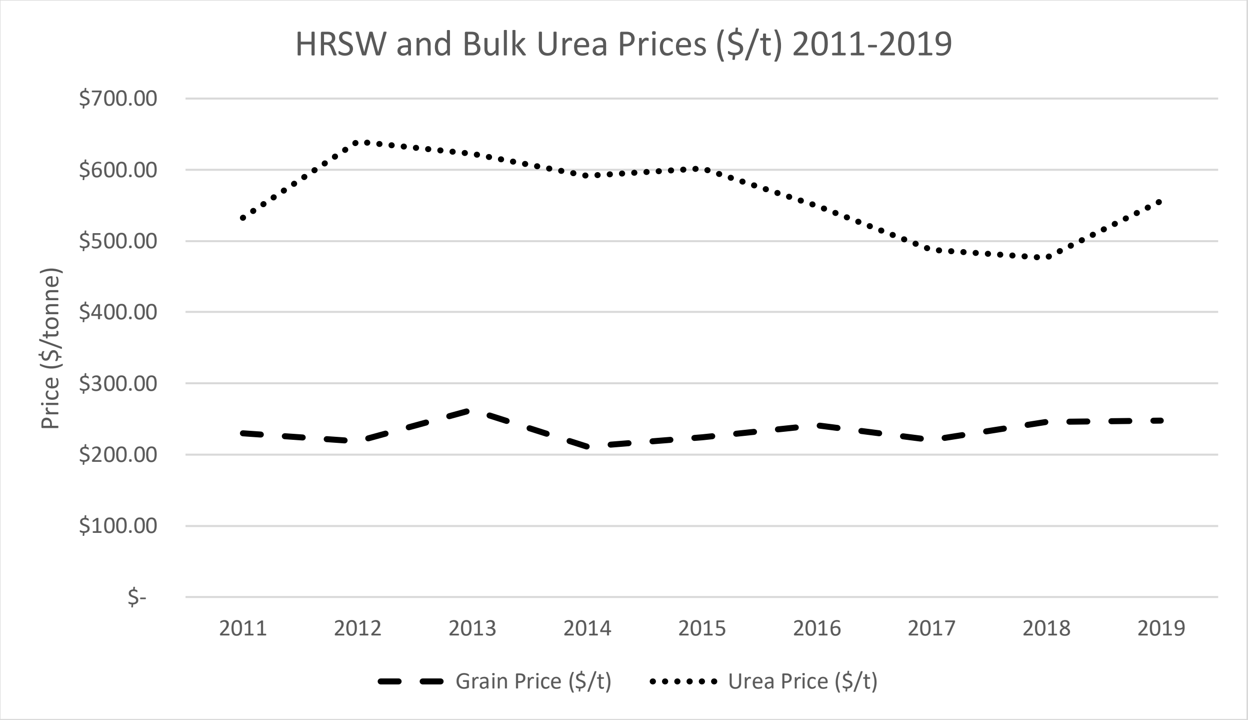
Image source: Saskatchewan Crop Insurance Corporation (SCIC), 2022

4.2.2 Price Data

To capture the price that producers pay for nitrogen fertilizer, data from the monthly farm input price survey collected by Alberta Agriculture and Forestry was used due to a lack of equivalent data for Saskatchewan farms. The monthly farm input price survey tracks both urea and anhydrous-ammonia forms of nitrogen. The reported urea price was chosen as a proxy for nitrogen price. Given that producers typically purchase fertilizer in the fall and early spring before planting, the yearly urea price was calculated as the average of the reported monthly prices from October to December of the previous year and January to March of the current year. The bulk price for urea ($ per ton), which is 46% nitrogen was converted to the price per kilogram of actual nitrogen by dividing the bulk urea price by a factor of 460. The bulk urea price is reported by the dotted line in Figure 4.2 below.

The expected farm gate price for wheat was collected from the Saskatchewan Crop Planning Guide which is published annually by the Saskatchewan Ministry of Agriculture. The expected farm gate price for the upcoming year is based on input from AAFC's annual winter farm income forecast, and from local grain buyer reports that include both spot pricing and contract pricing (Saskatchewan Ministry of Agriculture, 2022). The annual expected farm gate price for wheat is shown by the dashed line in Figure 4.2 below. The expected farm gate price for wheat is historically more stable than the price for nitrogen but appears to faintly lag the up and downward trend for the price of nitrogen fertilizer.

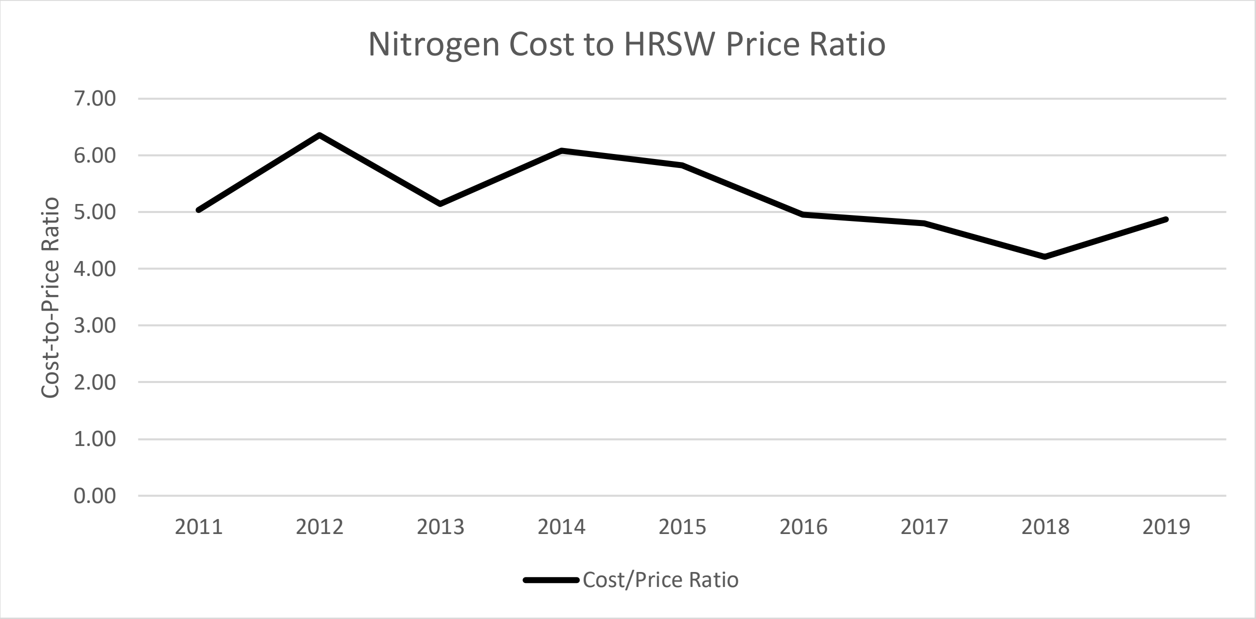
Figure 4.2: Annual CWRS and Bulk Urea Prices per ton from 2011-2019



Source: Author's calculations based on 2019 Alberta Agriculture and Forestry and SMA data.

Both the bulk urea price and the expected farm gate price for wheat were adjusted for inflation to 2019 values using the consumer price index (CPI) as reported by Statistics Canada (Statistics Canada, 2019). Using the historical annual prices for nitrogen and wheat, converted to the dollar per kilogram of actual nitrogen[2] and dollar per kilogram of wheat, annual nitrogen cost to wheat price ratios were calculated. The cost to price ratio, reported in Figure 4.3 below, remains relatively stable across years bottoming-out in 2013 and peaking in 2014.

Figure 4.3: Nitrogen Cost to Wheat Price Ratio, 2011 to 2019



Source: Author’s calculations based on 2019 Alberta Agriculture and Forestry and SMA data.

4.2.3 Rainfall and Topsoil Moisture Data

In-season rainfall data was sourced from the Saskatchewan Ministry of Agriculture’s (SMA) Weekly Crop Report. Each week from April to October, volunteer crop reporters provide information on rainfall, soil moisture, seeding and harvest progress, and yields at the RM level. Assuming that cereal crops reach near full maturity by September, any reported rainfall following September 1st of a given year was added to the next growing seasons rainfall total. It is also important to note that snowfall which is an important component of soil moisture for the

---

[2] The price for actual nitrogen as mentioned above was calculated by taking the bulk urea price ($ per ton) and dividing it by a factor of 460.

next growing season is not captured by the SMA Weekly Crop Report through accumulated rainfall or precipitation but may be reflected in topsoil moisture ratings early in the spring during snow melt. When no rainfall was reported for an RM due to a vacant reporter, the average of all bordering RM's was used to calculate a total. Additionally, rainfall from RM's with multiple reporters were averaged.

Topsoil moisture ratings were also sourced from the SMA Weekly Crop Report. Topsoil moisture ratings that represent the pre-seeding topsoil moisture conditions each spring were taken from the first annually published SMA Weekly Crop Report, usually published in early April of each year. Topsoil moisture ratings range from very short, short, adequate, and surplus.

4.3 Selection of Data

This section begins with a discussion of the representativeness of the SMP yield and management data to general SCIC data and provincial wheat production data. This is followed by a discussion of the data filtering process and criteria that was used to select the sample of data used for analysis.

4.3.1 Representation of SMP Data for Saskatchewan

Given that SMP data within the SCIC dataset must meet specific reporting requirements for producer enrolment, it is important to assess the representativeness of the SMP data relative to both general SCIC data and provincial wheat production data. Table 4.1A below highlights that from 2011-2019, roughly 7% of Saskatchewan's total seeded CWRS wheat hectares are insured through SCIC and are from grain risk zone 16. Of the CWRS insured hectares through SCIC within grain risk zone 16, over half of the seeded hectares are enrolled in the SMP program. Additionally, roughly one-fifth of the insured seeded hectares in grain risk zone 16 are annually planted to CWRS wheat.

Table 4.1A: Provincial, SCIC, and SMP Representative Hectares, CWRS, 2011-2019.

Year	% Of Total Saskatchewan Spring-Wheat Ha. Enrolled in SCIC (from Risk Zone 16)[1]	% Of Insured Spring-Wheat Ha. Enrolled in SMP program (Risk Zone 16)	% Of Total Risk Zone 16 Ha. Planted to Spring-Wheat, annual avg.
2011	9%	52%	23%

2019	9%	54%	24%
2011-2019	7%	56%	20%

Source: [1]Total Saskatchewan Spring-wheat hectares taken from Statistics Canada *Table 32-10-0359-01: Estimated areas, yield, production, average farm price and total farm value of principal field crops, in metric and imperial units* from 2011-2019 for Saskatchewan.

Additionally, Table 4.1B below highlights that SCIC data from the SMP program is upwardly biased given higher average wheat yields from 2011-2019 when compared to both risk zone 16 SCIC non-SMP and provincial (Statistics Canada) wheat yield data. It is important to note that this may result in biased results when estimating the wheat yield response to nitrogen using SMP data from risk zone 16 and extrapolating to the larger Saskatchewan wheat production scope. However, this rich dataset, unique to the SMP program, provides several benefits for understanding wheat production in Saskatchewan.

Table 4.1B: Provincial, SCIC, and SMP Representative Hectares, CWRS, 2011-2019.

Year	**Saskatchewan Avg. Spring Wheat Yield (kg/ha)[1]**	**SCIC Reported Spring Wheat Yield (kg/ha) (from Risk Zone 16)**	**SMP Program Reported Avg. Spring Wheat Yield (kg/ha) (from Risk Zone 16)**
2011-2019	2,935	2,955	3,124

Source: [1]Total Saskatchewan Spring-wheat hectares taken from Statistics Canada *Table 32-10-0359-01: Estimated areas, yield, production, average farm price and total farm value of principal field crops, in metric and imperial units* from 2011-2019 for Saskatchewan.

4.3.2 Selected Data Sample

The data used for this analysis was selected for by focusing specifically on CWRS wheat fields within grain risk zone 16 for the period of 2011-2019, from producers enrolled in the SMP program. As previously discussed in section 1.4, grain risk zone 16 most closely corresponds to AAFC Buffer Zone 21, the region where AAFC is estimating wheat yield potential, focusing this analysis on that particular area. Observations with missing information or extreme values, especially for variables required for analysis were removed.

Data was filtered for several criteria. Irrigated fields were excluded due to the distinction between dryland and irrigated production and the subsequent differences between potential yield and water-limited potential yield, as well as that the majority of wheat production in grain risk zone 16 dryland production. Extreme values for observed yields as well as applied nitrogen, phosphorus, potassium, and sulphur rates were excluded based on reasonable rates for Saskatchewan production. There is the potential that producers mistakenly reported applied rates rather than actual nutrient content in some instances. As noted in section 4.2.2 above, nitrogen in the form of urea for example is 46% nitrogen meaning that a producer applying 100 kg per hectare of nitrogen is applying 46 kg per hectare of actual nitrogen. SCIC asks producers to report actual nutrient content applied which requires producers to do their own conversions. Additionally, because nitrogen is of interest for this analysis, fields were selected for that had positive rates of nitrogen applied and recorded yields above 670 kg/ha^{-1}. Fields that were reported seeded approximately 60 days before or after May 14th of each year were excluded. Field size was filtered for fields that were at least 32 hectares in size to effectively remove any potential trial fields. As previously mentioned, observations with missing information for necessary variables including the fields previous crop or crop variety were also excluded. These criteria were selected by observing potential errors in data entry, the decision to avoid any field level plots or experimental trials that may record small field sizes or extreme values, and the need to have specific previous cropping and varietal information. Given that the same fields are not necessarily observed each year, the panel dataset is unbalanced, containing observations from 732 farms on 5,019 spring-wheat fields over 9 years (2011-2019), for a total of 7,537 observations.

4.4 Description of the Model Variables

This section highlights the variables included in the model followed by a brief discussion of each variable including their respective means or weight by each year. Table 4.2 below describes each variable, what the variable measures, and at what level the variable is collected at.

Table 4.2: Description of variables included in production function.

	Variable	**Description**
Continuous	Spring Wheat Yield[1]	Average spring wheat yield for field (kg ha^{-1})
	Nitrogen[1]	Applied actual nitrogen rate (kg ha^{-1})
	Phosphorous[1]	Applied actual phosphorous rate (kg ha^{-1})
	Potassium[1]	Applied actual potassium rate (kg ha^{-1})
	Sulphur[1]	Applied actual sulphur rate (kg ha^{-1})
	Seeding Date[1]	Seeding date as days before or after May 14th
	Variety Yield Index[1]	Yield index of chosen variety relative to Carberry, the check variety
	Management Index[1]	Ratio of Customer Long-Term Average Yield relative to Area Long-Term Average Yield
	Rainfall[2]	Total in-season rainfall (April-September) measured in millimeters (mm), by RM
Dummy	Topsoil Moisture Rating[2]	"Short" or "Adequate" rating for topsoil moisture prior to seeding (early April), by RM, 1=Short, 0=Surplus
	Herbicide Use[1]	Dummy variable, 1=herbicide applied, 0=herbicide not applied
	Fungicide Use[1]	Dummy variable, 1=fungicide applied, 0=fungicide not applied
	Insecticide Use[1]	Dummy variable, 1=insecticide applied, 0=insecticide not applied
	Seed Treatment Use[1]	Dummy variable, 1=seed treatment applied, 0=seed treatment not applied
	Cereal Previous Crop t-1[1]	Cereal crop grown on the field in the previous year
	Oilseed Previous Crop t-1[1]	Oilseed crop grown on the field in the previous year
	Pulse Previous Crop t-1[1]	Pulse crop grown on the field in the previous year
Categorical	Soil Group[1]	Grouped SCIC Soil Productivity Rating, 1=high production, 3=low production
	Year[1]	Year of observation

[1]Source: (SCIC, 2019), [2]Source: (Saskatchewan Ministry of Agriculture, 2022).

4.4.1 Yield

The dependent variable, *actual yield* is measured as the observed yield per hectare of field *i* in year *t*. It is calculated by dividing the fields total production (in kilograms) by the reported seeded hectares for field *i*. This calculation differs slightly from that used by the Global Yield Gap Atlas approach (van Dijk et al., 2017), which uses yield per hectare harvested rather than seeded. While using yield per hectare harvested rather than seeded has the benefit of accounting for those hectares that do not produce a crop due to flooding, hail, or pest damage and are not harvested, SCIC only requires producers to report seeded acres per LLD. Therefore, it is difficult to remove field observations that were only partially harvested due to flooding, hail, or pest damage.

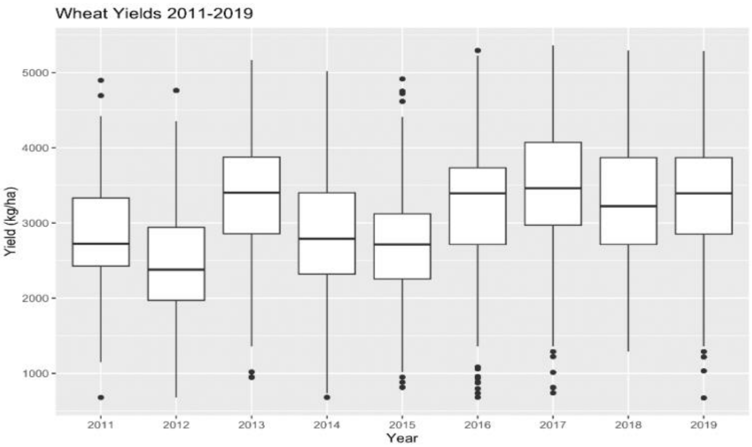
Table 4.2 shows the computed average spring-wheat yield per hectare by year. While yields appear to be trending upwards, there appears to be no significant jump in average yields from one year to the next. The average spring-wheat yield is lowest in 2012 and highest in 2017. It is important to note that grain quality is also a component of successful wheat production, even at high yields. However, grain quality is not a part of this analysis.

Table 4.3: Mean yield of Spring-wheat by year (kg ha^{-1}), RZ 16.[3]

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Yield (kg ha^{-1})	2827	2455	3354	2874	2718	3242	3466	3291	3352

Source: (SCIC, 2019)

Figure 4.4: Wheat yields (kg ha^{-1}), 2011-2019



Source: (SCIC, 2019).

---

[3] Note: The mean yield above, is calculated for fields that meet the criteria discussed in section 4.3.

4.4.2 Fertilizer

This section highlights fertilizer application rates for nitrogen, phosphorous, potassium, and sulphur from 2011-2019 for wheat production in grain risk zone 16. Additionally, the distinction between applied fertilizer and residual nutrients already in the soil is discussed.

4.4.2.1 Applied Actual Nutrients and Residual Nutrients

SCIC allows producers to report by LLD the amount of actual nutrients applied for nitrogen, phosphorous, potassium, and sulphur. Depending on the form of nutrients applied, whether it be granular, liquid, or gas, producers must convert their applied rate to the actual nutrient rate to calculate how much nutrients are actually being applied to the crop. For example, a producer applying 100 kg ha^{-1} of urea (46% nitrogen) is in fact applying 46 kg ha^{-1} of actual nitrogen to the crop. However, the applied fertilizer variables (N, P, K, S) fail to capture the amount of nutrients completely available in the soil (referred to as residual nutrients), or the amount absorbed by the plants during the growing season. While this is a potential weakness of the SCIC data and our subsequent analysis, the fields soil characteristics and previous crop type which impact soil nutrient availability have been included as controls. Additionally, over the past five years only 31% of wheat growers in western Canada reported soil sampling for soil nitrogen availability annually, indicating that most producers do not base their fertility decisions off residual nutrient availability (Fertilizer Canada, 2022). It is also challenging to accurately estimate how much of the applied nutrients are absorbed by the plant throughout the growing season. Nitrogen is considered to be quite mobile throughout the soil, largely being used up within the growing season. Therefore, it is assumed that producers view nitrogen as necessary to be applied each crop year. Phosphorous, however, is considered to be fairly immobile in the soil (Government of Saskatchewan, 2021). Additionally, potassium typically only shows yield responses under severe potassium deficiencies (Saskatchewan Ministry of Agriculture, 2022). Therefore, applied phosphorous, potassium, and sulphur are assumed to be less available during the growing season but positively contribute to their respective nutrient reservoirs. Therefore, this analysis focuses only on the optimization of applied nitrogen while still controlling for the contributions of other nutrients.

4.4.2.2 Nitrogen

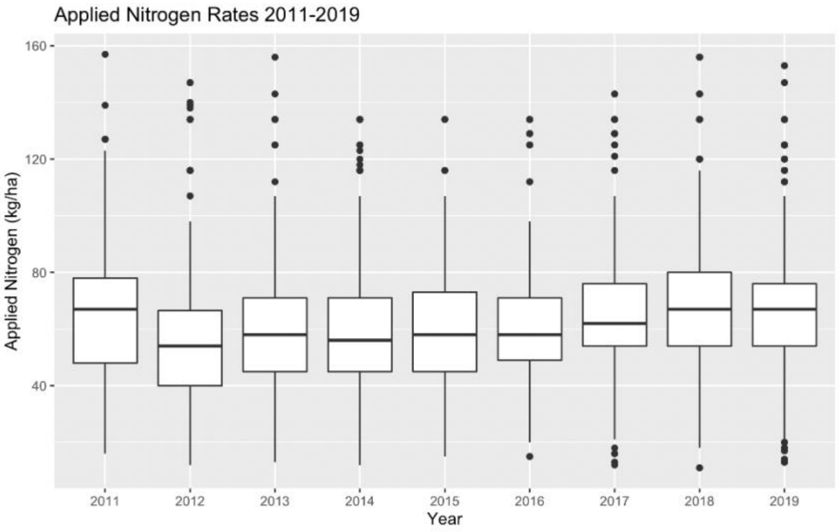
Following water in dryland agriculture production, nitrogen is considered the most common limiting factor in crop production across western Canada (Farrell, 2020). As previously discussed, the SCIC data includes producer reported applied actual rates of nitrogen for each field, in each year. The average applied rate of actual nitrogen has been increasing since 2012, as shown in Table 4.4. The high average applied rate of actual nitrogen rate in 2011, followed by a significant decrease in the average applied rate is consistent with the price trend for bulk urea as shown in Figure 4.2 above. Bulk urea prices peaked in 2012 followed by a down trend to 2019 which may help to explain the upward trending average applied rate.

Table 4.4: Mean applied rate of Nitrogen by year (kg ha^{-1}), RZ 16.[4]

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Applied N (kg ha^{-1})	66	56	58	60	60	63	65	67	67

Source: (SCIC, 2019)

Figure 4.5: Applied nitrogen (kg ha^{-1}), 2011-2019



Source: (SCIC, 2019).

---

[4] Note: the mean applied rates of nitrogen represent actual applied nitrogen.

4.4.2.3 Phosphorous, Potassium, and Sulphur

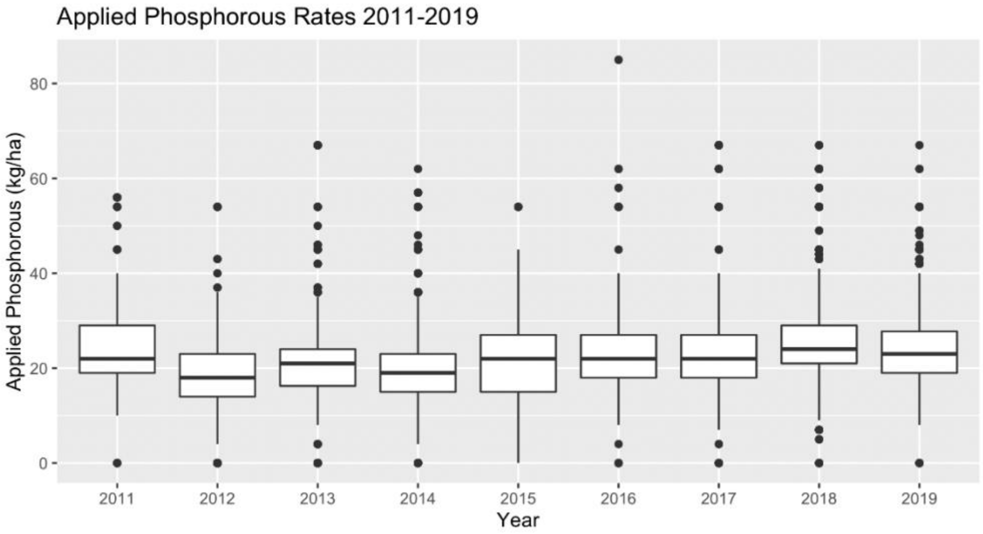
Like nitrogen, phosphorous (P), potassium (K), and sulphur (S) are producer reported actual applied rates. The average applied rates of actual phosphorous, potassium, and sulphur have been increasing since 2012, as shown in Table 4.5. Phosphorous is the second most applied nutrient, with rates significantly higher than that of both potassium and sulphur. While the average rates appear quite low relative to nitrogen, there is significant variation in the amount of phosphorous, potassium, and sulphur observed being applied.

Table 4.5: Mean applied rate of Phosphorous, Potassium, and Sulphur by year (kg ha^{-1}), RZ 16.[5]

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Applied P (kg ha^{-1})	26	19	21	21	22	24	24	27	24
Mean Applied K (kg ha^{-1})	1	1	2	1	2	1	2	2	3
Mean Applied S (kg ha^{-1})	3	3	4	3	4	5	4	5	6

Source: (SCIC, 2019)

Figure 4.6: Applied phosphorous (kg ha^{-1}), 2011-2019

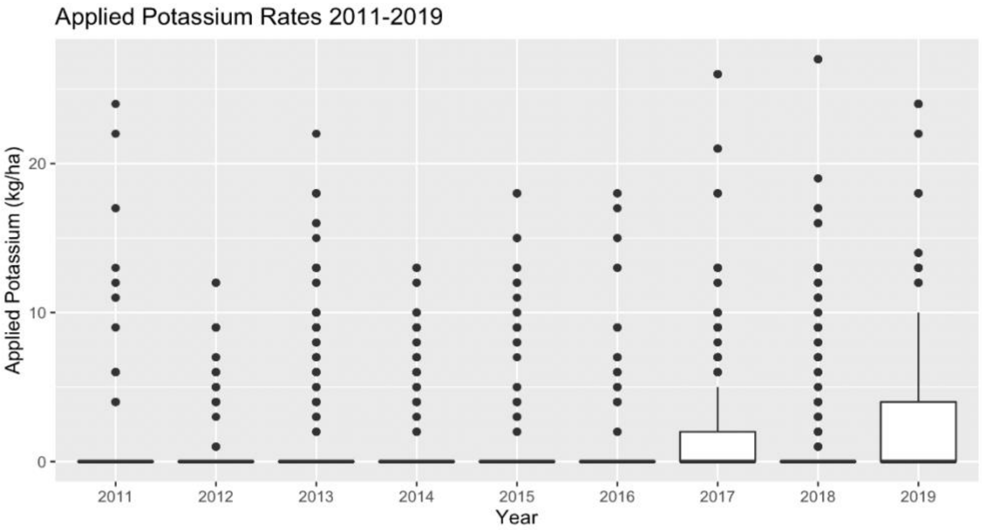


Source: (SCIC, 2019).

---

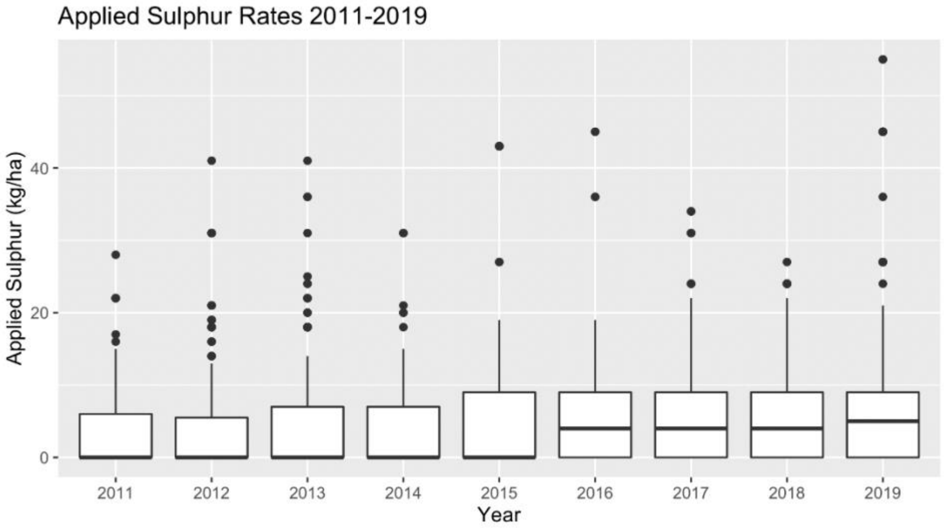
[5] Note: the mean applied rates of phosphorous, potassium, and sulphur represent actual applied nutrients.

Figure 4.7: Applied potassium (kg ha^{-1}), 2011-2019



Source: (SCIC, 2019).

Figure 4.8: Applied sulphur (kg ha^{-1}), 2011-2019



Source: (SCIC, 2019).

4.4.3 Seeding Date

Seeding date is an important management factor and is expected to have an impact on yield. The optimal seeding date is considered to be subjective to both environmental conditions in a given year and the logistical constraints of the entire farm including any other crops that are planted.

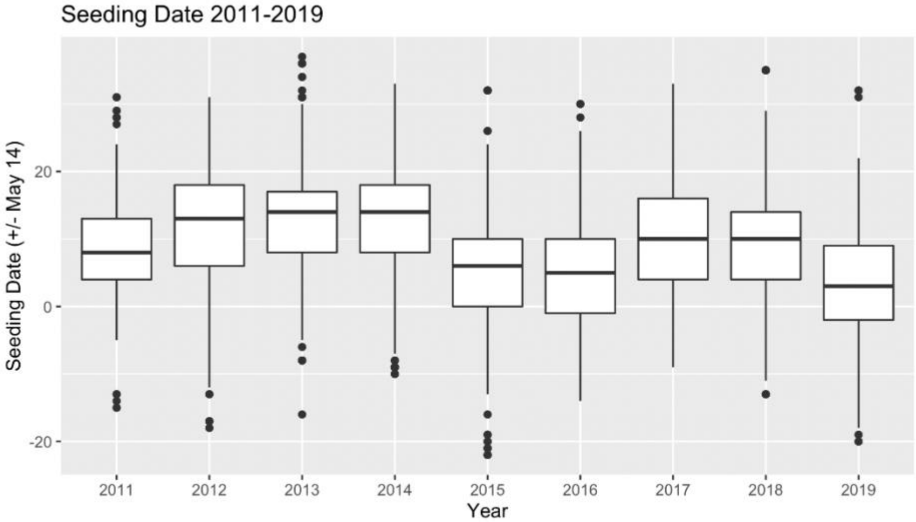
Fields that are seeded too early are at potential risk of frost and cooler temperatures as well as the potential for parts of the field to be impassable early in the spring. However, ultra-early seeding is a relatively new practice being researched. Fields seeded too late have increased risk of late season frost and heat blast during flowering. Seeding date is calculated as the days before or after May 14th of each year. As shown in Table 4.6, average seeding date varies each year and is likely to be heavily influenced by the environmental conditions present in that year.

Table 4.6: Average Seeding Date (days before/after May 14), 2011-2019.[6]

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Seeding Date (+/- days from May 14)	9	12	13	12	5	5	10	9	3

Source: (SCIC, 2019)

Figure 4.9: Seeding date as days before or after May 14, 2011-2019



Source: (SCIC, 2019).

---

[6] Note: the mean seeding date is calculated as positive days after or negative days before May 14 of each year.

4.4.4 Variety Yield Index

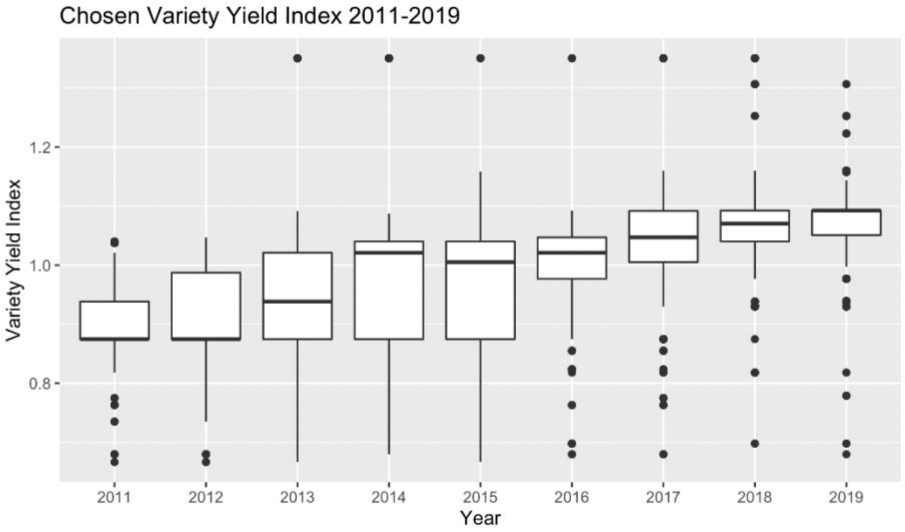
Variety choice is an important management decision cereal producers must make each season, with different varieties providing different genetic characteristics such as increased disease or pest resistance and lodging resistance for example. Producers in Saskatchewan have access to over 60 different spring wheat varieties (Syme, An, & Torshizi, 2023). To account for the role that variety plays in production, a variety yield index variable, which measures how the average yield of the chosen variety compares to a chosen check variety (AC Carberry) is calculated. The variety yield index is calculated as the average yield for a specific variety *v*, divided by the average yield for the check variety, Carberry. As shown in Table 4.7, the average variety yield index is increasing over time as producers adopt newer and higher yielding varieties, relative to the check variety, AC Carberry which was released in 2011.

Table 4.7: Average Variety Yield Index relative to Check Variety (AC Carberry), 2011 to 2019.

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Variety Yield Index	0.90	0.90	0.94	0.97	0.97	0.99	1.04	1.07	1.08

Source: (SCIC, 2019)

Figure 4.10: Variety yield index, 2011-2019



Source: (SCIC, 2019).

4.4.5 Management Index

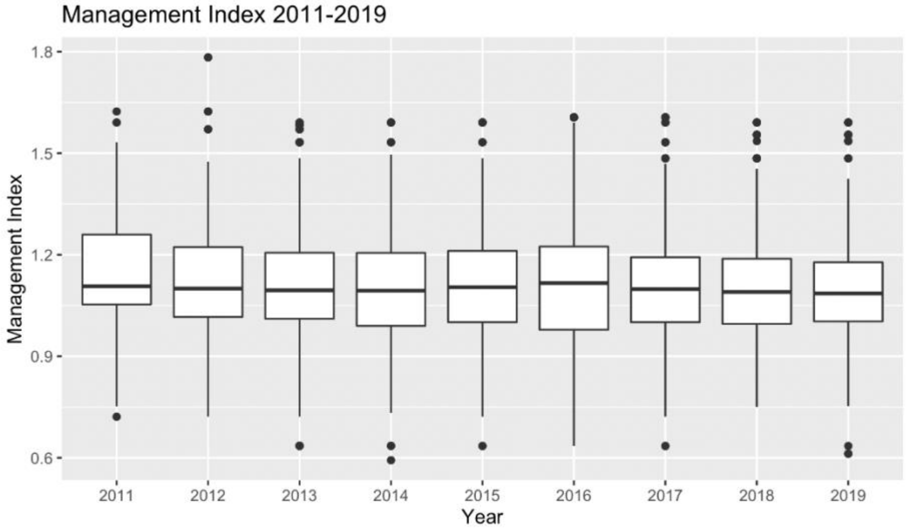
Characteristics of the farm and farmer that influence management decisions and ultimately crop yields commonly include age, education, value of assets, seeding equipment, and farm size to attempt to capture unobservable management skills and ability (Sheahan et al., 2012; van Dijk et al., 2017). Due to a lack of available data for specific farm and farmer characteristics in Saskatchewan, a management experience index variable is included that acts as a proxy for farm and farmer characteristics. The management experience index variable is calculated by taking the ratio of a particular producer's long-term individual average yield relative to the long-term area average yield. The SCIC dataset provides long-term individual yields for each producer for each crop, calculated by exponentially weighting historical producer-specific yields. Similarly, the SCIC dataset provides long-term area average yields for each risk zone for each crop using the same method of exponentially weighting historical yields. For each producer, the management experience index is calculated for each crop recorded in the SCIC dataset from 2011-2019. Therefore, the calculated management experience index represents the all-crop aggregate management index for each producer. Producers who have a management experience index greater than 1 are considered to have better yield experience than the area average. While this proxy variable does not capture characteristics such as those listed above, it does allow for a more direct comparison across farmers within a region, through their respective yield experiences.

Table 4.8: Average Management Index, 2011 to 2019.

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Management Index	1.13	1.11	1.12	1.10	1.11	1.12	1.11	1.10	1.10

Source: (SCIC, 2019)

Figure 4.11: Management index, 2011-2019



Source: (SCIC, 2019).

4.4.6 In-season Rainfall

In-season rainfall, especially in the months of May, June, and July has been shown to be important to determining wheat yields (Robertson, 1974). While rainfall is important for plant growth early in the year, significant rainfall later in the year can increase disease pressure reducing yield or reduce the quality of the grain through increased lodging or sprouting of mature wheat kernels. In-season rainfall (mm), taken from the SMA's Weekly Crop Reports for the months of April, May, June, July, and August is included in the model to capture rainfall during the crop growth stage. Any recorded rainfall in the months of September and October is included in the next years rainfall totals to account for contributions to soil moisture reserves for the next crop.

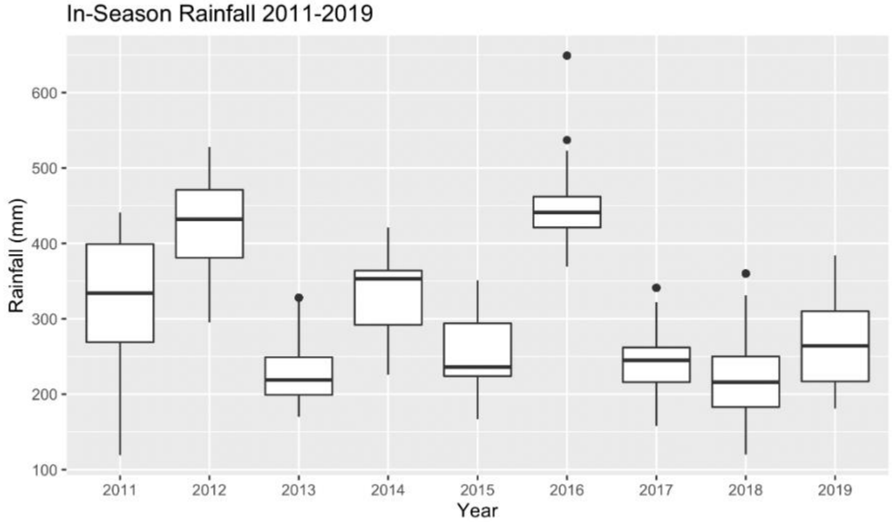
In-season rainfall data from the SMA's Weekly Crop Report is reported at the RM level. Table 4.9 below shows the annual average in-season rainfall, averaged across all RM's within risk zone 16. In the rainfall data, 2018 reported the least amount of rainfall while 2016 reported the most rainfall.

Table 4.9: Average In-season Rainfall (mm), 2011 to 2019.

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019
Mean Rainfall (mm)	316	422	229	332	249	445	244	228	266

Source: (SMA, 2019)

Figure 4.12: Average In-season Rainfall (mm), 2011 to 2019.



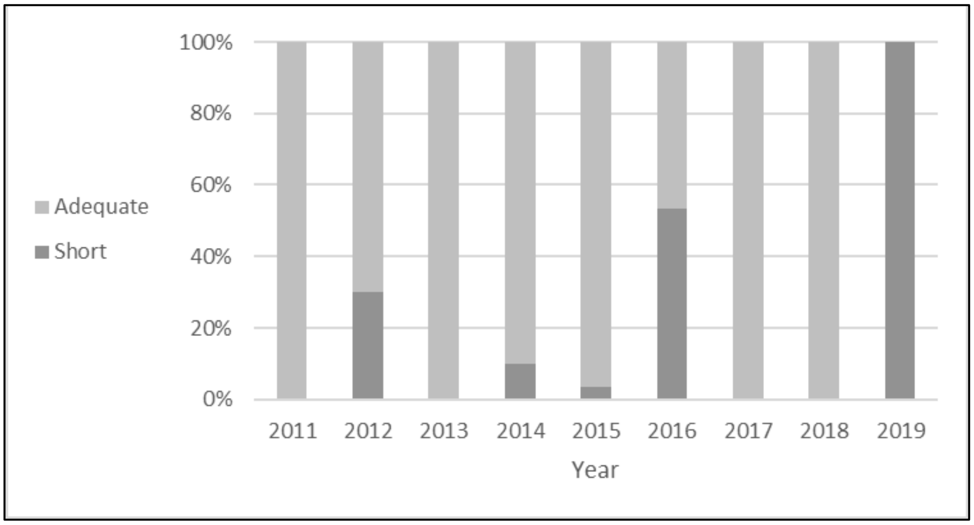
Source: (SMA, 2019).

4.4.7 Topsoil Moisture Rating

Topsoil moisture ratings, sourced from the SMA's Weekly Crop Report are also included in the model to capture soil moisture conditions prior to seeding. Topsoil moisture conditions prior to seeding likely influence producers' fertility and cropping decisions.

Topsoil moisture ratings that represent the pre-seeding topsoil moisture conditions each spring were taken from the first annually published SMA Weekly Crop Report, usually published in early April of each year. Topsoil moisture ratings range from very short to short to adequate to surplus. From 2011 to 2019, most reported ratings in grain risk zone 16 were either short or adequate, therefore the few observations that were rated as very short were categorized as short while those rated surplus were categorized as adequate. Topsoil moisture rating is included as a dummy variable (1=short topsoil moisture, 0=surplus topsoil moisture). Figure 4.13 highlights the variation from 2011 to 2019 in topsoil moisture, dominated largely by adequate or better ratings early in the growing season until 2019.

Figure 4.13: Topsoil moisture rating observations, 2011 to 2019.[7]



Source: (SMA, 2019)

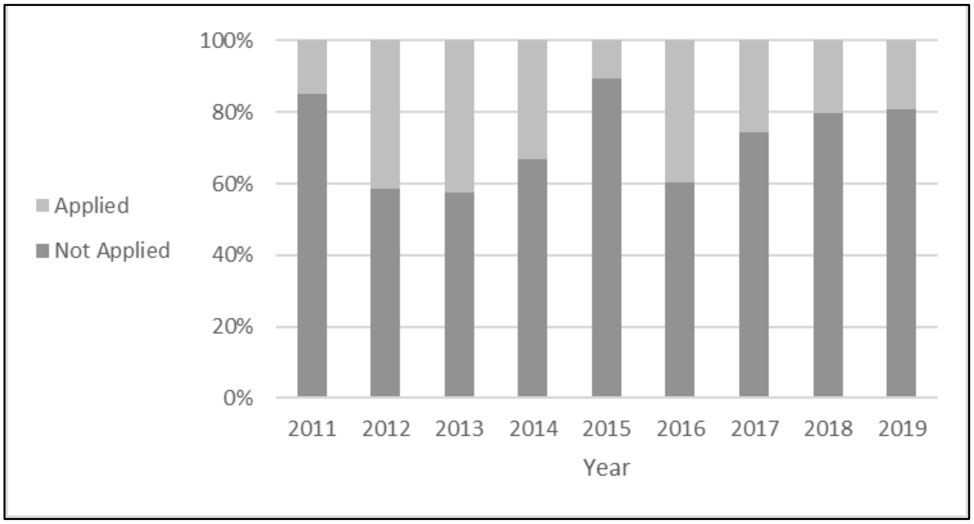
4.4.8 Fungicide

Fungicide application is considered an important management characteristic that can impact wheat yield and quality through the prevention or mitigation of disease. To capture fungicide application, a dummy variable (1=applied, 0=not applied) is included in the model. The years 2012, 2013, and 2016 reported the highest proportion of fungicide application in the dataset. This is not surprising given that 2012 and 2016 also reported the highest in-season rainfall amounts in the dataset leading to the increased risk of such disease as fusarium head blight (FHB). Fungicide application in 2013 may have been in response to widespread applications of fungicide in 2012, rainfall events, or increased humidity later in the growing season, especially at the flowering stage prompting disease prevention measures.

---

[7] Note: Adequate rating is considered adequate or better while short is considered short or worse.

Figure 4.14: Proportion of fungicide applications, 2011 to 2019.

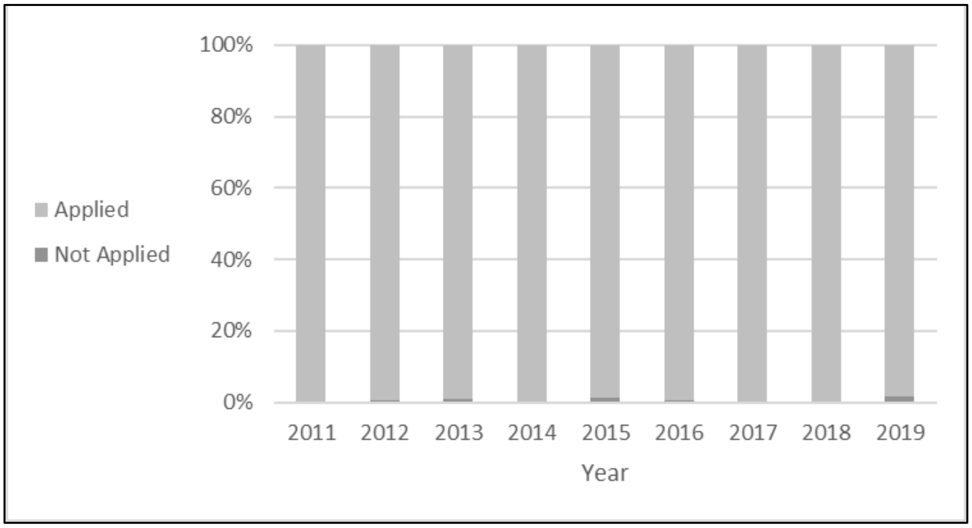


Source: (SCIC, 2019)

4.4.9 Herbicide, Insecticide, and Seed Treatment

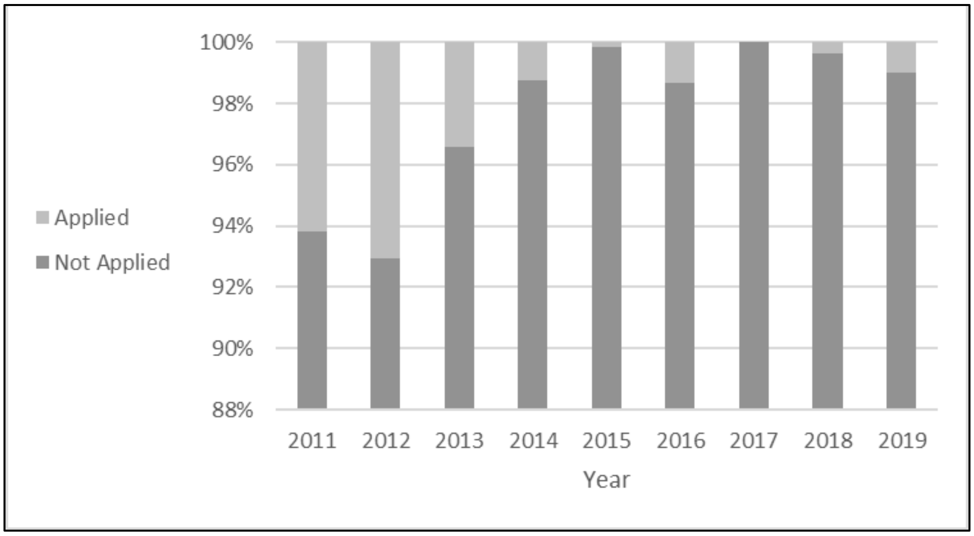
Herbicides, insecticides, and seed treatments, which are used to manage yield reducing factors are also included in the model to capture the impact of application on wheat yields. To capture herbicide, insecticide, and seed treatment applications, dummy variables (1=applied, 0=not applied) are included in the model. Herbicides are applied on almost all fields in each year in the dataset. Insecticides are rarely reported being applied according to the dataset, with a slightly higher percentage of application in 2011 and 2012, potentially indicating increased insect pressure or more susceptible wheat varieties, or both. Similarly, seed treatment is rarely reported being applied according to the dataset, but it is unclear if this is due to a lack of reporting by producers or a lack of the use of seed treatment.

Figure 4.15: Proportion of herbicide applications, 2011 to 2019.



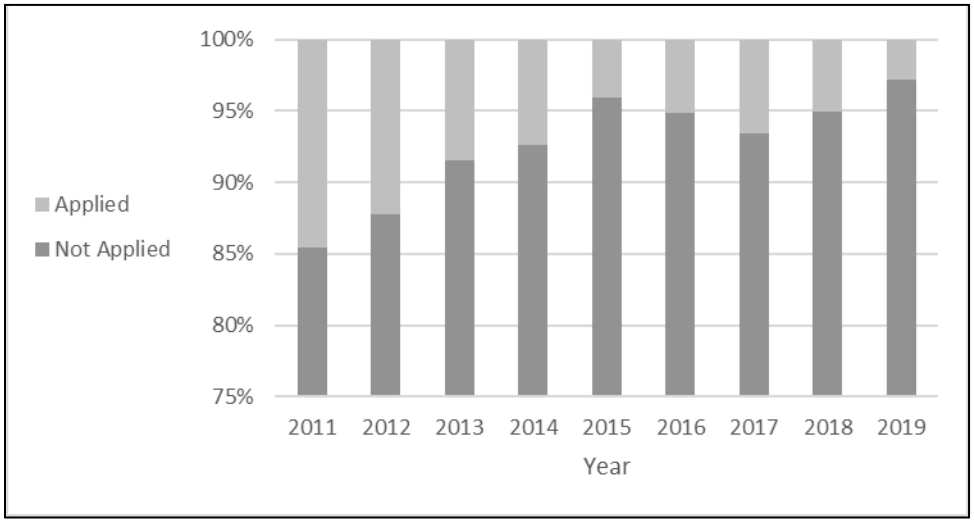
Source: (SCIC, 2019)

Figure 4.16: Proportion of insecticide applications, 2011 to 2019.



Source: (SCIC, 2019)

Figure 4.17: Proportion of seed treatment applications, 2011 to 2019.

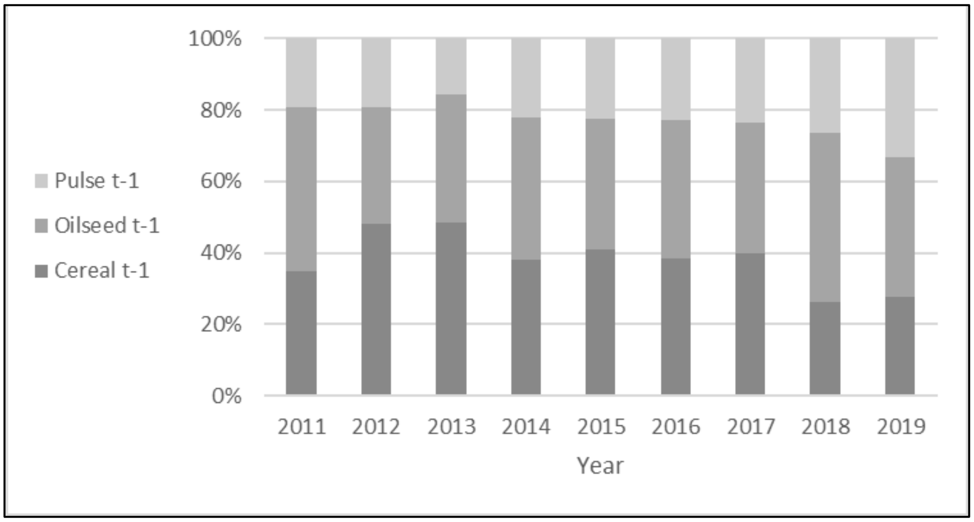


Source: (SCIC, 2019)

4.4.10 Previous Crop

Crop rotation or the fields previous crop is considered an important characteristic that impacts the yield and quality of subsequent crops. Crop rotations can provide both agronomic and economic benefits and costs for the fields next crop, impacting disease and weed pressure, and input requirements. The fields previous crop is classified as a cereal crop, oilseed crop, or pulse crop. Cereal crops found in the SCIC dataset include wheat (HRSW, CPS, CWES, HWS, Winter), durum, barley, oats, rye, and triticale. Oilseed crops include canola and flax. Pulse crops include lentils and field peas. Oilseed previous crops are the most common previous crop in the dataset followed by cereal previous crops and lastly pulse previous crops.

Figure 4.18: Previous crop observations, 2011 to 2019.



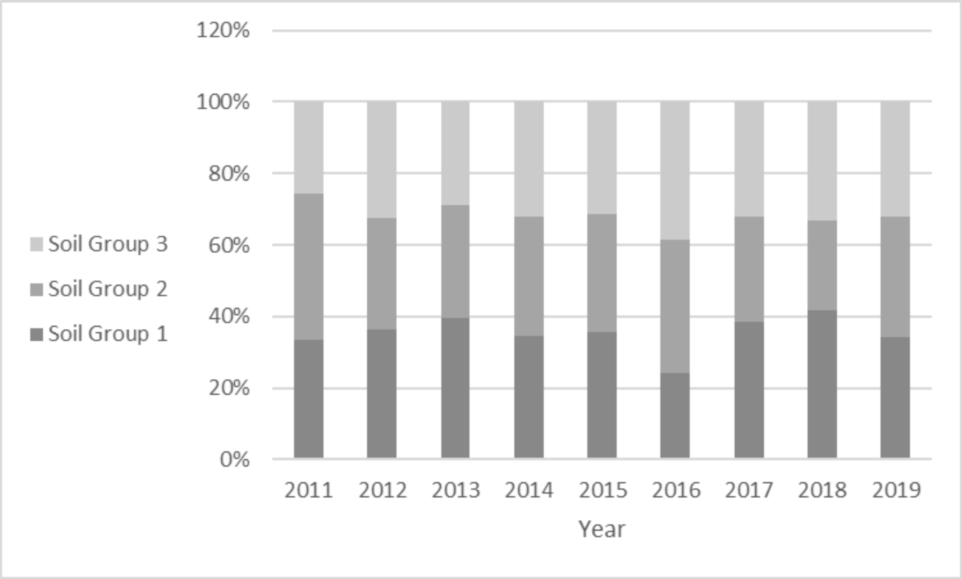
Source: (SCIC, 2019)

4.4.11 Soil Productivity Grouping

Soil productivity characteristics are important for understanding and the yield of a crop. SCIC assigns a soil productivity rating to each LLD to adjust the program including for yields and premiums. The SCIC soil productivity rating classifies LLD's on a scale from A-P, with A assigned to the most productive soils and P assigned to the least productive soils. The soil productivity ratings were provided by the Saskatchewan Assessment Management Agency (SAMA) based on soil surveys and field characteristics with a criterion regarding climate, soil association, soil texture, organic matter, salinity, drainage, and flooding history all influencing the rated productivity. Specific to risk zone 16, no A, I, or N soils are reported in the dataset, while soils H, G, and J were most reported.

To better handle estimating the response of 13 discrete soil productivity ratings, soil ratings are placed into three groups. The groupings are determined based on maintaining ordinal ranking and ensuring balance in the number of observations within each group. The specific characteristics of each soil productivity rating are unknown, yet fields with a soil rating closer to A are classified as more productive relative to fields with a soil rating closer to P. Soil group 1 contains field observations from rated soils B, C, D, E, and F, accounting for 2,749 observations. Fields with an A ranked soil productivity rating were not observed. Given that risk zone 16 is predominantly made up of medium or average productive soils, soil grouping 2 contains observations from productivity rated soils G and H, accounting for 2,393 observations. The least productive soil grouping, soil group 3 contains observations from productivity rated soils J, K, L, M, O, and P, accounting for 2,395 observations.

Figure 4.19: Soil group observations, 2011 to 2019.



Source: (SCIC, 2019)

4.5 Fixed Effects

Additional factors that likely impact crop yields include machinery, labor, seeding rate and depth, timing of fertility and chemical applications, temperature (commonly measured as growing degree days or GDD), and numerous soil characteristics. Due to data constraints, several such additional factors that may impact yield could not be included in the model.

To account for variation across growing conditions and other factors that vary by year, *year* fixed effects for 2011 to 2019 are included in the model to control for differences in growing conditions from year to year. Tables 4.3 and 4.4, and figures 4.4 and 4.5 above highlight the fluctuations and increasing trends for both wheat yields and applied nitrogen rates over time.

To account for variation across the different soil characteristics in the dataset, soil productivity rating fixed effects are considered in the model for productivity ratings B to P.

Table 4.10 below presents the variation in average wheat yields and average nitrogen rates by soil productivity rating from the dataset, as the average from 2011 to 2019 for both measures.

Table 4.10: Average wheat yields and nitrogen rates by soil productivity rating, risk zone 16, averaged for 2011 to 2019.

Soil Rating	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	*L*	*M*	*O*	*P*
Avg. Yield (kg/ha)	3226	3221	3148	3132	3366	3153	3088	3055	2985	2994	2760	2688	2854
Avg. N (kg/ha)	60	60	58	58	64	63	63	65	60	65	69	66	88

Source: (SCIC, 2019)

Additionally, to account for variations in farms and farmers such as in farm machinery, labor, and management including the timing of different fertilizer and chemical applications, producer fixed effects are considered in the model. In the selected sample of data, from 2011-2019 there are 723 different producer ID's. However, not all producers grow wheat each year, and not all producers have multiple fields of wheat when they do appear in the dataset for a given year. Of those producers that do have multiple wheat fields in a given year, the majority do not vary nitrogen fertilizer rates across fields. Table 4.11 below shows that no more than 10% of producers within each year from 2011 to 2019 vary nitrogen rates across wheat fields. Given that the producer ID fixed effect selects for producers who vary nitrogen rates across their wheat fields within a year and that there are very few producers who report varying nitrogen rates across fields each year, the resulting dataset would be quite limited in observations.

Table 4.11: Producer N Variation by Year, 2011-2019

Year	# of Producers	% of Prod. with N Variation	Mean S.D. for Prod. with N Variation	Mean Variance for Prod. with N Variation
2011	142	10%	1.75	24.9
2012	227	10%	1.46	31.6
2013	320	9%	1.14	19.3
2014	240	10%	1.00	12.4
2015	192	8%	0.976	14.7
2016	120	6%	0.709	8.3
2017	267	9%	1.14	18

2018	236	6%	0.562	6.28
2019	229	9%	0.908	14
Average		***8.6%***	***1.1***	***16.6***

Source: Author's calculations based on SCIC 2019 data.

4.6 Summary Statistics

Summary statistics for the included model variables are shown in Table 4.10 below. The average wheat yield in the sample was 3,124 kg ha^{-1}. The average applied nitrogen rate was 62 kg ha^{-1}, which was the highest among all fertilizer types in the sample. The average seeding date was 9 days after May 14th. The average variety had a yield index of 0.997 which was slightly below the yield index for the check variety AC Carberry. The average management index was 1.11 signifying that the average producer in the sample have a yield experience greater than the area average. The average amount of rainfall in-season was 287 mm but ranged from 119 mm up to 649 mm. Most field observations had a surplus topsoil moisture rating (78%) and had herbicide applied (99%). Less than a third of the field observations received a fungicide application (28%), and very few fields received either an insecticide (2%) or seed treatment (7%) application. Oilseed crops (39%) were the most common previous crop, followed by cereal crops (38%) and pulse crops (23%).

Table 4.12: Summary Statistics, 2011 to 2019

Variable	N	Mean/Mode	Std. Dev.	Min	Max
Spring Wheat Yield (kg ha^{-1})	7,537	3,124	830	674	5363
Nitrogen (kg ha^{-1})	7,537	62	23	11	157
Phosphorous (kg ha^{-1})	7,537	23	11	0	85
Potassium (kg ha^{-1})	7,537	1.7	4	0	27
Sulphur (kg ha^{-1})	7,537	4.2	5.7	0	55
Seeding Date	7,537	9	9	-22	37
Variety Yield Index	7,537	0.997	0.11	0.66	1.35
Management Index	7,537	1.11	0.16	0.6	1.78
Rainfall (mm)	7,537	287	89	119	649
Topsoil Moisture Rating Short	1,679	22%			
Topsoil Moisture Rating Surplus	5,858	78%			
Herbicide Applied	7,488	99%			
Fungicide Applied	2,108	28%			

Insecticide Applied	149	2%			
Seed Treatment Applied	515	7%			
Cereal Previous Crop t-1	2,855	38%			
Oilseed Previous Crop t-1	2,930	39%			
Pulse Previous Crop t-1	1,752	23%			
Soil Group 1	2,749	36%			
Soil Group 2	2,393	32%			
Soil Group 3	2,395	32%			

Source: (SCIC, 2019).

4.7 Summary

This chapter discussed the various sources of data used for analysis including the representativeness of the selected sample and the trends and descriptions of each of the variables included in the model. Lastly, a description of the summary statistics for the selected sample of data was included.

Chapter 5: Model Estimation and Results

5.1 Introduction

This chapter discusses the model estimation process and regression results. The estimation process discussion highlights the choice of the functional form for estimation, the econometric approach used, and the specification of the final model. Additionally, the estimated wheat production function coefficients are presented in the regression results section with a detailed discussion of each of the model components.

5.2 Production Function Estimation

5.2.1 Functional Form

The production function, described earlier in equation 3.1, provides the foundation for the estimation of the economically optimal nitrogen rate for spring-wheat production. Given the importance of the production function for this analysis, the choice of functional form is crucial to accurately capture spring-wheat production in west-central Saskatchewan and provide unbiased parameter estimates. Given that there are multiple true functional forms representing a relationship and many are largely unknown, Griffin, Montgomery, & Rister (1987) establish a criteria for selecting a representative functional form for econometric analysis, including (1) consideration regarding the maintained hypotheses, (2) parameter estimation constraints including data availability and estimation procedures, (3) goodness-of-fit and data conformity, and (4) the application of the results. This section describes the process of selecting a functional form for the estimation of the production function.

Many functional forms have appeared in the literature and are commonly based on agronomic or theoretical considerations. Polynomial functions, usually of the second degree, are often used because they are smooth, differentiable functions that allow for the diminishing marginal return of inputs which is often observed (Dhakal & Lange, 2021). Quadratic functions are considered a good first order approximation when aggregating across space, even when there are nonlinear responses within the field. This is especially the case when there exists significant heterogeneity across a field. Criticisms of the quadratic function include that estimated yields and corresponding optimal input recommendations are often overestimated, and these functions imply that crop yields fall dramatically at excessive fertilizer rates. Additionally, these functions fail to consider the agronomic concept of Liebig's law of the minimum where plant growth is dictated by the most limiting input and not the total amount of inputs available.

The second functional form found in response literature is linear- quadratic plateau functions. These functions are largely based on Liebig's law of the minimum where plants require nutrients and other inputs in fixed ratios. Yield is impacted linearly with the addition of an input until it is no longer limiting relative to other inputs, where the corresponding yield is then plateaued. For yield to increase beyond this plateau point, other inputs, which are now considered limiting must be increased. Nonuniformities across the field will cause plants in different areas to be limited by different inputs at different amounts. While a plateau function makes agronomic sense, empirically, only one input is considered as the limiting input. In the context of experimental plot data where inputs, soil, and environmental conditions are made as homogenous as possible to analyze yield responses at various input levels, plateau functions seem appropriate. However, across heterogenous soil, climate, and input conditions, where different inputs, soil, or environmental conditions may be limiting in different areas, a quadratic function serves as a better approximation. This is further supported by Berck and Helfand (1990) findings that suggest polynomial and plateau functions converge under heterogenous conditions and across space.

Given the discussion above and the criteria described by (Griffin et al., 1987), a modified quadratic production function of the second degree is estimated. The production function is considered "modified" because not all inputs are estimated with a second squared term, and not all possible interactions are estimated. Instead, only those inputs and interaction relationships that were considered conceptually significant are included, described by equations 5.1 and 5.2 below.

5.2.2 Econometric Approach in Estimation

A significant challenge in the estimation of production function parameters is the simultaneity problem, first described by Marschak & Andrews (1944) where output and chosen input quantities are simultaneously chosen. The particular area of concern is the correlation between chosen input levels and any unobserved firm specific shocks in productivity that biased the parameter estimates of the production function (Levinsohn & Petrin, 2003). Any unobserved positive productivity shocks will most likely increase the chosen level of inputs leading to biased estimates of productivity.

To combat the simultaneity problem, researchers have implemented firm specific fixed effects using panel data, focusing specifically on within-firm variation in chosen inputs over time for the estimation of production parameters. A potential area of concern, as pointed out by Griliches & Mairesse (1995) when focusing on within-firm variation in chosen inputs over time to alleviate the simultaneity problem is the fact that many chosen input levels within a firm do not vary much over time. Unique to this study, Table 4.11 above highlights that the chosen input of focus for this study, nitrogen, showcases very little variation within producers (firms) confirming this concern.

Realizing the legitimate concern raised by Griliches & Mairesse (1995), Olley and Pakes (1996) introduced a novel estimation method to address the simultaneity problem and the lack of within firm variation. Olley and Pakes (1996) assume that firms choose variable inputs such as labor and a level of investment or investment proxy, which is in itself a function of the firms capital and productivity shock. Positive investments signal positive productivity shocks without the constraints of focusing on within firms over time (Kim, Petrin, & Song, 2016).

A potential area of criticism for the Olley Pakes method is that the level of investment proxy must be positive (Levinsohn & Petrin, 2003). They instead suggest the use of intermediate input proxies such as electricity as they are likely always positive and provide the same benefits and more as the Olley Pakes method (Levinsohn & Petrin, 2003). Levisohn and Petrin (2003) compare the estimates of OLS, fixed effects, the Olley Pakes method, a Blundell-Bond GMM estimator (lagged input instrumental variable (IV) estimator with fixed effects), and their intermediate input proxy method. They find that their method of estimation of the production function provides different results than those compared but builds off the Olley Pakes method in addressing and solving the simultaneity problem (Levinsohn & Petrin, 2003).

Due to a lack of sufficient and available data, the Olley Pakes or investment proxy method or the Levisohn and Petrin or intermediate input proxy method was not employed in the estimation of the wheat production function in this study. Additionally, the use of fixed effects at the producer or firm level was considered but not used in the final model due to the concern discussed above regarding a significant lack of variation within producers. A more detailed discussion of this takes places in section 5.2.4.3. While it is assumed that producers do not vary nitrogen inputs after seeding when additional shocks may occur, producers may respond to unobserved positive productivity shocks from one season to the next. Therefore, the estimated results that are discussed later on in this chapter are assumed not to be free from bias and do not fully address the simultaneity problem.

5.2.3 The Data Analysis Process

Data analysis and model estimation was performed using RStudio statistical software version 1.3.1093 (Rstudio Team, 2022). The OLS regression model was estimated using the *fixest* package from Berge et al., 2022. The economic optimal nitrogen rates (EONR) were estimated using the *deltamethod* function from the *marginaleffects* package from Arel-Bundock, 2022.

5.2.4 The Model

This section includes the specification of the model, a detailed discussion of the inclusion of interaction terms and fixed effects, the correlation between the included continuous variables, the choice of standard errors, and the model residuals. Following this section, the regression results are presented and discussed.

5.2.4.1 Model Specification

The estimated production function from chapter 3 is:

$$y_{it} = f(v_{it}, m_{it}, s_{it}) \tag{5.1}$$

where v_{it} is a vector comprised of variable inputs including applied fertilizers and crop protection products. Vector m_{it} is comprised of management factors including chosen crop rotation, seeding date, and the variety of spring wheat chosen to grow. Lastly, vector $s_{it}$is comprised of the physical conditions of spring wheat field i at time t including beginning topsoil moisture conditions, accumulated rainfall, and soil productivity. Table 4.2 describes the full list of variables included in the estimated production function below.

The production function described by equation 5.1 above is modeled and estimated by a modified quadratic production function, described by equation 5.2 below:

$$\begin{aligned} y_{it} = {} & \beta_1 N_{it} + \beta_2 N_{it}^2 + \beta_3 P_{it} + \beta_4 K_{it} + \beta_5 S_{it} + \beta_6 \mathit{SeedingDate}_{it} + \beta_7 \mathit{SeedingDate}_{it}^2 \\ & + \beta_8 \mathit{ManagementIndex}_{it} + \beta_9 \mathit{VarietyIndex}_{it} + \beta_{10} \mathit{Rainfall}_{it} \\ & + \beta_{11} \mathit{Rainfall}_{it}^2 + \beta_{12} \mathit{Fungicide}_{it} + \beta_{13} \mathit{ShortTopsoilMoisture}_{it} + \beta_{14} N_{it} \\ & * \mathit{ShortTopsoilMoisture}_{it} + \beta_{15} \mathit{Pulse}(t-1)_{it} + \beta_{16} \mathit{Oilseed}(t-1)_{it} \\ & + \beta_{17} N_{it} * \mathit{Pulse}(t-1)_{it} + \beta_{18} N_{it} * \mathit{Oilseed}(t-1)_{it} + \beta_{19} \mathit{SoilGroup}2_i \\ & + \beta_{20} \mathit{SoilGroup}3_i + \beta_{21} N_{it} * \mathit{SoilGroup}2_i + \beta_{22} N_{it} * \mathit{SoilGroup}3_i + \beta_{23} \tau_t \\ & + \varepsilon_{it} \qquad 5.2 \end{aligned}$$

where nitrogen (N), phosphorous, potassium, sulphur, seeding date, variety index, fungicide use, previous crop, and soil grouping all vary by field. Topsoil moisture rating and in-season rainfall also vary by field but are reported at the rural municipality (RM) level. Additionally, the squared terms for nitrogen, seeding date, and rainfall were included. Year fixed effects, represented by tau (τ) in equation 5.2 above, are included to control for variation across time that is not already captured in the model. Fixed effects for producer and soil productivity rating were considered for the model but not included due to a lack of variation in management within producers, and the benefit of including a nitrogen by soil grouping interaction term over just a soil productivity rating fixed effect. A more detailed discussion regarding these alternative models and the exclusion of these fixed effects takes place in section 5.2.4.3.

The variables that make up the production function are based on existing literature, both from previous yield gap studies, yield response studies, and fertilizer-use profitability research, as well as underlying economic and agronomic theory. The inclusion of squared terms for nitrogen, seeding date, and rainfall were also based on previous literature and agronomic theory. Intuitively, too little or too much rainfall negatively impacts crop yields and quality. Similarly, early and late seeding dates expose the crop to different growing conditions that impact yield such as frost or dry conditions later in the growing season. These particular variables are expected to have a positive but diminishing marginal impact on yield. As previously mentioned in section 5.2.1, the estimated production function is considered “modified” because not all inputs are estimated with an included squared term.

5.2.4.2 Included Interaction Terms

Several interaction terms were tested based on existing literature and underlying agronomic theory. Given that the wheat yield response to nitrogen was of particular interest for this study, nitrogen interacted with several input and management decisions were tested for their statistical significance at p value < 0.05.

Nitrogen was interacted with phosphorous, potassium, and sulphur, however, no statistically significant relationships were found. Additionally, nitrogen was interacted with seeding date, variety index, management index, fungicide application, and rainfall, however, no statistically significant relationships were found. This suggests that the wheat yield response to applied nitrogen is largely independent of the other included inputs and management characteristics.

Additionally, nitrogen was interacted with soil productivity grouping, which was statistically significant for soil group 3, the least productive soils. A Wald Test was performed using the *waldtest* function from the *lmtest* package in Rstudio from Hothorn et al., (2022), to determine if the nitrogen by soil grouping interaction term should remain in the model. The null hypothesis that the nitrogen by soil grouping interaction term should not be included in the model was rejected given the p value of less than 0.05 (p value = 0.0007741).

Nitrogen was also interacted with the fields previous crop (cereal, oilseed, or pulse in t-1). While the interaction term was only significant at 0.1 (p value < 0.1), a Wald Test was performed to determine if the nitrogen by previous crop interaction term should remain in the model. The null hypothesis that the nitrogen by previous crop interaction term should not be included in the model was rejected given the p value of less than 0.05 (p value = 0.003397).

Lastly, nitrogen was interacted with topsoil moisture rating (short or adequate). The interaction term was not statistically significant. A Wald Test was performed to determine if the nitrogen by topsoil moisture rating interaction term should remain in the model. The null hypothesis that the nitrogen by topsoil moisture rating interaction term should not be included in the model could not be rejected given the p value of 0.75. Given agronomic theory and the conceptual importance of topsoil moisture prior to planting for making fertilizer decisions, a dummy variable for topsoil moisture rating remained in the model to control for differences in beginning soil moisture.

Although the wheat yield response to applied nitrogen was of interest for this study, fungicide and rainfall were also interacted to test the statistical significance of the relationship. However, the interaction was not significant. Rainfall, rainfall^2, and fungicide use were all statistically significant and had coefficient signs that were consistent with agronomic theory.

Table 5.1: Wald Test results for Nitrogen by Soil Grouping, Previous Crop, and Topsoil Moisture Rating Interaction. [8]

Interaction	**F Statistic**	**H_0: Do not include in the model**
Nitrogen: Soil Grouping	7.1706	Reject the null (p value < 0.05)
Nitrogen: Previous Crop t-1	5.6892	Reject the null (p value < 0.05)
Nitrogen: Topsoil Moisture Rating	0.0991	Do not Reject the null (p value > 0.05)

Source: Author's calculations.

5.2.4.3 Included and Considered Fixed Effects

Fixed effects for year, soil productivity rating, and producer ID were considered for the model to control for unobservable factors that impact crop yields such as those that vary over time, differing soil characteristics and qualities, and management abilities. Wald tests were performed to determine if the fixed effects for year, soil productivity rating, and producer ID should be included in the model. The null hypotheses that year, soil productivity rating, and producer ID should not be included in the model as fixed effects were rejected (p values < 2.2e-16).

Year fixed effects were included in the model for 2011 to 2019 to account for factors that impact all producers but vary over time. Additionally, an alternative model that does not include year fixed effects is included in Appendix A.

Unobserved heterogeneity in different soil productivity ratings were also considered to be controlled for by the use of fixed effects for ratings A to P. However, the conditioning of the yield response to nitrogen was considered to be of interest leading to the development of the soil groups 1, 2, and 3 and the interaction terms discussed in section 5.2.4.2 above. An alternative model with the inclusion of soil productivity rating fixed effects is shown in Appendix A.

---

[8] Note: The Wald Test results were estimated using the *waldtest* function from the *lmtest* package in Rstudio from Hothorn et al., 2022.

As discussed in sections 4.5 and 5.2.2, while a producer ID fixed effect given available data would help to partially alleviate concerns regarding the simultaneity problem, there is the concern of a significant lack of variation in nitrogen use at the producer level. Producer ID fixed effects would also help to account for unobservable differences across producers such as machinery, management, and labor as examples. Given this lack of variation, a producer ID fixed effect is not included in the final model but is shown in Appendix A below. Instead, to account for unobserved differences across producers, a management index proxy is implemented. The management index, as discussed in section 4.4.5, is used by SCIC to differentiate producers yield experiences. Additionally, to account for pre-seeding conditions that may influence fertility decisions at the start of each growing season or production cycle, a topsoil moisture rating is included.

5.2.4.4 Continuous Variable Correlation

To test for the presence of multicollinearity among the included variables, a correlation matrix is generated. The correlation of the continuous variables included in the model are presented in Table 5.1 below. Yield is most correlated with variety index, management index, and nitrogen. There does not appear to be any highly correlated variables, therefore multicollinearity is not considered to be a problem in the model.

Table 5.1: Correlation of continuous variables included in the model. [9]

	Yield	N	P	K	S	Seed Date	Mng. Index	Variety Index	Rainfall
Yield	1.00	0.31	0.19	0.18	0.12	-0.05	0.38	0.42	-0.13
N	0.31	1.00	0.45	0.09	0.12	-0.08	0.28	0.25	-0.03
P	0.19	0.45	1.00	0.04	0.06	-0.02	0.16	0.22	-0.04
K	0.18	0.09	0.04	1.00	0.26	-0.12	0.19	0.17	-0.08
S	0.12	0.12	0.06	0.26	1.00	-0.12	0.12	0.14	-0.07
Seed Date	-0.05	-0.08	-0.02	-0.12	-0.12	1.00	-0.0007	-0.14	0.09

---

[9] Note: Correlation values calculated using the *cor* function from the *stats* package in Rstudio from the R Core Team, 2022.

Mng. Index	0.38	0.28	0.16	0.19	0.12	-0.0007	1.00	0.20	0.09
Variety Index	0.42	0.25	0.22	0.17	0.14	-0.14	0.20	1.00	-0.16
Rainfall	-0.13	-0.03	-0.04	-0.08	-0.07	0.09	0.09	-0.16	1.00

Source: Author's calculations based on compiled dataset.

5.2.4.5 Homoskedasticity

A classic assumption of OLS regressions is that the errors of the estimated coefficients have constant variance, known as homoskedasticity (Wooldridge, 2010). When this assumption is not met, models exhibit heteroskedasticity. This is a common occurrence in time-series data, where estimates are still unbiased but considered inefficient. To test for heteroskedasticity, a Breusch-Pagan test is employed. The resulting test statistic is 218.8 and the corresponding p-value is significant (< 2.2e-16) therefore, the null hypothesis that homoskedasticity is present should be rejected. Therefore, heteroskedasticity is statistically present in the model.

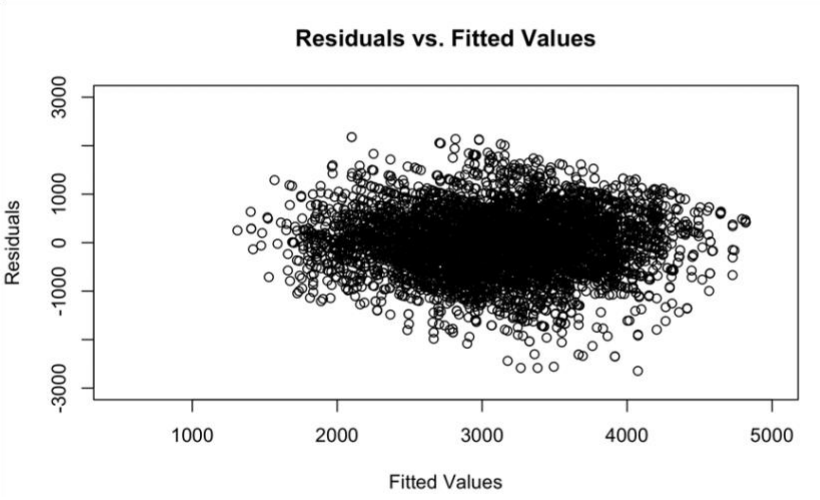
The use of robust standard errors is a common solution to heteroskedasticity, but clustered standard errors are also considered as a solution. The common motivation for clustering is that unobserved components in outcomes for units within clusters are correlated (Abadie, Athey, Imbens, & Wooldridge, 2023). However, it is often unclear when to use clustered standard errors. Abadie et al. (2023) outline conditions for when standard errors should be clustered when using fixed effects including if there is both clustering in the assignment and there is heterogeneity in the treatment effects, or if there is both clustering in the sampling and there is heterogeneity in the treatment effects.

In the case of the model used for this analysis, it is expected that there is unexplained variation across producers indicating that there is heterogeneity in the producer effect. Additionally, it is expected that there is clustering in the sampling of producers used in the dataset, further warranting the use of clustered standard errors at the producer ID level, as suggested by the criteria set out by Abadie et al. (2023).

5.2.4.6 Model Residuals

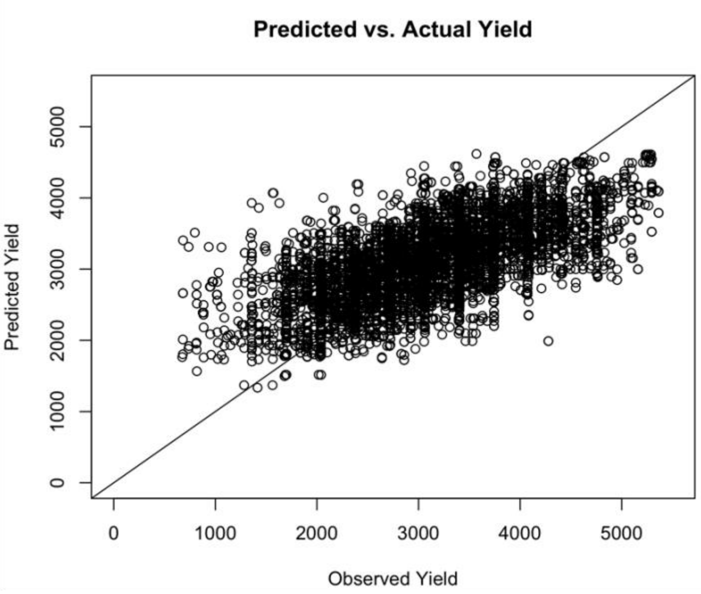
To assess the appropriateness of the model's functional form, Figure 5.1 shows a scatterplot of the model's residuals versus fitted values. Figure 5.2 below shows the models predicted yields versus observed yields. As mentioned in the Section 5.2.4.5, heteroskedasticity is present in the model and is somewhat visible in Figure 5.1 below. Additionally, it appears from Figure 5.2 below that the model slightly over-predicts yields at the lower end of those that are observed while under-predicting yields at the higher end of those that are observed.

Figure 5.1: Model residuals vs. fitted values.



Source: Author's calculations.

Figure 5.2: Model predicted yields vs. observed yields.



Source: Author's calculations.

5.3 Regression Results

This section discusses the results of the estimated production function, presented in Table 5.2 below. A more detailed discussion of the regression results takes place in the next section, 5.4.

Table 5.2: Production function estimation results.

Dependent Variable	*Yield*
Independent Variables	Estimated Coefficient
Nitrogen	13.75 *** (2.90)
Nitrogen^2	-0.062245 *** (0.02)
Phosphorous	0.36 (1.99)
Potassium	10.16 * (4.02)
Sulphur	-0.17 (2.72)
Seeding Date	1.92 (2.79)
Seeding Date^2	-0.22 * (0.13)
Management Index	1,199 *** (139.6)
Variety Yield Index	1,740 *** (159.4)
Topsoil Moisture Short (=1)	57.98 (69.71)
Soil Group 2 (=1)	-85.17 (72.27)
Soil Group 3 (=1)	-144.89 * (87.79)
Rainfall	6.42 *** (1.34)
Rainfall^2	-0.01 *** (0.002)
Fungicide (=1)	295 *** (38.7)
Previous Crop Oilseed (=1)	-124.9 (97.2)
Previous Crop Pulse (=1)	-41.2 (83.2)

Nitrogen × Previous Crop Oilseed	2.15 (1.57)
Nitrogen × Previous Crop Pulse	2.26 * (1.35)
Nitrogen × Soil Group 2	-1.39 (1.14)
Nitrogen × Soil Group 3	-2.99 ** (1.34)
Observations	7,537
Adj. R^2	0.451
Within R^2	0.351
Standard Errors	Clustered (Producer ID)
Fixed Effects	Year
Significance Codes	***: 0.01, **: 0.05, *: 0.1

Source: Author's calculations.

5.4 Discussion of Regression Results

5.4.1 Nitrogen, Phosphorous, Potassium, and Sulphur

Nitrogen, which accounts for the effect of nitrogen applied on soil group 1 and previous cereal crop fields is significant and of the expected sign. Additionally, the squared term for nitrogen is also significant and of the expected sign. The other included fertilizer inputs: phosphorous and sulphur produce non-significant coefficients, while the applied potassium coefficient is significant and of a similar magnitude to that of nitrogen. However, potassium is less commonly applied and at a far lower rate on average than that of applied nitrogen. The model results suggest that applying 1 kilogram of actual potassium per hectare increases yield by roughly 10 kilograms per hectare, holding all other inputs constant. The average observed rate of potassium is nearly 2 kilograms per hectare, estimated to increase spring-wheat yields by approximately 20 kilograms per hectare when applied. Phosphorous and sulphur do not exhibit statistically significant coefficients and are small in magnitude, especially when considering the agronomic importance of phosphorous in plant growth. Additionally, while the estimated coefficient for applied sulphur is not significant, it is also unexpectedly negative. It is unclear as to why, as it was expected that all fertilizer applied would have a positive impact on yield. As mentioned in the previous chapter, any applied phosphorous, potassium, and sulphur is assumed to add to the soils total reserve and not necessarily used by the plant in the given year, therefore a diminishing marginal return for these fertilizer inputs was not considered.

5.4.2 Nitrogen Interactions

The response of nitrogen is included as conditional on the fields specific soil grouping and previous crop. The individual nitrogen term represents the yield response to applied nitrogen for fields that fall into soil group 1 (highest productivity) and have previously had a cereal crop grown on them. Focusing on the nitrogen by soil group interaction, the yield response is greatest on soil group 1, which is expected given that this soil is considered to be the most productive. The response to nitrogen on soil group 3 is almost 50% as responsive as soil group 2.

When the yield response to nitrogen is conditioned by the fields previous crop, the estimated coefficients are statistically significant. The a priori expectation was that the yield response to applied nitrogen would be greatest following an oilseed or cereal crop given that these crops typically demand more nitrogen for growth and do not naturally fixate nitrogen like that of pulse crops. Given that residual fertilizer is not accounted for in the model, the estimated coefficients for fertilizer inputs reflect the yield response to only applied fertilizer. If residual nitrogen was present in the soil after a pulse crop, conventional wisdom would suggest that the estimated response to applied nitrogen would be less responsive than for fields with less residual nitrogen. Instead, the estimated yield response to nitrogen, conditioned by the fields previous crop shows that pulse crops have the greatest response to applied nitrogen, followed by oilseed stubble[10]. This may be the case that the previous crop variable is also picking up non-nitrogen related benefits from previous pulse crops.

While not expected, the benefits that pulse crops have on subsequent crops have been documented in agronomic research. Arcand, Lemke, Farrell, & Knight (2014), find that decomposing pulse crop materials can add to the supply of residual nitrogen. Miller, Gan, McConkey, & McDonald (2003), find additional soil moisture available to subsequent crops following pulses. Additionally, pulse crops have the non-nitrogen related benefit to subsequent crops of breaks in pest cycles (Stevenson & Van Kessel, 1996). It appears non-nitrogen related benefits are being captured in the estimated results, the estimated marginal product of applied nitrogen, and the determined economically optimal nitrogen rates for spring-wheat production following pulse crops.

---

[10] Stubble refers to what was grown on the field in the previous year.

5.4.3 Fungicide Use

Fungicide, which was the only applied chemical included in the final model, was significant and positive, as expected. The choice to apply fungicide increased yields by almost 300 kilograms per hectare, holding all other inputs constant. Similar benefits to using fungicide on wheat have been reported in winter wheat in Kansas, with yield increases on wheat fields that had a fungicide application (Cruppe et al., 2021). The decision to apply fungicide to cereal crops is largely dependent on environmental conditions, especially during the time of flowering. While a yield benefit is observed here, quality is also an important factor when focusing on the return of fungicide use which is not included in our analysis.

5.4.4 Rainfall and Topsoil Moisture

Focusing on the included climate characteristic, rainfall is both positive and significant and exhibits a small diminishing marginal return on wheat yield. This result is consistent with the a priori expectation that as rainfall accumulates beyond a certain threshold, increased disease and lodging pressure begins to reduce the overall yield and quality of the crop. From the estimated results, it appears that the yield impact of rainfall is largest around 350 mm. Additionally, the estimated yield gap between the minimum observed rainfall and the average observed rainfall is nearly 500 kilograms per hectare, holding all other inputs constant. This result suggests that variation in received rainfall, even just between the minimum and average amounts observed, has a significant impact on the yield gap for spring-wheat in Saskatchewan. While it was hypothesized that rainfall may also influence the yield response to nitrogen, the resulting interaction term was not significant and therefore not included in the final model. Rainfall, measured in millimeters, is the in-season total amount of rain at the rural municipality (RM) level. While this variable fails to capture the impact that rainfall timing has on yield, it does capture the impact that variation in rainfall has on yields. Topsoil moisture, which is also reported at the RM level is of the expected sign but is not significant. It was included in the final model based on its conceptual significance as an indicator of pre-seeding moisture conditions that may impact fertility and other management decisions.

5.4.5 Seeding Date

Several of the included management characteristics exhibit the expected coefficient signs. Seeding date, which is measured as days before or after May 14th of each year that seeding took place is positive and shows a significant diminishing response meaning that delayed seeding has a slight yield penalty. Research conducted at the Indian Head Agriculture Research Foundation (IHARF) also found a minimal spring-wheat yield penalty for delayed seeding date on black, dark brown, and brown soil zones beyond the middle of May (Catellier, 2022).

5.4.6 Variety Index

The yield index of the spring-wheat variety grown relative to AC Carberry (check variety) was significant and positive. Given that it is measured as an index, the coefficient which is 1,740 should be interpreted as a yield increase of approximately 17 kilograms per hectare or about 0.6% of the average observed yield, for each 1% increase in variety yield-index holding all other inputs constant. The impact of the chosen variety on yield is quite significant given the range of varieties observed being used. Holding all other inputs constant, moving from the lowest yielding variety to the highest had an estimated increase in yield of almost 1,200 kilograms per hectare. The impact of the variety chosen for production may explain a sizeable portion of the yield gap for spring-wheat production in Saskatchewan.

5.4.7 Management Index

To control for differences at the farm and manager level, a management index variable described in Chapter 5, is included in the model. The estimated management index coefficient is significant, indicating that it is controlling for select differences across farms or farm managers. Given that is measured as an index, the coefficient of 1,199 should be interpreted as a yield increase of roughly 12 kilograms per hectare or roughly 0.4% of the average observed yield, for each 1% increase in management index level, holding all other inputs constant. The observed management index variable ranges from 0.6 up to 1.8. While the impact of increasing the management index by 1% is small, there is a significant gap between the lowest and highest management indices of nearly 1,500 kilograms per hectare. Additionally, the estimated model controls for several input and management characteristics already, meaning that there are still significant differences in farms or farm managers potentially contributing to the yield gap.

Chapter 6: The Economically Optimal Nitrogen Rate

6.1 Introduction

This chapter discusses calculating the economically optimal nitrogen rate (EONR) from the estimated production function results in the previous chapter. Annual average cost-to-price scenarios as well as the mean cost-to-price ratio are presented to capture the potential different prices and costs individual producers may face. Additionally, the theoretical yield maximizing nitrogen rate is discussed as well as the pure-nitrogen related change in profitability moving from observed average nitrogen rates to the estimated optimal nitrogen rate.

6.2 Calculating the Economically Optimal Nitrogen Rate (EONR)

Following estimation of the wheat production function in section 5.3 of the previous chapter, the economically optimal nitrogen rate (EONR) can be determined from the estimated marginal product of nitrogen. As discussed in section 3.2.6, the EONR is dependent on the estimated marginal product (MP_N) of nitrogen multiplied by the expected output price of spring-wheat, equalling the input price of nitrogen. Equation 3.8 simplifies this optimal condition where at the EONR, the marginal product of nitrogen (MP_N) is equal to the ratio of the unit cost for nitrogen to the expected unit price for spring-wheat.

The price producers pay for nitrogen plays an important role in the calculation of the EONR. While there are several costs associated with nitrogen use including purchasing, transportation, and storage, this analysis only considers the purchase price. Additionally, producers may face different nitrogen prices depending on when during the year they purchase. This analysis assumes that the nitrogen price will be the same for all producers in a given year. A more detailed discussion of the nitrogen price data can be found in section 4.2.2.

The price producers expect to receive for spring-wheat is also important in determining the EONR. Producers in Saskatchewan have access to several options when selling grain including forward contracting and spot pricing. For this analysis, it is assumed that all producers expect the same price for spring-wheat in a given year, and specific to wheat there are no protein premiums or discounts applied. However, it is noted that nitrogen does contribute both to yield and protein levels in wheat. Additionally, protein premiums or discounts do impact the price that producers may receive depending on the spreads each year. A more detailed discussion of the expected wheat price data can be found in section 4.2.2. Additionally, to account for individual producers paying and receiving different prices, the minimum, mean, and maximum cost-price ratios are presented in Table 6.2.

Focusing specifically on the economically optimal nitrogen rate, a risk neutral producer will optimize nitrogen usage where the value of the marginal product for nitrogen is equal to the unit cost for nitrogen. The estimated marginal product of nitrogen is conditional on both the previous crop and soil grouping. Using the estimated regressions coefficients and the decision rule developed in Chapter 3, equation 3.5 can be altered to determine the optimal nitrogen rate, $N_{it}*$

$$N_{it}^{*} = \frac{\left(w_{N_t}/p_{y_t}\right) - \beta_1 - \beta_{15} * prevpulse_{it} - \beta_{16} * prevoilseed_{it} - \beta_{17} * soilg3_i - \beta_{18} * soilg3_i}{2 * \beta_2} \qquad (6.1)$$

where the economically optimal nitrogen rate is dependent on the cost-price ratio and the fields previous crop and soil grouping.

6.2.1 The Economically Optimal Nitrogen Rate for Annual Cost-Price Ratios

While the marginal product calculation for nitrogen does not account for any dynamic factors beyond the fields previous crop and soil grouping, the historical annual cost-price ratios for 2011 to 2019 and the estimated EONR's are presented in Table 6.1 below. The estimated EONR's are separated by the fields previous crop. The highest estimated EONR's was on previous pulse crop fields.

Table 6.1: Annual Cost-Price Ratio and Estimated EONR by Soil Group, Previous Crop, 2011-2019.[11]

Year	**CWRS Price ($/kg)**	**N Cost ($/kg)**	**Cost-Price Ratio**	**Soil Group**	**EONR (kg ha^{-1})**	**Std. Error**	**Obs. Mean N Rate (kg ha^{-1})**
Cereal Previous Crop							
2011	$ 0.26	$ 1.31	5.04	*1*	70	12.4	51
				2	59	12.5	65
				3	46	13.7	54
2012	$ 0.24	$ 1.55	6.36	*1*	59	12.5	52

---

[11] Note: *, **, *** indicate significantly different observed mean nitrogen rates relative to the estimated EONR at the 0.10, 0.05, and 0.01 level.

				2	48	13.4	56
				3	35	15.3	55
2013	$ 0.29	$ 1.50	5.15	*1*	69	12.4	50
				2	58	12.5	54
				3	45	13.8	59
2014	$ 0.23	$ 1.40	6.09	*1*	62	12.4	52
				2	50	13.1	53
				3	38	14.9	55
2015	$ 0.24	$ 1.41	5.83	*1*	64	12.3	57
				2	53	12.9	64
				3	40	14.6	59
2016	$ 0.26	$ 1.27	4.95	*1*	71	12.4	60
				2	60	12.5	57
				3	47	13.6	66
2017	$ 0.23	$ 1.11	4.8	*1*	72	12.5	62
				2	61	12.4	62
				3	48	13.4	63
2018	$ 0.25	$ 1.06	4.21	*1*	77	12.8	63
				2	65	12.3	77
				3	53	12.9	68
2019	$ 0.25	$ 1.21	4.87	*1*	71	12.5	67
				2	60	12.4	63
				3	47	13.5	65
Year	**CWRS Price ($/kg)**	**N Cost ($/kg)**	**Cost-Price Ratio**	**Soil Group**	**EONR (kg ha^{-1})**	**Std. Error**	**Obs. Mean N Rate (kg ha^{-1})**
Oilseed Previous Crop							
2011	$ 0.26	$ 1.31	5.04	*1*	87	13.9	76
				2	76	12.5	71
				3	63	12.3	72
2012	$ 0.24	$ 1.55	6.36	*1*	77	12.8	58
				2	66	12.3	60
				3	53	12.9	65
2013	$ 0.29	$ 1.50	5.15	*1*	86	13.8	61*
				2	75	12.7	64
				3	62	12.4	65
2014	$ 0.23	$ 1.40	6.09	*1*	79	12.9	62
				2	68	12.3	72
				3	55	12.7	66
2015	$ 0.24	$ 1.41	5.83	*1*	81	13.2	62
				2	70	12.4	64
				3	57	12.6	66
2016	$ 0.26	$ 1.27	4.95	*1*	88	14.1	57**
				2	77	12.8	62
				3	64	12.3	74

2017	$ 0.23	$ 1.11	4.8	*1*	89	14.3	68
				2	78	12.9	69
				3	65	12.3	67
2018	$ 0.25	$ 1.06	4.21	*1*	94	15	67*
				2	83	13.4	67
				3	70	12.4	67
2019	$ 0.25	$ 1.21	4.87	*1*	89	14.2	68
				2	78	12.9	66
				3	65	12.3	68
Year	**CWRS Price ($/kg)**	**N Cost ($/kg)**	**Cost-Price Ratio**	**Soil Group**	**EONR (kg ha^{-1})**	**Std. Error**	**Obs. Mean N Rate (kg ha^{-1})**
				Pulse Previous Crop			
2011	$ 0.26	$ 1.31	5.04	*1*	88	14.1	63*
				2	77	12.8	64
				3	64	12.3	67
2012	$ 0.24	$ 1.55	6.36	*1*	78	12.9	51**
				2	66	12.3	52
				3	54	12.8	48
2013	$ 0.29	$ 1.50	5.15	*1*	87	13.9	60*
				2	76	12.7	62
				3	63	12.3	53
2014	$ 0.23	$ 1.40	6.09	*1*	80	13.1	58
				2	69	12.4	55
				3	56	12.7	62
2015	$ 0.24	$ 1.41	5.83	*1*	82	13.3	49**
				2	71	12.4	57
				3	58	12.5	53
2016	$ 0.26	$ 1.27	4.95	*1*	89	14.2	59**
				2	78	12.9	65
				3	65	12.3	62
2017	$ 0.23	$ 1.11	4.8	*1*	90	14.4	57**
				2	79	12.9	64
				3	66	12.3	73
2018	$ 0.25	$ 1.06	4.21	*1*	95	15.2	68*
				2	84	13.5	68
				3	71	12.4	67
2019	$ 0.25	$ 1.21	4.87	*1*	90	14.3	71
				2	78	12.9	67
				3	71	12.4	63

Source: Author's calculations.

As evident in Table 6.1 above and in the discussion of data trends in Chapter 4, from 2011 to 2019 observed average nitrogen rates appear to be increasing. Table 6.1 shows that average annual nitrogen rates are steadily increasing while the annual cost-price ratio fluctuates. As previously discussed, producers may be applying more fertilizer over time to respond to the increased yield potential of newer varieties, or the increased rates recommended by extension agronomists. In some years, higher rates may also be explained by increased optimism about the crop year due to optimal conditions early on in the growing season.

Unsurprisingly, the EONR is estimated to be lowest on soil group 3, the least productive soils as classified by SCIC. It is interesting to note however, the observed average annual nitrogen rate in several years is highest on the least productive soil group 3. This may be indicative of producers responding to certain characteristics or the perception of certain characteristics of the less productive soils such as less retention or availability of soil nutrients. In many years, the average nitrogen rate on soil group 3 is at or above the estimated EONR and for previous cereal crop fields. The estimated EONR is highest on previous pulse crop fields and only slightly higher than previous oilseed crop fields, while the cereal previous crop fields had the lowest estimated EONR. In several of the years, the observed average nitrogen rate is lower for previous pulse crop fields especially on the more productive soil groups 1 and 2.

6.2.2 The Economically Optimal Nitrogen Rate for the Mean Cost-Price Ratio

As previously discussed, individual producers may face vastly different prices for both nitrogen and the wheat that is produced and sold. Additionally, the estimated marginal product for nitrogen is conditional on the fields previous crop and soil group and is therefore static in the sense that it does not include any factors that change year to year. To better capture this, Table 6.2 below presents the estimated economically optimal nitrogen rate's (EONR) for the mean annual cost-price ratio from 2011 to 2019. The mean cost-price ratio, calculated as 5.26 most closely resembles the observed cost-price ratio from the year 2013, as shown in Table 6.1 above.

Additionally, Table 6.2 below presents the estimated 90, 95, and 99% confidence intervals for the estimated EONR's. Given that it is also of interest in this study to test if individual producers are applying nitrogen rates that are consistent with what is estimated to be economically optimal, Table 6.2 presents the percentage of observations that reject this null hypothesis for each level of confidence. Using a two-tailed hypothesis test for each level of significance, observations by previous crop and soil group are determined to be below, above, or within the estimated optimal range given the selected level of confidence.

Table 6.2: Mean Cost-Price Ratio and Estimated EONR, % of Observed Suboptimal N rates at 90, 95, and 99% confidence.[12]

Cost-Price Ratio	**Soil Group**	**EONR (kg ha^{-1})**	**Obs. Mean N Rate (kg ha^{-1})**	**# of Obs. for Soil Group and Prev. Crop**	**90% Confidence Interval**			**95% Confidence Interval**			**99% Confidence Interval**		
					90% CI Range (kg ha^{-1})	*% of LLD's Observed Below EONR*	*% of LLD's Observed Above EONR*	*95% CI Range (kg ha^{-1})*	*% of LLD's Observed Below EONR*	*% of LLD's Observed Above EONR*	*99% CI Range (kg ha^{-1})*	*% of LLD's Observed Below EONR*	*% of LLD's Observed Above EONR*
Cereal Previous Crop													
5.26	*1*	68	56	1,077	(48, 89)	38%	9%	(44, 92)	27%	6%	(36, 100)	24%	4%
	2	57	60	903	(36, 78)	20%	21%	(32, 82)	17%	15%	(25, 90)	8%	8%
	3	44	61	875	(21, 67)	3%	37%	(17, 72)	1%	22%	(8, 80)	0%	17%
Oilseed Previous Crop													
5.26	*1*	86	65	930	(63, 108)	51%	3%	(59, 113)	38%	2%	(50, 121)	22%	1%
	2	74	66	989	(54, 95)	30%	9%	(50, 99)	23%	7%	(42, 107)	11%	5%
	3	62	67	1,011	(41, 82)	8%	17%	(37, 86)	6%	14%	(30, 93)	3%	6%
Pulse Previous Crop													
5.26	*1*	86	61	742	(64, 109)	59%	3%	(59, 114)	46%	3%	(51, 122)	30%	2%
	2	75	62	501	(54, 96)	41%	5%	(50, 100)	32%	5%	(43, 108)	14%	3%
	3	62	62	509	(42, 83)	15%	16%	(38, 87)	11%	12%	(31, 94)	6%	8%

Source: Author's calculations.

---

[12] Note: EONR and related confidence intervals are calculated using the *deltaMethod* function from the *car* package in Rstudio, from Fox and Weisberg, 2019. Additionally, the percentage of LLD's observed below or above the estimated EONR for each previous crop and soil group are calculated relative to number of observations by previous crop and soil group found in column 5 of Table 6.2 above.

Table 6.2 above presents the percentage of observations that reject the null hypothesis that producers are applying nitrogen rates that are consistent with what is economically optimal at the mean cost-to-price ratio. Interestingly, even at the 99% confidence level, 30% of the observed nitrogen rates for wheat production following a pulse crop on soil group 1 rejected the null hypothesis and were estimated to be below what was considered optimal. This indicates that a portion of the yield gap for wheat in west-central Saskatchewan could potentially be reduced by producers increasing nitrogen rates towards the EONR when producing wheat on pulse stubble. Additionally, across all confidence intervals, the percentage of observations that rejected the null hypothesis that producers are applying nitrogen rates that are consistent with what is economically optimal increased when moving to less productive soils for nitrogen rates beyond what is optimal. The percentages were highest for the least productive soil group 3. This was especially the case for wheat production on cereal stubble on soil group 3 where 17% of observations rejected the null hypothesis with 99% probability. This was previously noted that the observed average annual nitrogen rate in several years appears highest on the least productive soil group 3. This may be a result of producers increasing nitrogen rates to reach a target yield with increased nutrient demands over the more productive soils. The opposite can be observed for soil group 1, the most productive soils, where the percentage of observations that rejected the null hypothesis are considerably higher for those below what is estimated to be optimal. This was remarked previously that producers appear to be on average under applying on more productive soils and over applying on less productive soils relative to what is estimated to be optimal.

6.3 Yield Maximizing Nitrogen Rate

As discussed in section 3.2.7, the estimation of the yield gap also requires the consideration of the objective of maximizing yield and not necessarily profit which is what this study assumes. Following a similar process as described above in section 6.2 for the estimation of the EONR, using the first order condition (equation 3.4), the maximum condition for the variable input nitrogen can also be determined where output is maximized for the given variable input. The maximum condition for variable input nitrogen is described by equation 3.9, where the maximum condition for nitrogen is where the expected marginal product is equal to zero, where an additional unit of the variable input nitrogen does not produce an additional unit of output.

From the estimated wheat production function results, the yield-maximizing nitrogen rate can be determined by altering equation 6.1 above,

$$N_{it}{}^{M} = \frac{(0) - \beta_1 - \beta_{15} * prevpulse_{it} - \beta_{16} * prevoilseed_{it} - \beta_{17} * soilg2_i - \beta_{18} * soilg3_i}{2 * \beta_2} \quad (6.2)$$

where the marginal product and cost-to-price ratio is set equal to zero. This represents the point where an additional unit of nitrogen does not produce an additional unit of wheat. Table 6.5 below presents the estimated yield maximizing nitrogen rates, holding all other inputs constant.

Table 6.5: Estimated Yield Maximizing Nitrogen Rates by Soil Group and Previous Crop.

Soil Group	**Estimated Yield Maximizing N Rate (kg ha^{-1})**	**Std. Error**	**Obs. Mean N Rate (kg ha^{-1})**
Cereal Previous Crop			
1	110	18.4	56
2	99	16.1	60
3	87	13.9	61
Oilseed Previous Crop			
1	128	22.5	65
2	117	19.8	66
3	104	16.9	67
Pulse Previous Crop			
1	129	22.7	61
2	118	19.9	62
3	104	17.1	62

Source: Author's calculations.

Consistent with the estimated EONR's in Tables 6.1 to 6.4, soil group 1 has the highest estimated yield maximizing nitrogen rates as does the pulse previous crop fields. In some cases, the estimated yield maximizing nitrogen rates are double the observed average nitrogen rates and far exceed the estimated EONR's.

6.4 Change in Pure Nitrogen Producer Surplus

What is also of interest in this study is the change in producer surplus and profitability moving from observed nitrogen rates to estimated economically optimal rates. Figure 6.1 below highlights the change in producer surplus moving from the observed nitrogen rate to the estimated optimal rate. The area under the value of the marginal product (VMP) curve, denoted as the price of wheat (P_Y) multiplied by the marginal product of nitrogen (MP_N) represents the nitrogen related revenue. The area under the cost of nitrogen, denoted by W_N, represents the cost of the applied nitrogen rate.

Under applying nitrogen relative to the optimal rate N^* at N^{OB} forgoes the additional area under the VMP (P_YMP_N) and the net loss to producer surplus is denoted by triangle area a. P_Y represents the average price per kilogram of wheat from 2011 to 2019 of \$0.25. This area is calculated as

$$\Delta\pi_{N_{it}} = 0.5 \cdot 2 \cdot \beta_2 \cdot P_Y(N^* - N^A)^2 \tag{6.3}$$

where the area of a is 0.5 multiplied by the slope of the VMP curve, the deviation from the optimal rate, and the base. Similarly, over-applying nitrogen relative to the optimal nitrogen is represented by the triangle area b. Using the regression results from Table 5.2 above, the slope of the VMP curve is 0.062245. Therefore, the change in nitrogen profitability is calculated as:

$$\Delta\pi_{N_{it}} = 0.5 \cdot 0.12449 \cdot 0.25(N^* - N^A)^2 \tag{6.4}$$

where the change in producer surplus represented by triangle a or b in Figure 6.1 below represents the loss in producer surplus from not optimizing nitrogen rates.

Figure 6.1: Graphical representation of change in producer surplus moving from observed to optimal nitrogen rate.

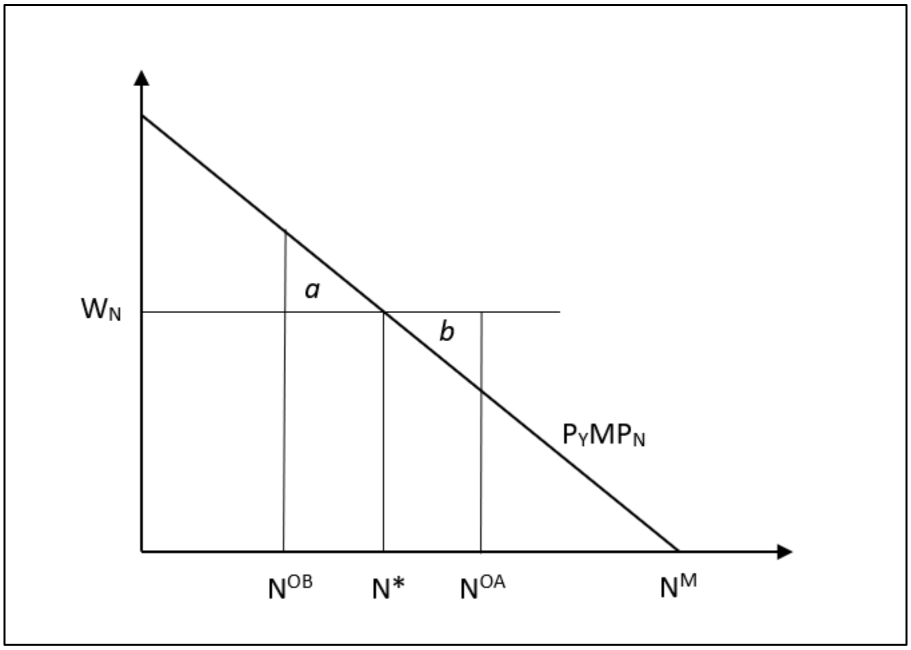


Image Source: Author.

Table 6.6 below highlights the estimated change in strictly nitrogen related profitability moving from actual to optimal nitrogen use. Changes in profitability are calculated at the field level and averaged by previous crop and soil group. The change in nitrogen related profitability ranges from \$0 per hectare at the observed optimal rate to \$195 per hectare. The largest change in profitability occurs on cereal previous crops and on soil group 3 for both cereal and oilseed previous crops. Under these conditions, the observed nitrogen rates exceed the estimated economically optimal nitrogen rates as seen in Table 6.3 above. Overall, the average changes in nitrogen related profitability are small. A similar result was also found by Rajsic and Weersink (2008).

Table 6.6: Change in Pure Nitrogen Profitability moving from Observed to Optimal Nitrogen Rates.

Soil Group	**Mean Change in Producer Surplus by Optimizing N Rate (\$ per ha)**	**Min Change in Producer Surplus by Optimizing N Rate (\$ per ha)**	**Max Change in Producer Surplus by Optimizing N Rate (\$ per ha)**	**Standard Deviation**
		Cereal Previous Crop		
1	10.20	0	121	13.4
2	10.10	0	153	18.8
3	13.40	0	195	25.4
		Oilseed Previous Crop		
1	13.60	0	76	15.7
2	8.14	0	105	11.7
3	6.98	0	140	14.5
		Pulse Previous Crop		
1	16.40	0	85	16
2	9.13	0	81	11.5
3	8.04	0	112	16.2

Source: Author's calculations.

6.5 Summary

This chapter discussed the estimation process of the economically optimal nitrogen rate for various soil groups and previous crop combinations and under various cost-price ratio scenarios including both the annually observed ratios and the minimum, mean, and maximum ratios. While producers are on average increasing nitrogen rates over time, the majority of producers, even under the minimum, mean, and maximum ratio scenarios, were found to be applying suboptimal nitrogen rates. It appears that producers on average are found to be over-applying nitrogen on less productive soils. Additionally, the highest estimated EONR's were found on fields that were previously a pulse crop. Producers appear to on average apply the least amount of nitrogen on fields that were previously cereals, followed by fields that were previously pulse crops. Lastly, this chapter discussed the estimation of the theoretical yield maximizing nitrogen rate for each soil group and previous crop combination, presented in Table 6.5 above. The next chapter will conclude the study and discuss these results in more context.

Chapter 7: Conclusion

7.1 Introduction

This chapter begins with a discussion of the background and main objective of this thesis. Following this, the approach and the major findings are discussed including the broader implications for both spring-wheat production and the yield gap in Saskatchewan. Additionally, the results are discussed in the frame of the current policy landscape in Canada. Lastly, any noted limitations and challenges of the study are discussed.

7.2 Summary

The demand for wheat is expected to increase towards 2050, driven largely by population growth and rising per capita incomes. Additionally, interest towards renewable energy sources such as biofuels has the potential to further increase the demand for cereal grains. Given that cropland expansion is not a realistic solution to increasing crop production, focus has shifted to the sustainable intensification of crop yields on existing cropland to their potential. The difference between actual yields and potential yield is referred to as the *yield gap*. Existing yield gap studies estimate that globally, wheat yields range from 20 to 70% of their potential (R. Fischer et al., 2009; Licker et al., 2010; Lobell et al., 2009; Mueller et al., 2012). One of the most common yield constraints reported is poor nutrient management, specifically that of nitrogen (Beza et al., 2017; Mueller et al., 2012; van Dijk et al., 2017; Waddington et al., 2010). While many current studies have signified that the yield gap for wheat can be reduced by increasing fertilizer usage, they have failed to address the economic component of observed and increased fertilizer rates for producers.

The main objective of this thesis was to determine if an economically exploitable gap for applied synthetic nitrogen fertilizer use exists for spring-wheat production, in Saskatchewan. Early results from Agriculture and Agri-Food Canada (AAFC) suggest that actual wheat yields range from 50 to 80% of their potential in Saskatchewan (Beres, 2022). Focusing specifically within risk zone 16, actual yields are approximately 70% of estimated yield potential (Beres, 2022). Given that actual yields are not expected to reach potential, this research focuses primarily on the economically optimal nitrogen rate (EONR) to help inform how much of the yield gap for spring-wheat in Saskatchewan is economically exploitable.

Based on producer theory, the assumed goal of spring-wheat producers in Saskatchewan is to maximize profits. Under this assumption, a producer will demand an input up to the point where the value of its marginal product is equal to its marginal cost, described by equation 3.5. A wheat yield response function is estimated using a modified quadratic production function, with data from the Saskatchewan Crop Insurance Corporation (SCIC). From the estimated production function, the marginal product of nitrogen is calculated by soil grouping and the fields previous crop. The yield response to nitrogen was determined to be greatest on the most productive soils and when spring-wheat was grown on pulse stubble. Additionally, producers appear to over apply nitrogen on less productive soils and under apply on the most productive soils.

7.3 Major Findings

With respect to the main objective of this thesis, a gap between average observed nitrogen rates and the determined economically optimal rates was observed. Producers on average, were observed underapplying nitrogen on more productive soils and when following a pulse crop, compared to what was determined to be optimal. This indicates that a portion of the yield gap for wheat in risk zone 16 could be reduced by optimizing nitrogen, specifically on soil groups 1 and 2 and when following pulse crops. On the least productive soil group 3, producers on average, were observed overapplying nitrogen compared to what was determined to be optimal.

The estimated response to applied nitrogen was found to vary significantly across soil groupings and by the field’s previous crop. Fields belonging to soil group 1 were among the most responsive to applied nitrogen, which is consistent with the fact that these soils are listed as the most productive by SCIC. Additionally, the yield response to applied nitrogen was greatest when following a pulse crop. It appears that pulse crops provide significant non-nitrogen related benefits to subsequent crops. The average observed nitrogen rates were also greater on less productive soils compared to the most productive soil group 1, signalling that these producers may be responding to the additional agronomic requirements of the field without fully considering the reduced potential of these fields.

Additional findings include the large roles that the chosen wheat variety and management ability of the producer have on the resulting yield. A difference in yield of almost 1,200 kg/ha was estimated between the highest yield-index variety and the lowest, holding all other inputs constant. Similarly, a difference in yield of almost 1,500 kg/ha was estimated between the highest and lowest observed management index measures, holding all other inputs constant. While these findings are not fully explored in this thesis, they are discussed further in the sections below.

7.4 Implications for Spring-Wheat Production in Saskatchewan and the Yield Gap

The results of this thesis provide several important implications for spring-wheat production and the approach of assessing of the yield gap in Saskatchewan. To my knowledge, this is one of the first studies to use large scale field-level yield, input, and management data, from SCIC to estimate the yield response to applied nitrogen for spring-wheat production in Saskatchewan. As previously mentioned, a limiting factor of many existing yield gap studies has been a lack of appropriate data. Additionally, this thesis is one of the few studies to look at how much of the estimated yield gap (completed by AAFC) is economically exploitable for a particular input, in this case nitrogen. As discussed in Chapter 2, exploitable yield is commonly represented as 80% of the estimated yield potential as a proxy, without accounting for any economic components in its calculation. It is however, unlikely that 80% of yield potential represents what is economically exploitable for producers.

While this thesis focuses specifically on the economically exploitable gap between observed and optimal nitrogen use for spring-wheat production in grain risk zone 16, 23 SCIC developed grain risk zones exist in Saskatchewan. Extending the approach of this study to the other grain risk zones as they align with AAFC buffer zones could provide significant insight into the role that both management and economic factors have in the estimated yield gap for spring-wheat production in Saskatchewan. Additionally, analyzing grain risk zones across different soil zones could provide insight into how different factors contribute to the yield gap across Saskatchewan.

The results suggest that the spring-wheat yield response to applied nitrogen varies by the fields specific soil classification and previous crop. Therefore, fertilizer recommendations and use decisions should also vary according to these characteristics. While the response to applied nitrogen was estimated to be greatest on fields classified in soil group 1 and those that had previously been a pulse crop, average nitrogen rates were observed to be higher on less productive soils. Agronomic recommendations should account for the fields unique yield potential and expected response to nitrogen, rather than solely the nutrient requirements to reach a specific yield target. Additionally, the non-nitrogen related benefits of previous pulse crops, as the model suggests, should be considered when addressing fertility recommendations. Given that observed nitrogen rates were highest on the least productive soil group 3, it appears recommendations or use level decisions are responding to the increased nutrient demand of these less productive fields rather than the limited response to applied nitrogen and thus lower optimal rates. Lastly, the results also indicate that a portion of the yield gap for wheat in risk zone 16, estimated by AAFC, could be reduced by optimizing nitrogen, specifically on soil groups 1 and 2 and when wheat production follows pulse crops.

7.4.1 Further Research

The results suggest that the selected wheat variety, the management ability of the producer, and the amount of rainfall received also have a significant impact on spring-wheat yields. Producers in Saskatchewan can select from over 60 different spring-wheat varieties to plant or can choose to replant spring-wheat production from a previous year (Syme et al., 2023). Within risk zone 16, spring wheat varieties with a yield index of 66 to 135% of the check variety were observed being used. While more research is needed, this result suggests that encouraging the use of higher yielding wheat varieties could have a significant impact on the yield gap for spring wheat in Saskatchewan.

Given that several inputs and management characteristics are already accounted for in the model, the management index variable still captures significant differences impacting yields, at the producer level. Such differences may include access to equipment, the education level of the producer, or the timing of various applications throughout the growing season, all contributing to the yield gap for spring-wheat in Saskatchewan. Additional data on unique producer attributes such as the timing of fertilizing or the placement of fertilizer by seeding equipment, gathered through a producer survey would further benefit the yield gap analysis in Saskatchewan. Some of this data may be inferenced through reported statistics on the uptake of the 4R nutrient management strategy of right rate, right time, right place, and right source. However, data regarding 4R adoption is still being improved in terms of regionality of adoption.

Applied potassium (K) also had a large and significant estimated impact on spring-wheat yields but is not largely applied in risk zone 16. While this result suggests that significant yield gains can be realized by increasing potassium use, it is not optimized like nitrogen in this analysis and ultimately, more research is needed to assess the benefits of increasing potassium rates in spring-wheat production in Saskatchewan.

Lastly, while uncontrollable by producers, the amount of rainfall received was estimated to have a significant impact on yields. Therefore, variation in rainfall from year to year and across space within a given year may explain a considerable portion of the yield gap. However, this portion of the yield gap should be considered unexploitable by producers. While rainfall is uncontrollable by producers, more detailed information on the timing of rainfall throughout the growing season may help to better explain the impact it has on the yield responsiveness to other inputs. This could further inform management strategies that are better suited to variable climatic inputs. The application of fungicide variable had a significant and positive impact on spring-wheat yields but did not account for rainfall or relative humidity within the crop canopy which are both important components of the decision to apply fungicide. Additional data on rainfall and other climatic variables, more spatially explicit than at the RM level would benefit the yield gap analysis in Saskatchewan.

Overall, the application of SCIC data to analyze management practices impacting the yield gap has proved to be extremely valuable, especially in the context of fertilizer use in western Canada. Extending this framework across the rest of the SCIC risk zones in coordination with AAFC buffer zones would continue to increase the yield gap analysis in western Canada. Additionally, increased data collection by SCIC, specifically for management practices would strengthen any future uses of this data and the implementation of the crop insurance program in Saskatchewan with linkages between management and insurance premiums.

7.5 Policy Implications

The results of this thesis provide several important policy implications for spring-wheat producers in Saskatchewan. Canada has committed to addressing global food and nutritional security through increased food production, while also committing to reducing fertilizer related greenhouse gas (GHG) emissions by 30% over the next 8 years (Government of Canada, 2022). The nature of these two commitments creates a unique challenge for Canadian agriculture producers tasked with increasing production through increased yields while also reducing GHG emissions.

The results also suggest that a portion of the yield gap for wheat in risk zone 16 could be reduced by optimizing nitrogen, specifically on soil groups 1 and 2 and when following pulse crops. This result proposes a significant challenge, as any increases in the costs to use nitrogen fertilizer may reduce yields thus increasing the relatively small, estimated yield gap for spring wheat production in risk zone 16, in Saskatchewan. An important observation to note from the data is that while nitrogen use is not observed to be increasing significantly over the data period from 2011 to 2019, observed average spring wheat yields over the same time period appear to be increasing. This observation requires more research, as the factors that are driving this increase in yields may become increasingly important for producers if production is further demanded yet fertilizer is further disincentivized.

Further research should also focus on spring wheat variety adoption in Saskatchewan. As the results suggest, using varieties with a higher yield index results in higher yields. Encouraging producers to adopt newer and higher yielding varieties could have a significant impact on the yield gap for spring wheat in Saskatchewan. Furthermore, research on how different varieties impact the yield response to inputs, such as nutrient-use efficiency could have positive benefits for closing the yield gap and reducing fertilizer related emissions jointly. SCIC could play a role in not only identifying local best management practices (BMP's) given existing data but also in encouraging the adoption of such BMPs through the crop insurance program such as increased coverage or reduced premiums for adopting. It is important; however, that the program remain actuarially sound for all producers.

7.6 Study Limitations

The framework for analysis drew largely from fertilizer-use profitability studies. While this approach was deemed appropriate given the definition of the economically optimal nitrogen rate and economically exploitable yield, a major limitation was the lack of localized nitrogen and spring-wheat price data. Instead, it is assumed all producers pay and receive the same price for inputs and output when actual prices fluctuate across the year. To address this limitation, an average cost-to-price ratio for 2011 to 2019 was included. Additionally, the estimated marginal response of nitrogen is static in the sense that data from all years is used in estimation. While the estimated model captures variation in the response to applied nitrogen across soil groupings and the fields previous crop, it fails to capture variation in the response to applied nitrogen across time. This adds to the challenge of choosing accurate and reflective cost-to-price ratios for the optimal rate calculation.

This thesis focuses specifically on optimizing applied nitrogen use as it relates to the economically exploitable yield gap for spring wheat production in Saskatchewan. First, this is somewhat of a partial approach given that there are several inputs involved in crop production. While many of these other inputs are controlled for in the model, they are not economically optimized. Therefore, the results pertain only to the nitrogen related yield gap. Secondly, this thesis does not account for any residual nitrogen already present in the soil. While this is not considered a limitation in the case of this study, it is important to note for comparing plot level yield response studies in which residual fertility is commonly measured and accounted for. It is not considered a limitation in this study given that a low percentage of producers have reported to soil test their fields regularly, raising the question if they are in fact accounting for it in their decision-making.

The same comparison challenge can be said for the included soil productivity rating variable used in this study. While the SCIC-developed soil productivity rating provides an aggregated measure of several factors that characterize a field's expected productivity including soil characteristics, its use limits the application of this approach beyond Saskatchewan. As discussed in section 7.4.1, further research on unique producer characteristics and practices could include gathering information on soil testing behaviour and soil characteristics. Lastly, this study focuses on the relationship between nitrogen use and spring wheat yields but does not account for grain protein content, which is an important component of nutrient use decisions in wheat production. Due to a lack of data regarding protein content at the field level, protein levels and the corresponding protein premium or discounts was excluded from this analysis.

Lastly, the results of this study are specific to risk zone 16 and therefore not considered directly applicable or comparable to other risk zones. As mentioned in the discussion for areas for further research, expanding this empirical approach to other risk zones could provide valuable insight into the yield gap for wheat across Saskatchewan.

7.7 Study Conclusions

This thesis finds that a gap does exist between observed average nitrogen use levels and the estimated economically optimal nitrogen rates for spring-wheat production in Saskatchewan. The majority of observed nitrogen applications under various cost-price ratio scenarios, and soil group and previous crop combinations, were at suboptimal levels to what was estimated to be economically optimal. Average observed nitrogen rates were greatest on the least productive soils compared to soil group 1. Therefore, agronomic recommendations should account for the fields unique yield potential and expected response to nitrogen, rather than solely the nutrient requirements to reach a specific yield target. Additionally, producers on average, were observed underapplying nitrogen on more productive soils and when following a pulse crop and were observed overapplying nitrogen on the least productive soils, compared to what was determined to be economically optimal. A portion of the yield gap for wheat in risk zone 16 could be reduced by increasing nitrogen rates to the estimated optimal rate on soil groups 1 and 2 and when following pulse crops and oilseed crops.

Works Cited

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When Should You Adjust Standard Errors for Clustering? *Quarterly Journal of Economics*, *138*(1), 1–35. https://doi.org/10.1093/qje/qjac038

Alexandratos N. and J. Bruinsma. (2012). WORLD AGRICULTURE TOWARDS 2030 / 2050 The 2012 Revision PROOF COPY. *ESA Working Paper*, *12*(03), 146.

Arcand, M. M., Lemke, R., Farrell, R. E., & Knight, J. D. (2014). Nitrogen supply from belowground residues of lentil and wheat to a subsequent wheat crop. *Biology and Fertility of Soils*, *50*(3), 507–515. https://doi.org/10.1007/s00374-013-0873-8

Beddow, J., Hurley, T., Pardey, P., & Alston, J. (2015). Rethinking yield gaps. *Staff Papers*, (April), 1–44.

Berck, P., & Helfand, G. (1990). Reconciling the von Liebig and Differentiable Crop Production Functions. *American Journal of Agricultural Economics*, *72*(4), 985–996. https://doi.org/10.2307/1242630

Beres, B. L. (2022). Wheat yield gaps in the Canadian Prairies: magnitude, implications, and opportunities. *Top Crop Summit 2022 – Western Summit*.

Beza, E., Silva, J. V., Kooistra, L., & Reidsma, P. (2017). Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, *82*(May 2019), 206–222. https://doi.org/10.1016/j.eja.2016.06.016

Canada, S. (2019). Consumer Price Index. Retrieved from https://www.statcan.gc.ca/en/subjects-start/prices_and_price_indexes/consumer_price_indexes

Cassman, K. G., Dobermann, A., Walters, D. T., & Yang, H. (2003). Meeting cereal demand while protecting natural resources and improving environmental quality. *Annual Review of Environment and Resources*, *28*, 315–358. https://doi.org/10.1146/annurev.energy.28.040202.122858

Catellier, C. (2022). Evaluating the interaction of management and environment on crop production in western Canada using producer-reported data. Retrieved from https://wgrf.ca/projects/evaluating-the-interaction-of-management-and-environment-on-crop-production-in-western-canada-using-producer-reported-data/

Cereals Canada. (2022). Wheat. Retrieved from https://cerealscanada.ca/wheat/#:~:text=CWRS-,Canada Western Red Spring,noodles and Asian product applications.

Chapagain, T., & Good, A. (2015). Yield and production gaps in rainfed wheat, barley, and canola in Alberta. *Frontiers in Plant Science*, *6*(NOVEMBER), 1–10. https://doi.org/10.3389/fpls.2015.00990

Cruppe, G., DeWolf, E., Jaenisch, B. R., Andersen Onofre, K., Valent, B., Fritz, A. K., &

Lollato, R. P. (2021). Experimental and producer-reported data quantify the value of foliar fungicide to winter wheat and its dependency on genotype and environment in the U.S. central Great Plains. *Field Crops Research*, *273*(April), 108300. https://doi.org/10.1016/j.fcr.2021.108300

Dhakal, C., & Lange, K. (2021). Crop yield response functions in nutrient application: A review. *Agronomy Journal*, *113*(6), 5222–5234. https://doi.org/10.1002/agj2.20863

Dumanski, J., Cann, M. and Wolynetz, M. S. (1992). *Identification of crop production risk areas in Manitoba based on agroecological resources.*

Evans, L. T., & Fisher, R. A. (1999). Yield potential: Its definition, measurement, and significance. *Crop Science*, *39*(6), 1544–1551. https://doi.org/10.2135/cropsci1999.3961544x

Farrell, R. (2020). Environmental and Agronomic Benefits of Enhanced Efficiency Nitrogen Fertilizers. *2020 Soils and Crops Workshop*.

Fertilizer Canada. (2022). Fertilizer Use Survey. Retrieved from https://fertilizercanada.ca/wp-content/uploads/2022/08/SPARK-FERTILIZER-USE-IN-CANADA-REPORT-2022-VF_08_04_2022.pdf

Fischer, R. a, Byerlee, D., & Edmeades, G. O. (2009). Can technology deliver on the yield challenge to 2050? *FAO: Expert Meeting on How to Feed the World in 2050*, *2050*(June), 1–48. Retrieved from ftp://ftp.fao.org/docrep/fao/012/ak977e/ak977e00.pdf

Fischer, T., Byerlee, D., & Edmeades, G. (2014). Crop yields and global food security: will copyright Act 1968 yield increase continue to feed the world? *Australian Centre for International Agricultural Research*, 634.

Global Yield Gap Atlas. (n.d.). Common definitions of terms frequently used in GYGA. Retrieved from https://www.yieldgap.org/web/guest/glossary

Gómez-Limón, J. A., Riesgo, L., & Arriaza, M. (2004). Multi-criteria analysis of input use in agriculture. *Journal of Agricultural Economics*, *55*(3), 541–564. https://doi.org/10.1111/j.1477-9552.2004.tb00114.x

Gomez, K. (1977). On-Farm Assessment of Yield Constraints: Methodological Problems. *International Rice Research Institute Repository*, 1–17.

Government of Canada. (2022). Canada's fertilizer emissions reduction target. Retrieved from https://agriculture.canada.ca/en/department/transparency/public-opinion-research-consultations/share-ideas-fertilizer-emissions-reduction-target/canadas-fertilizer-emissions-reduction-target

Government of Saskatchewan. (2021). Phosphorous Fertilization in Crop Production. Retrieved from https://www.saskatchewan.ca/business/agriculture-natural-resources-and-industry/agribusiness-farmers-and-ranchers/crops-and-irrigation/soils-fertility-and-nutrients/phosphorus-fertilization-in-crop-production

Griffin, R. C., Montgomery, J. M., & Rister, E. (1987). Production , Technology and Cost Functions (Translog Production Function and Estimation of Efficiency). *Western Journal of Agricultural Economics, 12*(2), 460–468. Retrieved from http://www.luc.edu/orgs/meea/volume7/khalil.pdf%0Ahttp://dx.doi.org/10.1016/j.energy.2016.12.119%0Ahttp://dx.doi.org/10.1016/j.rser.2015.12.160%0Ahttps://academic.oup.com/restud/article-lookup/doi/10.2307/2296293%0Ahttp://link.springer.com/10.1007/978-1-3

Griliches, Z., & Mairesse, J. (1995). *Production Functions: The Search for Identification*.

Hatfield, J. L., & Beres, B. L. (2019). Yield Gaps in Wheat: Path to Enhancing Productivity. *Frontiers in Plant Science*, *10*(December), 1–7. https://doi.org/10.3389/fpls.2019.01603

Horowitz, J. K., & Lichtenberg, E. (1993). Insurance, Moral Hazard, and Chemical Use in Agriculture. *American Journal of Agricultural Economics*, *75*(4), 926–935. https://doi.org/10.2307/1243980

Kelly, V., Adesina, A. A., & Gordon, A. (2003). Expanding access to agricultural inputs in Africa: A review of recent market development experience. *Food Policy*, *28*(4), 379–404. https://doi.org/10.1016/j.foodpol.2003.08.006

Kim, K. Il, Petrin, A., & Song, S. (2016). Estimating production functions with control functions when capital is measured with error. *Journal of Econometrics*, *190*(2), 267–279. https://doi.org/10.1016/j.jeconom.2015.06.016

Levinsohn, J., & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, *70*(2), 317–341. https://doi.org/10.1111/1467-937X.00246

Licker, R., Johnston, M., Foley, J. A., Barford, C., Kucharik, C. J., Monfreda, C., & Ramankutty, N. (2010). Mind the gap: How do climate and agricultural management explain the "yield gap" of croplands around the world? *Global Ecology and Biogeography*, *19*(6), 769–782. https://doi.org/10.1111/j.1466-8238.2010.00563.x

Lobell, D. B., Cassman, K. G., & Field, C. B. (2009). Crop Yield Gaps: Their Importance, Magnitudes, and Causes. *Annual Review of Environment and Resources*, *34*(1), 179–204. https://doi.org/10.1146/annurev.environ.041008.093740

Marschak, J., & Andrews, W. (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica*, *12*(3), 143–205.

Meijer, S. S., Catacutan, D., Ajayi, O. C., Sileshi, G. W., & Nieuwenhuis, M. (2015). The role of knowledge, attitudes and perceptions in the uptake of agricultural and agroforestry innovations among smallholder farmers in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, *13*(1), 40–54. https://doi.org/10.1080/14735903.2014.912493

Miller, P. R., Gan, Y., McConkey, B. G., & McDonald, C. L. (2003). Pulse crops for the northern Great Plains: II. Cropping sequence effects on cereal, oilseed, and pulse crops. *Agronomy Journal*, *95*(4), 980–986. https://doi.org/10.2134/agronj2003.9800

Monfreda, C., Ramankutty, N., & Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(1), 1–19. https://doi.org/10.1029/2007GB002947

Monjardino, M., Hochman, Z., & Horan, H. (2019). Yield potential determines Australian wheat growers' capacityto close yield gaps while mitigating economic risk. *Agronomy for Sustainable Development*, *39*(6). https://doi.org/10.1007/s13593-019-0595-x

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., & Foley, J. A. (2012). Closing yield gaps through nutrient and water management. *Nature*, *490*(7419), 254–257. https://doi.org/10.1038/nature11420

Neumann, K., Verburg, P. H., Stehfest, E., & Müller, C. (2010). The yield gap of global grain production: A spatial analysis. *Agricultural Systems*, *103*(5), 316–326. https://doi.org/10.1016/j.agsy.2010.02.004

Olley, G. S., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, *64*(6), 1263. https://doi.org/10.2307/2171831

Rajsic, P., & Weersink, A. (2008). Do farmers waste fertilizer? A comparison of ex post optimal nitrogen rates and ex ante recommendations by model, site and year. *Agricultural Systems*, *97*(1–2), 56–67. https://doi.org/10.1016/j.agsy.2007.12.001

Reader, M. A., Revoredo-Giha, C., Lawrence, R. J., Hodge, I. D., & Lang, B. G. (2018). Farmers' spending on variable inputs tends to maximise crop yields, not profit. *International Journal of Agricultural Management*, *7*(1), 30–40. https://doi.org/10.5836/ijam/2018-07-30

Robertson, G. (1974). Wheat Yields for 50 Years at Swift Current, Saskatchewan in Relation to Weather. *Canadian Journal of Plant Science*, *54*, 625–650.

Saskatchewan Ministry of Agriculture. (2022). Saskatchewan Crop planning Guide Archive. Retrieved from https://publications.saskatchewan.ca/#/categories/1465

Sheahan, M., Black, R., & Jayne, T. S. (2012). Are Farmers Under-Utilizing Fertilizer? Evidence from Kenya. *International Association of Agricultural Economist (IAAE) Triennial Conference*, (January), 5.

Shiferaw, B., Smale, M., Braun, H. J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security*, *5*(3), 291–317. https://doi.org/10.1007/s12571-013-0263-y

Snyder, K. A., Miththapala, S., Sommer, R., & Braslow, J. (2017). THE YIELD GAP:

CLOSING the GAP by WIDENING the APPROACH. *Experimental Agriculture*, *53*(3), 445–459. https://doi.org/10.1017/S0014479716000508

Stevenson, F. C., & Van Kessel, C. (1996). The nitrogen and non-nitrogen rotation benefits of pea to succeeding crops. *Canadian Journal of Plant Science*, *76*(4), 735–745. https://doi.org/10.4141/cjps96-126

Syme, J., An, H., & Torshizi, M. (2023). Estimating the effect of time-invariant characteristics in panel data: wheat adoption in Western Canada. *American Journal of Agricultural Economics*, (March 2022). https://doi.org/10.1111/ajae.12400

Tong, Q., Swallow, B., Zhang, L., & Zhang, J. (2019). The roles of risk aversion and climate-smart agriculture in climate risk management: Evidence from rice production in the Jianghan Plain, China. *Climate Risk Management*, *26*(November), 100199. https://doi.org/10.1016/j.crm.2019.100199

van Dijk, M., Morley, T., Jongeneel, R., van Ittersum, M., Reidsma, P., & Ruben, R. (2017). Disentangling agronomic and economic yield gaps: An integrated framework and application. *Agricultural Systems*, *154*(February), 90–99. https://doi.org/10.1016/j.agsy.2017.03.004

Van Ittersum, M. K., & Rabbinge, R. (1997). Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crops Research*, *52*(3), 197–208. https://doi.org/10.1016/S0378-4290(97)00037-3

Van Ittersum, Martin K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance-A review. *Field Crops Research*, *143*, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Waddington, S. R., Li, X., Dixon, J., Hyman, G., & de Vicente, M. C. (2010). Getting the focus right: Production constraints for six major food crops in Asian and African farming systems. *Food Security*, *2*(1), 27–48. https://doi.org/10.1007/s12571-010-0053-8

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (Vol. 58). https://doi.org/10.1515/humr.2003.021

APPENDIX A

Table A1: Alternative Estimated Production Function Models.

	Full Model as seen in Chapter 5	No Fixed Effects	Fixed Effects for Producer ID and Year	Fixed Effects for Producer ID, Year, and Soil
Variable	**Estimate**	**Estimate**	**Estimate**	**Estimate**
Intercept	N/A	-1356 *** (127.3)	N/A	N/A
Nitrogen	13.75 *** (2.90)	16.04 *** (1.43)	7.65 *** (2.28)	5.61 ** (2.16)
Nitrogen^2	-0.06 *** (0.02)	-0.08 *** (0.01)	-0.03 ** (0.02)	-0.03 ** (0.02)
Phosphorous	0.36 (1.99)	1.64 * (0.84)	2.76 ** (1.4)	2.56 * (1.39)
Potassium	10.16 * (4.02)	11.17 *** (1.99)	6.69 ** (3.4)	6.4 * (1.39)
Sulphur	-0.17 (2.72)	-0.2 (1.38)	0.2 (1.98)	0.14 (1.92)
Seeding Date	1.92 (2.79)	4.69 ** (1.49)	5.64 ** (2.07)	5.2 ** (2.1)
Seeding Date^2	-0.22 * (0.13)	-0.26 *** (0.07)	-0.36 *** (0.08)	-0.36 *** (0.08)
Management Index	1,199 *** (139.6)	1,128 *** (56.8)	N/A	N/A
Variety Yield Index	1,740 *** (159.4)	2,084 *** (74.5)	1,273 *** (115.8)	1,277 *** (115.6)
Topsoil Moisture Short (=1)	57.98 (69.71)	13.83 (20.5)	17.4 (37.3)	14.68 (37.3)
Soil Group 2 (=1)	-85.17 (72.27)	-54.5 (53.1)	-17.1 (53.3)	N/A
Soil Group 3 (=1)	-144.89 * (87.79)	-170.6 (54.5)	24.36 (63.14)	N/A
Rainfall	6.42 *** (1.34)	4.15 *** (0.60)	6.62 *** (0.63)	6.48 *** (0.62)
Rainfall^2	-0.01 *** (0.002)	-0.007 *** (0.0009)	-0.009 *** (0.0009)	-0.008 *** (0.001)
Fungicide (=1)	295 *** (38.7)	359.5 *** (18.9)	220.4 *** (23.5)	222 *** (23.3)
Previous Crop Oilseed (=1)	-124.9 (97.2)	-109.9 ** (51.8)	-177.3 *** (48.7)	-178.9 *** (48.3)

Previous Crop Pulse (=1)	-41.2 (83.2)	-112.1 ** (54.6)	-125.2 ** (57.6)	-126.3 ** (57.03)
Nitrogen × Previous Crop Oilseed	2.15 (1.57)	1.72 ** (0.79)	3.03 *** (0.75)	3.01 *** (0.74)
Nitrogen × Previous Crop Pulse	2.26 * (1.35)	3.13 *** (0.87)	3.34 *** (0.95)	3.4 *** (0.94)
Nitrogen × Soil Group 2	-1.39 (1.14)	-1.63 * (0.84)	-1.72 ** (0.83)	N/A
Nitrogen × Soil Group 3	-2.99 ** (1.34)	-2.21 ** (0.85)	-4.2 *** (0.97)	N/A
Observations	7,537	7,537	7,537	7,537
Adj. R^2	0.451	0.377	0.616	0.621
Within R^2	0.351		0.097	0.080
Standard Errors	Clustered (Producer ID)	Hetero-Robust	Hetero-Robust	Hetero-Robust
Fixed Effects	Year	None	Producer ID (723), Year	Producer ID, Year, Soil (13)
Significance Codes	***: 0.01, **: 0.05, *: 0.1			

Source: Author’s calculations.

As can be seen in Table A1, the inclusion of fixed effects for producer ID significantly lowers the estimated coefficients for the nitrogen related terms and many of the management characteristics including rainfall, fungicide, and variety index. However, as previously discussed the producer ID fixed effects models may lack necessary variation given a lack of variation in nitrogen rates within producers.

APPENDIX B

Figure B.1: AAFC Buffer Zone Map for Saskatchewan.

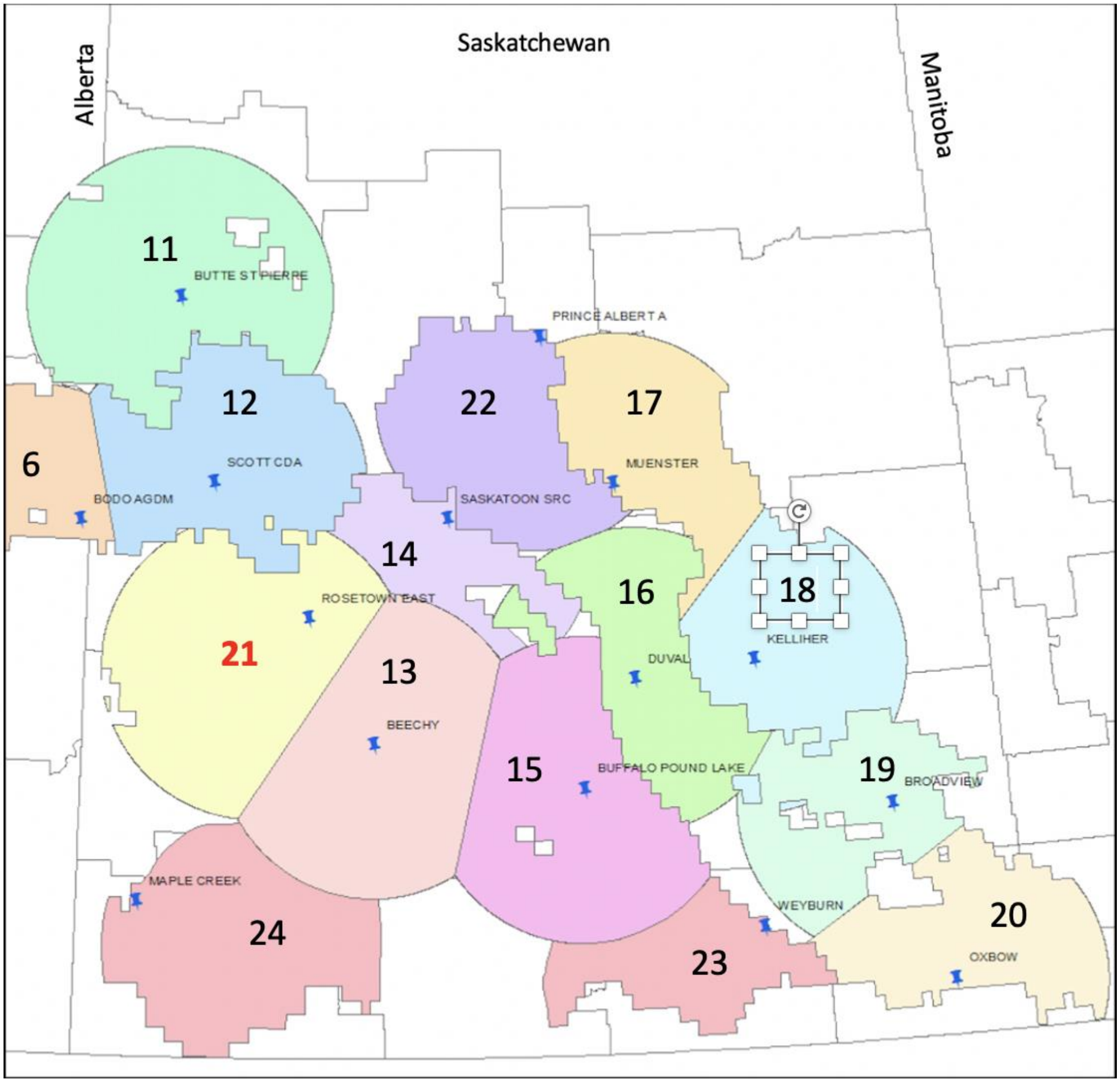


Image source: (Beres, 2022).